Detecting Flow Experiences in Cognitive Tasks
A Neurophysiological Approach

Zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaften

(Dr. rer. pol.)

von der Fakultät für
Wirtschaftswissenschaften
am Karlsruher Institut für Technologie (KIT)

genehmigte
DISSERTATION

von
Michael Thomas Knierim (M.Sc.)

Tag der mündlichen Prüfung: 18.08.2020
Referent: Prof. Dr. Christof Weinhardt
Korreferent: Prof. Dr. Marc T. P. Adam

Karlsruhe, 2020

KARLSRUHE INSTITUTE OF TECHNOLOGY

Abstract

Department of Economics and Management
Institute of Information Systems and Marketing

Dr. rer. pol.

Detecting Flow Experiences in Cognitive Tasks - A Neurophysiological Approach

by Michael Thomas KNIERIM

The experience of flow is a unique sensation of complete task absorption and effortless action that is highlighted as a correlate of peak performances, personal and social growth, and general well-being. For organisations, higher flow frequencies, therefore, relate to a more engaged, skilled, and productive workforce. Especially as global phenomena like increasing knowledge work demand and low worker engagement are developing, organisations could strongly benefit from fostering workers' flow experiences. However, facilitating flow represents a substantial challenge due to the variety of workers' abilities, tasks and workplace configurations. Knowledge workers are faced with unstructured and complex tasks, that require numerous domain-specific abilities and cooperation with others. Workplaces are diversifying with boundaries disappearing between centralized and digitally-mediated workspaces. This variety means that only person-, task- and situation-independent approaches can deliver comprehensive flow support. For this reason, research on the experiences neurophysiological basis is increasingly pursued. On this basis, adaptive Neuro-Information Systems (NeuroIS) could be developed that are able to detect flow continuously (especially through wearable sensor systems), and that can provide flow-supporting mechanisms. Presently, despite these efforts, the knowledge on how to detect flow with neurophysiological measures is sparse, highly fragmented, and lacks experimental variety. On the individual level, competing propositions exist that have not been consolidated through cross-situational, and multi-sensor observation. On the group level, almost no research has been conducted to investigate neurophysiological correlates in social interactions, particularly not in digitally-mediated interactions. This dissertation addresses these gaps through the cross-situational observation of flow using wearable ECG and EEG sensor systems. In doing so, limitations in the present state of experimental flow research are addressed that refer to central shortcomings of established paradigms for the controlled elicitation of flow experiences. Specifically, two experiments are conducted with manipulations of difficulty, naturalism, autonomy, and social interaction to investigate the question of how flow elicitation can be intensified, and the experience detected more robustly across situations. These investigations are based on an extensive integration of the theoretic and empiric literature on flow neurophysiology. Altogether, the results suggest flow to be represented by moderate physiological activation and mental workload, by increased attentional task engagement and by affective neutrality. Especially EEG features indicate a diagnostic potential to separate lower from higher flow intensities by the reflection of optimal and non-optimal (individual and group) task difficulties. To catalyse, that the positive promises of fostering flow in individuals and social units, can be realised, avenues to advance flow facilitation research are outlined.

KARLSRUHE INSTITUTE OF TECHNOLOGY

Zusammenfassung

Fakultät für Wirtschaftswissenschaften
Institut für Wirtschaftsinformatik und Marketing

Dr. rer. pol.

Detecting Flow Experiences in Cognitive Tasks - A Neurophysiological Approach

von Michael Thomas KNIERIM

Das Flow-Erlebnis beschreibt einen Zustand vollständiger Aufgabenvertiefung und mühelosen Handelns, der mit Höchstleistungen, persönlichem Wachstum, sowie allgemeinem Wohlbefinden verbunden ist. Für Unternehmen stellen häufigere Flow-Erlebnisse der ArbeitnehmerInnen daher auch eine produktivitäts- und zufriedenheitsfördernde Basis dar. Vor allem da sich aktuell globale Phänomene wie die steigende Nachfrage nach Wissensarbeit und das niedrige Arbeitsengagement zuspitzen, können Unternehmen von einer Förderung von Flow profitieren. Die Unterstützung von Flow stellt allerdings aufgrund der Vielfalt von Arbeitnehmerfertigkeiten, -aufgaben, und -arbeitsplätzen eine komplexe Herausforderung dar. WissensarbeiterInnen stehen dynamischen Aufgaben gegenüber, die diverse Kompetenzen und die Kooperation mit anderen erfordern. Arbeitsplätze werden vielseitiger, indem die Grenzen zwischen ko-präsenten und virtuellen Interaktionen verschwinden. Diese Vielfalt bedeutet, dass eine solide Flow-Förderung nur durch personen-, aufgaben- und situationsunabhängige Ansätze erfolgen kann. Aus diesem Grund werden zunehmend die neurophysiologischen Grundlagen des Flow-Erlebens untersucht. Auf deren Basis könnten adaptive Neuro-Informationssysteme entwickelt werden, die mittels tragbarer Sensorik Flow kontinuierlich erkennen und fördern können. Diese Wissensbasis ist bislang jedoch nur spärlich und in stark fragmentierter Form vorhanden. Für das Individuum existieren lediglich konkurrierende Vorschläge, die noch nicht durch situations- und sensorübergreifende Studien konsolidiert wurden. Für Gruppen existiert noch fast keine Forschung zu neurophysiologischen Flow-Korrelaten, insbesondere keine im Kontext digital-mediierter Interaktionen. In dieser Dissertation werden genau diese Forschungslücken durch die situationsübergreifende Beobachtung von Flow mit tragbaren EKG und EEG Sensoren adressiert. Dabei werden zentrale Grenzen der experimentellen Flow-Forschung berücksichtigt, vor allem die Defizite etablierter Paradigmen zum kontrollierten Hervorrufen von Flow. Indem Erlebnisse in zwei kognitiven Aufgaben und mehreren Manipulationen (von Schwierigkeit, Natürlichkeit, Autonomie und sozialer Interaktion) variiert werden, wird untersucht, wie Flow intensiver hervorgerufen und wie das Erlebnis stabiler über Situationen hinweg beobachtet werden kann. Die Studienergebnisse deuten dabei insgesamt auf ein Flow-Muster von moderater physiologischer Aktivierung und mentaler Arbeitslast, von erhöhter, aufgabenorientierter Aufmerksamkeit und von affektiver Neutralität hin. Vor allem die EEG Daten zeigen ein diagnostisches Potenzial, schwächere von stärkeren Flow-Zuständen unterscheiden zu können, indem optimale und nicht-optimale Aufgabenschwierigkeiten (für Individuen und Gruppen) erkannt werden. Um das Flow-Erleben weiter zu fördern, werden geeignete Wege für zukünftige Forschung abschließend diskutiert.

Acknowledgements

To use the terminology from flow theory: completing this dissertation was a continuous stream of challenges. I would like to say that it was mostly a series of rewarding flow experiences too, but that would overly romanticize the work. Some of the challenges were immensely rewarding, and some very frustrating. For the exciting times and for the other times, I am glad to have been surrounded by very knowledgeable, kind and supportive people. To them, I would like to say thank you here.

First of all, I would like to thank Prof. Dr. Christof Weinhardt for the opportunity to pursue my dissertation under his supervision. I would also like to thank my co-advisor Prof. Dr. Marc Adam and the other members of the PhD committee, Prof. Dr. Petra Nieken and Prof. Dr. Wolf Fichtner, for their time in reviewing the work at hand, for their helpful comments and the thorough and insightful discussions.

For his meticulous and tireless effort to discuss experiments, revise articles, brainstorm wild ideas and mostly, for believing in me and having my back, I want to thank Mario. Without your effort and support, I might not have completed this project. For his insightful remarks, his open ear in all matters related to science or else, I want to thank Simon. I could think of no better colleague to start the PhD journey. For his kindness, inspiring decisiveness and fighting spirit, I want to thank David for his help in dealing with the occasional hardship and for the uplifting consequences of his words and actions. I want to thank the people that sat with me in long lunch sessions and extended office coffee breaks or that shared conference travels and hikes. These are times that I remember very fondly. The times when we could talk intricately about all pressing and not so pressing matters. Ewa, Carola, Florian, Tim, Anke, Peyman, Claudia, Magie, thank you! This list is way to short as there have been many more people that I have had the fortune to get to know, to talk with and to learn from. Verena, Anuja, Raphael, Benedikt, Dominik, Claudia, Marc, Bent, Greta, Christian, Nico, Max, Miguel, Ivo, Celina, Sebastian, Pascal, Felix, and Balthasar you all shaped my PhD years into experiences that I don't want to miss.

A thesis does not only happen at work. It has, at many points, led to emotional spillovers in my personal life, and I am very grateful to have had the loving tolerance and support from my partner, friends, and family. I want to say thank you to my parents, who never had a single doubt about my aptitude to complete this dissertation. You gave me freedom, courage and peace of mind to pursue this challenging task. My brothers David and Matthias: thank you for being a safe and joyful haven to return to and for making me feel like I am a bit of an annoying know-it-all. I want to thank my good friends Stefan and Lena. You shared a lot of the experience that I went through and the discussions we have had were a great liberation from being stuck for too long. The one person that I want to mention last - because this is what people remember - is Viktoria. You probably experienced my highs and lows most vividly. I know that you have had to develop patience with me and I am very grateful that you did. With you by my side, I knew I had a partner to rely on and to grow with and I am happy that we could experience so many wonderful things during these years. Thank you!

Lastly, thank you to everyone that I might have forgotten to name. Thank you, everyone, who shared time, advice, a triumph or a shoulder.

Contents

List of Figures

List of Tables

List of Abbreviations

ACC	Anterior Cingulate Cortex
AF	AF3 & AF4
AI	Artificial Intelligence
AIC	Akaike Information Criterion
AM	Autonomy Manipulation
ANOVA	Analysis of Variance
ANS	Autonomous Nervous System
AUTO	Self-Selected Optimal Difficulty
BH	Benjamini-Hochberg
BCI	Brain Computer Interface
BIC	Bayesian Information Criterion
BS	Between Subject
cEEGrid	c-Shaped Electro-Encephalpography Grid
cESM	controlled Experience Sampling Method
CA	Cognitive Absorption
CAL	Calibrated Optimal Difficulty
CI	Confidence Interval
CNS	Central Nervous System
COM	Communication Manipulation
CS	Corrugator Supercilii
CPM	Cooperation Manipulation
DM	Difficulty Manipulation
DMC	Digitally-Mediated Communication
DMN	Default Mode Network
DV	Dependent Variable
ENG	Engagement
EASY	Low Difficulty

ECG Electrocardiography

EDA Electrodermal Activation

EEG Electroencephalography

ECM Estimated Conditional Means

EMG Electromyography

ERP Event Related Potential

ESM Experience Sampling Method

GEQ Game Experience Questionnaire

GG Greenhouse-Geisser

F2F Face-to-face

FAA Frontal Alpha Asymmetry

FB Frequency Band

FC FC5 & FC6

FFT Fast Fourier Transformation

fGEQ14 Flow Dimension of the Short Game Experience Questionnaire

fGEQ36 Flow Dimension of the Long Game Experience Questionnaire

FI Flow Index Scale

FKS Flow Kurzskala

F-L F7 & F8

F-M F3 & F4

fMRI functional Magnetic Resonance Imaging

fNIRS functional Near-Infrared Spectroscopy

FOI Features of Interest

FS Feature Selection

FSS Flow Short Scale

FWHM Full-Width-At-Half-Maximum of Pulse Pressure

HARD High Difficulty

HF-HRV High Frequency Heart Rate Variability

HR Heart Rate

HRV Heart Rate Variability

IAF Individualized Alpha Frequency

IBI Inter-Beat Interval

ICA Independent Component Analysis

ICC Intra-Class Correlation Coefficient

IDPM Interdependence Manipulation

IQR Interquartile Range

IS Information Systems

IT Information Technology

IV Independent Variable

KIT Karlsruhe Institute of Technology

KW Knowledge Work

LH Left Hemisphere

LF-HRV Low Frequency Heart Rate Variability

LM Linear Regression Model

LMM Linear Mixed Model

LOESS Locally Estimated Scatterplot Smoothing

MAS Mastery

mHR Mean Average Heart Rate

ML Machine Learning

MP Multi Person Condition

MWT Morlet Wavelet Transformation

NeuroIS Neuro-Information Systems

O O1 & O2

OO Orbicularis Oculi

P P7 & P8

PFC Prefrontal Cortex

PNS Peripheral Nervous System

PNN50 Percent of Adjacent NN Intervals not Differing More than 50ms

PRNG Pulse Range

PSD Power Spectral Density

PWA Pulse Wave Amplitude

RCY Respiratory Cycle

aRD	Abdominal Respiratory Depth
tRD	Thoracic Respiratory Depth
RG	Research Goal
RH	Right Hemisphere
RM	Repeated Measures
RMSSD	Root Mean Square of Successive Differences
ROI	Region of Interest
RQ	Research Question
RR	Respiratory Rate
SAM	Self Assessment Manikin
SCL	Skin Conductance Level
SCM	Social Context Manipulation
sFSS	Short Version of the Flow Short Scale
SIM	Statistical Inferential Modelling
SLR	Structured Literature Review
SD	Standard Deviation
SDNN	Standard Deviation of NN-Intervals
SE	Standard Error
SP	Single Person Condition
ST	Synchronization Theory
T	T7 & T8
TD	Time Dynamics
THT	Transient Hypofrontality Theory
VHF	Very High Frequency
VLF	Very Low Frequency
W1	Writing Round 1
W2	Writing Round 2
W3	Writing Round 3
WM	Working Memory
WS	Within-Subject
WOLF	WOrk-reLated Flow Inventory
ZM	Zygomaticus Majors

Chapter 1

Introduction

1.1 Motivation

Flow is described as a unique experience of complete task immersion, in which action and awareness merge, concentration feels effortless, and that is accompanied by peak performances and exhilarating satisfaction (Csikszentmihalyi, 1975). Repeated flow experience has been related to individual and collective benefits, as flow is strongly linked to general well-being in life, and the strengthening of social relationships (Keeler et al., 2015; Tse, Nakamura, and Csikszentmihalyi, 2020). In the work domain, flow experiences have been associated with better job performances (through increased productivity and creativity) and more worker satisfaction, leading to reduced employee turnover and shielding from burnout (Fullagar and Delle Fave, 2017; Yotsidi et al., 2018). Similarly, research on the social dimension of flow at work highlights the positive links of flow to workgroup performances, interaction satisfaction and collective efficacy development (Keith et al., 2016; Zumeta et al., 2016; de Moura Jr and Bellini, 2019). Due to these numerous beneficial relationships of flow to life and work experiences, the facilitation of flow represents a desirable goal for scholars and organisational practitioners (de Moura Jr and Bellini, 2019). At the same time, global polls indicate a severe lack of engagement of workers across professions (Gallup, 2017; Parent-Thirion et al., 2015). As flow is considered as a primordial instance of deep (task or profession) engagement (Tse, Nakamura, and Csikszentmihalyi, 2020; de Moura Jr and Bellini, 2019), and as these polls similarly highlight the benefit of engagement for worker productivity and well-being, the need to foster flow at work is emphasised further. Therefore, the research on flow in the workplace has seen increasing attention in recent years (Fullagar and Delle Fave, 2017; de Moura Jr and Bellini, 2019). However, the facilitation of these experiences still represents a significant challenge.

The situational requirements for flow are complex and rooted in the cognitive-affective dynamics of the individual. The facilitation of concentration, alertness and recovery, the shielding from self-criticism and the balance of workload, are amongst some of the flow requirements that are difficult to manage in today's hectic workplaces (Ceja and Navarro, 2012; Spurlin and Csikszentmihalyi, 2017; Peifer et al., 2019). Specifically, in these workplaces, a vast diversity of tasks and configurations is present that poses significant challenges to the structured facilitation of flow. Demands for unstructured Knowledge Work (KW) are growing, due to the rise of Artificial Intelligences (AI) that are replacing the repetitive work in sales, administrative support, or service tasks (Frey and Osborne, 2017). Additional developments including flat hierarchies, self-directed work and job-crafting further extend this KW complexity (Spurlin and Csikszentmihalyi, 2017; Bakker and Woerkom, 2017). These developments require increased individual expertise and entrepreneurship, and the utilization of small groups (Wuchty, Jones, and Uzzi, 2007; Keith et al., 2016). Also, more complex work environments like as open offices and digitally-mediated collaboration are pervading the KW domain and are supposed to facilitate worker interaction and flexibility (Spurlin and Csikszentmihalyi, 2017). However, these trends are accompanied by problematic phenomena such as information overload (e.g. through high frequencies of electronic messaging), or increases in professional ambiguities due to requirements of more self-organisation (Bakker and Woerkom, 2017). These phenomena stand in contrast to flow experience requirements as they represent attention-competing stimuli, unclear goals, a lack of feedback, and the elicitation of frustration or anxiety (Spurlin and Csikszentmihalyi, 2017; Bakker and Woerkom, 2017). Similarly, social interactions at work have become more complex. Concepts such as open offices and virtually distributed teams bring opportunities to communication, but also impede senses of social connectedness through the alteration of the communication of social information (Derks, Fischer, and Bos, 2008; Chanel and Mühl, 2015). At the time of writing this introduction, this shift towards decentralised, remote workplace configurations is strikingly emphasised by the quarantine measures due to the global pandemic of SARS-CoV-2. The nature of such digital tools, therefore, alters how flow can be experienced in workgroup interactions, and further complicates the provision of flow facilitation. Altogether these developments mean that comprehensive flow facilitation at work must revolve around a person-, task- and situation-independent approach. One such approach is the development of adaptive Neuro-Information Systems (NeuroIS).

Adaptive NeuroIS leverage the information from (wearable) neurophysiological sensors to provide continuous insight into the presence, intensity, and dynamism of cognitive-affective experiences (Riedl and Léger, 2016; Krol, Haselager, and Zander, 2019; Brouwer et al., 2015). This characteristic means that an adaptive NeuroIS reacts to the user's internal state to assist across various situations. Eventually, systems able to adapt to flow intensities could, for instance, reduce flow interruptions (e.g. by blocking incoming messages - see, e.g. Rissler et al., 2018) or provide feedback information to improve flow regulation (e.g. by adjusting task difficulty, or by optimising arousal levels through bio- or neurofeedback - see Lux et al., 2018; Knierim et al., 2017a). Especially electrophysiological methods, like Electrocardiography (ECG) to observe changes in the heart, or Electroencephalography (EEG) to observe the electrical activity of neuron assemblies in the brain, provide promising means for continuous user state detection (Blankertz et al., 2016; Wascher et al., 2019). These promises are based on the low costs, high portability and high temporal resolution of these measures, which makes them both valuable for fundamental research and the eventual transfer of findings into real-world applications. Yet, despite their appeal, adaptive NeuroIS are still far from real-life applicability (Blankertz et al., 2016; Brouwer et al., 2015). Their limitation is to date based in questions of construct measurement validity (which constructs can be identified, with high specificity, by the used sensors), but also in challenges from measurement sensitivity, objectivity, and reliability. This limitation means that more research is required that bridges fundamental and applied settings to enable and instantiate adaptive NeuroIS. For the facilitation of flow experiences, in particular, more research at these intersections is needed. While survey-based research has identified a variety of individual, group, and organisational variables that influence flow, experimental and neurophysiological flow research are still in early stages. Importantly, despite scholastic efforts, there is no robust method available yet to detect flow experiences continuously (Moneta, 2012; de Moura Jr and Bellini, 2019). Currently, the only way to detect flow is to interrupt someone's task and ask about their current experience - which disrupts the (potential) flow experience altogether (Moneta, 2012; de Moura Jr and Bellini, 2019). Therefore, to enable and instantiate flow-fostering, adaptive NeuroIS, more fundamental knowledge about how to observe flow using wearable neurophysiological sensors needs to be developed.

1.2 Dissertation Research Agenda

The present state of flow neurophysiology research is marked by fragmented findings from paradigmatic silos, especially in terms of task contexts and research designs. Having surveyed the available literature on flow neurophysiology (see Chapter 4), it can be summarised, that there has so far been a rather large absence of the study of flow neurophysiology in KW. Instead, a focus on highly controlled game tasks that are manipulated in difficulty is present (the so-called Difficulty Manipulation - DM paradigm). This focus has led to a problematic common-method bias that limits the transferability of previous findings to the work context. Thus, to provide a foundation for flow-facilitating adaptive NeuroIS, first and foremost, research needs to be advanced on flow in cognitive tasks. Consequently, two central limitations of the Difficulty Manipulation (DM) paradigm need to be overcome. First, this paradigm has been criticised for eliciting only shallow flow experiences (Delle Fave, Massimini, and Bassi, 2011; Hommel, 2010). Second, the focus on single tasks and paradigms limits the reliability of identified neurophysiological patterns for the detection of flow intensities. These biases have led to calls for more creative laboratory research that can intensify the flow elicitation in controlled settings (Harris, Vine, and Wilson, 2017b; Hommel, 2010) and for more cross-situational experiments (Barros et al., 2018; Katahira et al., 2018). In addition, a low degree of integration of research is found (e.g. 80% of flow-related Electroencephalography (EEG) studies cite one or fewer of their related studies, nor theoretic flow neurophysiology work - see Section 4.3). On the individual level, this lack of research integration is a likely reason why to date, the understandings of the neurophysiological configuration during flow experiences are often contradictory and sparse. On the group level, almost no research has been conducted to investigate neurophysiological correlates in social interactions. Research that aims to facilitate flow in the context of KW must, therefore, conduct cross-situational and multi-sensor experiments that are integrated with related work. First such work needs to consolidate an understanding of the individual neurophysiology of flow using wearable sensor systems. Afterwards, given the growing prevalence of decentralised workgroups in KW, these foundations must be extended to the understanding of flow in small groups, particularly in digitally-mediated interaction scenarios. This understanding includes not only how individuals' flow is affected by social interaction but also how group-level flow experiences can be described and detected continuously.

The research in this dissertation aims to provide a foundation for the future development of flow-fostering adaptive NeuroIS by advancing the understanding of how to elicit flow under novel (more and less) controlled settings, and by observing neurophysiological data across these situations with wearable sensors. Thereby, the two main limitations of the dominant DM paradigm for flow research are addressed, namely the elicitation of shallow flow experiences and the lack of cross-situational research. Therefore, in a first experiment a more naturalistic (i.e. closer to a real-world KW setting) laboratory observation approach is developed that integrates controlled environments and the original flow field research method, the Experience Sampling Method (ESM). In doing so, improvements to both internal validity and external validity of flow neurophysiology research are proposed. For the same purpose, in a second experiment, two variables are then integrated into a DM paradigm, the provision of more autonomy, and the inclusion of a social interaction. The latter approach provides not only a means to intensify flow (as flow is reported to be deepened through social interaction - see Magyaródi and Oláh, 2017; Tse et al., 2016) but also allows to expand the previous work into the context of flow in digitally-mediated social interaction - a highly relevant workplace interaction format. Spanning these scenarios, sensor systems are integrated into a measurement instrumentation, that is likely to become usable in real-world settings in the near future (i.e. simple and wearable ECG and EEG sensors). These combinations critically extend the knowledge base on cross-situational neurophysiological flow observation, with high-potential measurement tools. Formally, the Research Goals (RG) for this dissertation are:

- **RG1:** Integrate the present body of knowledge on how neurophysiological data can be used to detect flow experiences.
- **RG2:** Identify how flow experiences can be intensified in the laboratory in cognitive tasks.
- **RG3:** Consolidate which neurophysiological patterns of flow can be detected with wearable sensors across different situations, including simplistic, naturalistic, and social interaction scenarios.

To reach these goals, the present work provides extensive integrations of the related literature and the results from multiple experimental instances and a variety of measurement instruments. The specific contributions are embedded in the description of the structure of the dissertation that concludes this introductory chapter.

1.3 Dissertation Structure

It is common in academic research that in the continuous discussion and exchange with other researchers, a research project is improved, shaped, evaluated, and even sometimes completely inverted. Research is to considerable extent teamwork, as in many cases, a single person would not be able to perform the data collection alone or know all available literature by heart. The same is true for presented research, many hands and heads shaped the results to small or sometimes even significant extents. Some sections of this dissertation have for that purpose been submitted to conferences for additional peer review. For each section where this applies, the related publications are disclosed at the start of the section. All referenced publications, including co-authors and attribution of contributions, are further detailed in the Appendix Section A.1.

To establish a background, Chapter 2 describes the state of research on flow theory. This chapter covers how flow is conceptualised and researched across a variety of domains like sports, arts, and work. Importantly, this chapter discusses how flow is currently measured and highlights self-report instruments as the present standard to establish a ground truth for flow observation. Afterwards, in Chapter 3, the latest developments in experimental flow research are summarised. In particular, the primordial role of DM paradigms is critically appraised, and the two major alternatives and extensions are presented. The first is the provision of more natural tasks through controlled Experience Sampling Method (cESM) and the provision of more task autonomy (= Autonomy Manipulation - AM). The second is the observation of flow in social interaction (= Social Context Manipulation - SCM). Mainly due to the comparison of physiological findings across these paradigms, the utilisation of multiple paradigms represents a valuable contribution to the field of flow research. The herein conducted neurophysiological study of flow is rooted in comprehensive reviews of its related work. Chapter 4 presents the results from two Structured Literature Reviews (SLR) that summarise the state of knowledge on the Peripheral Nervous System (PNS) and Central Nervous System (CNS) configurations during flow. The reviews highlight that to date, no distinct markers are known that allow to directly identify the occurrence of flow in someone's body or brain. This gap is likely due to a low degree of integration, that has impeded the discovery of more detailed patterns. The integration of these works makes the Structured Literature Review (SLR)s valuable contributions for flow and NeuroIS research. This chapter also includes a section that discusses the principles and limitations of psychophysiological research that are at the core of the research in this dissertation.

Building on the established background, Chapters 5, and 6 document the conducted laboratory experiments. In the first experiment in Chapter 5, the intensification of flow in the laboratory is pursued by contrasting an established mental arithmetic DM task to the more naturalistic observation (cESM) of thesis writing under controlled conditions. Main results highlight a potential qualitative divergence of reported flow in the form of flow with reduced stress perceptions, but greater physiological demand in the cESM approach. The higher autonomy in the cESM task is considered as one potential cause of this divergence. Besides, analyses of EEG patterns across tasks provide further evidence that refutes a major theory in flow neurophysiology (namely Transient Hypofrontality Theory - THT) that posits downregulation of frontal brain areas during flow. Instead, during flow, frontal areas are found to show increased activity (in the form of decreased Alpha frequency power), and more stable frontal activation over time (less fluctuation) that is likely related to focused task attention. In the second experiment in Chapter 6, results from the first experiment are followed up by extending the DM task with a condition for self-selected optimal difficulty. This manipulation reviews the question if more autonomy intensifies flow (= Autonomy Manipulation - AM) in a more controlled form. In addition, the traditional observation of task performance in isolation is contrasted to DM in small groups (= Social Context Manipulation - SCM). Thereby, the potential of intensifying flow in the laboratory through the inclusion of social interaction is investigated. Also, this experiment represents the first step to understand the individual neurophysiological processes during flow in digitally-mediated social interaction. In contrast to related work, flow is not found to be increased in groups, which is considered to be caused by a lack of social information or a lack of opportunities for selecting sub-tasks that represent more optimal difficulty (a lack of autonomy). Neurophysiological analyses further consolidate previous findings. Specifically, they confirm propositions that flow is likely represented by moderate physiological activation (indicated by moderately reduced Heart Rate Variability (HRV) levels), elevated attentional engagement (indicated by reduced frontal EEG Alpha power), and moderate mental workload (indicated by elevated EEG Theta and HiBeta power). The vital contribution of this work is two-fold. First, as diverse findings on configurations of heart and brain are currently present in the related literature, the present findings that span different manipulations, provide valuable results to consolidate the body of knowledge on flow neurophysiology. Second, the results highlight limitations and a promising potential to detect boundary conditions (i.e. non-optimal workloads) for the emergence of flow in the employed wearable EEG system. Specifically, higher Beta frequency ranges are found as a workload-sensitive and specific feature that is more robust than other features in the used feature space, and that even shows similar influences through other group members as do the reported flow experiences.

Following up on the results on the individual level, Chapter 7 further explores the social interaction condition of the second experiment to understand better, which group-level flow dynamics unfold in these increasingly relevant interaction scenarios. Given the added complexity of social interactions, analyses are pursued that seek to identify potential causes and consequences of flow experience intensities in digitally-mediated interactions. First, this includes the study of measures of configurational and perceived diversity metrics of the groups (e.g. ages, genders, technical abilities, but also perceptions of diversity) and whether or not they show relationships to experienced flow. In this work flow experiences in groups appear to be independent of most of these diversity metrics. However, perceived diversity in group member effort is found to predict flow experience intensities. Second, the relationship of flow to desirable group interaction experiences (group performance, satisfaction and growth) is investigated. While such relationships have repeatedly been reported before, the novelty of the lower flow experience in small groups in this experiment and the digitally-mediated interaction setting warrant the question if such relationships are affected. The results confirm that flow is positively related to these group interaction experiences also in digitally-mediated interaction. Third, the possibility of an emergent group-level flow experience is investigated. This exploration represents a particularly interesting direction, given its sparseness in the related literature. The present experiment shows that a weak to moderately strong reciprocal influence of group members on individual flow (i.e. a shared flow) is likely emerging in the designed setting. One the one hand, this confirms propositions that a group-level flow experience might exist. However, the found levels are much lower than in the one study that conducted similar analyses. Following up the possible reasons for this difference, a particular contribution of this chapter is the analysis of possible covariates that enable the emergence of shared flow experiences. It is observed that reciprocal flow influences are stronger when group members perceive higher autonomy, and when the task difficulty is lower. The latter finding is confirmed through experiment design, report, behaviour, and neurophysiology features. This observation implies that neurophysiological data could be used to detect boundary conditions (i.e. non-optimal workloads on the group level) for the emergence of shared flow experiences. Lastly, by the observation of flow reports and EEG features, a similar pattern in both variables is found that could indicate a relationship between reciprocal group influences of flow and workload. As flow theory is strongly rooted in the argument that (elevated) optimal difficulty is required for more intense flow experiences, the theory could be extended by the proposition that (optimal) reciprocal influences of difficulty act similarly as a precondition for the emergence and intensification of shared flow experiences.

To consolidate both experiments, the findings are integrated and discussed in Chapter 8. Specifically, it is discussed how the different experimental approaches to elicit flow experiences compare regarding the goal of intensifying flow (for individuals and groups) under controlled conditions. As a central result, it is appraised, that the integration of higher levels of participant autonomy emerged as the most effective driver of flow intensification. Opportunities to further extend and integrate this factor into flow laboratory research are outlined. In terms of neurophysiological observations, the amalgamation of findings from the experiments in this dissertation (that relied primarily on wearable sensors) leads to a series of consolidating and novel results that are discussed extensively. In principle, it is described how wearable sensors allow describing the configuration of the brain and the heart during flow. From the present data, it can be summarised that flow appears to be represented by moderate physiological activation (moderate HRV) and mental workload (moderate HiBeta power - and tentatively elevated frontal Theta power), and by elevated attentional engagement (reduced and stable frontal Alpha). In addition, flow appears to be represented by an absence of variation in approach-avoidance motivation or affective valence (as indicated by the absence of Frontal Alpha Asymmetry (FAA) changes). Importantly, these results emerge through the inclusion of various mechanisms for the elicitation of flow experiences in the laboratory (DM, cESM, Autonomy Manipulation (AM), and Social Context Manipulation (SCM)), which represents the major contribution of this work to the flow neurophysiology literature. Of particular relevance is the finding that through frequency band personalisation and sub-segmentation promising new options for flow detection emerged. Specifically, the frequency band segmentation highlighted the particular sensitivity of the HiBeta frequency ranges with manipulations of difficulty. An additional absence of confounds with time, and a group level influence on HiBeta power levels, further indicate that this higher frequency range could have a valuable role for the observation of flow on the individual and group level. While a connection of Beta powers to flow is not entirely new, its sensitivity and emergence over a wider area of the scalp (i.e. also in wearable EEG) make it a promising feature to be leveraged in adaptive NeuroIS in the future. The opportunities for the development of adaptive NeuroIS based on current possibilities and under consideration of the (eventual) ethical limitations represent the last major topics of discussion. Herein, of high importance is an understanding, that flow must be balanced with episodes of recovery. Also, ethical ramifications, like the dangers of developing addictive tendencies and how to take preventative action in adaptive systems research are described, and prescriptive actions are suggested, like for example the facilitation of flow with diversification amongst activities.

Finally, in the concluding Chapter 9, the contributions of these works are summarised. The first major contribution is the integration of neurophysiological knowledge about flow in the electrophysiological domain (specifically ECG and EEG), from two SLR, and two multi-paradigm experiments. As a promising metric, in particular, high-frequency EEG features are found to provide the interesting potential to not only unobtrusively identify boundary conditions for individual flow, but also for shared flow experiences of small groups - all given using wearable sensor systems. The second major contribution is the extension of flow research by the exploration of alternative research paradigms and the identification of new hypotheses on how to elicit flow in the laboratory with increased internal and external validity. Hopefully, the provision of these findings contributes a piece to the larger puzzle that is the facilitation of flow at work that has been highlighted as such a desirable goal for individual, organisational and societal reasons (Bakker and Woerkom, 2017; Spurlin and Csikszentmihalyi, 2017; Gallup, 2017).

Chapter 2

Flow Theory

Contents of this section are in part adopted or taken from Knierim et al. (2017c), Knierim et al. (2018a), Knierim et al. (2018b), Knierim, Nadj, and Weinhardt (2019), and Knierim et al. (2019). See Section A.1 for further details.

2.1 Flow Importance & Components

In this dissertation, the central construct of interest, is the experience of flow, a primordial instance of (task and profession) engagement (Tse, Nakamura, and Csikszentmihalyi, 2020; de Moura Jr and Bellini, 2019). To foster flow experiences through adaptive Neuro-Information Systems (NeuroIS), a stable theoretic basis is required to serve as the conceptual reference for such an adaptive system in the future (Riedl, Davis, and Hevner, 2014; Brouwer et al., 2015). This chapter provides a comprehensive overview of flow theory's importance, its components, and its research approaches to establish the understanding on which the reviews and experiments in this dissertation are built.

Flow theory represents a body of theoretic and empiric works in the realm of positive psychology (Seligman and Csikszentmihalyi, 2000), conceptualised by Mihaly Csikszentmihalyi (1975; 1990; 1996). The theory focuses on experiential states that explain intrinsically motivated behaviour, peak performances, personal growth and general well-being in life (Tse, Nakamura, and Csikszentmihalyi, 2020; Nakamura and Csikszentmihalyi, 2009; Moneta and Csikszentmihalyi, 1996). Flow theory resulted from Csikszentmihalyi's interest in the question of why some individuals (e.g. professional musicians, artists, or athletes) pursued daily activities with the exertion of extensive mental and physical resources, without typically expected incentives (e.g. monetary compensation) (Csikszentmihalyi, 1975). Thousands of interview hours later, the author concluded, that intrinsic motivation and the prospect of experiencing a highly rewarding mental state cause this type of behaviour (Csikszentmihalyi, 1975; Sadlo, 2016). Thus, a central thesis of the theory emerged, that psychological well-being and happiness are not generated by material artefacts, but through the pursuit of intrinsically motivated behaviours. In doing so, a so-called "optimal" psychological experience (flow) is elicited, that creates a feeling of high intrinsic reward and satisfaction (Nakamura and Csikszentmihalyi, 2009). Due to this lack of extrinsic rewards, flow experiences have been termed autotelic (greek autós ="self" and télos = "goal"). To pursue flow means to pursue an activity just for the sake of experiencing it (Asakawa, 2004; Moneta and Csikszentmihalyi, 1996; Nakamura and Csikszentmihalyi, 2009). Flow is, therefore, also termed an archetypal form of intrinsic motivation (Deci and Ryan, 2000). Because these flow states are experienced as so rewarding, individuals are incentivised to seek out flow-opportunities over and over again (Nakamura and Csikszentmihalyi, 2009). To experience flow, individuals need to seek out situations with increasing challenges, as the experience of capacity (skill) extension is considered a central driver of flow emergence (Moneta and Csikszentmihalyi, 1996; Nakamura and Csikszentmihalyi, 2009; Stavrou, 2008).

In the flow state, individuals act with total involvement and absorption in the activity, to the degree that awareness of the self, the environment and time are lost (Moneta and Csikszentmihalyi, 1996; Nakamura and Csikszentmihalyi, 2009). More explicitly, the experience of flow is conceptualised to encompass nine distinct characteristics: (1) difficulty-skill balance, (2) clear goals, (3) unambiguous feedback, (4) merging of action and awareness, (5) perception of total control, (7) loss of self-consciousness, (8) transformation of time, (9) concentration on the task at hand, and (9) autotelic experience (see Figure 2.1). These nine characteristics were later separated into two parts that are the preconditions of flow experiences (dimensions 1-3) and the core components (dimensions 4-9) (Nakamura and Csikszentmihalyi, 2009). Over the years, these dimensions and categorisations have been developed further by different research groups. For example, in an effort for more parsimonious conceptualisation and following empiric validation of survey instrument developments, Engeser and Rheinberg (2008) argue for a reduction of the core flow components to the dimensions of absorption and fluency. Besides, the authors have remarked that affective experience components, e.g. enjoyment or reward, are more appropriately considered as consequences to flow because the lack of self-awareness during flow makes the conscious experience of affect unlikely (Engeser and Schiepe-Tiska, 2012). As another example for the conceptual nuance, Bakker (2008) argues for a simplification of the flow construct to the dimensions of absorption, enjoyment and intrinsic motivation and therefore integrates affective and motivational constructs into the flow experience. Nevertheless, while these distinctions are not unanimously shared, flow scholars consider, that there is overall a high level of consensus about the components of flow (Engeser and Schiepe-Tiska, 2012). Flow exists on a continuum of intensity. That means, flow should not be considered as a binary experience (flow or non-flow), but rather as a state that can have no intensity, little intensity (i.e. shallow flow), or high intensity (i.e. deep flow) (Moneta, 2012). The differentiation can be seen in the emphasis of isolation from the environment, the lack of self-awareness, or the distortion of time (Moneta, 2012). In the most intense flow experiences, time seems to be altered severely, and situations can appear to occur in slow-motion like speeds (Csikszentmihalyi, 1975). In contrast, a less intense flow experience might lead to forgetfulness of time, when, for example, playing a video game in which one is absorbed. This distinction of flow intensity as a continuum is an important assumption in the measurement of flow.

Flow experiences are said to be possible in almost any task, as long as they are sufficiently complex (but structured) and that active engagement is required from the individual (i.e. passive tasks like watching TV have been found to be less conducive to flow) (Moneta and Csikszentmihalyi, 1996; Delle Fave and Massimini, 2005; Csikszentmihalyi, 1975; Csikszentmihalyi and LeFevre, 1989). In support of this proposition, studies have found support for the experience of flow in contexts like musical performance (Manzano et al., 2010; Jaque, Karamanukyan, and Thomson, 2015), academic learning (Engeser and Rheinberg, 2008), gaming (Harmat et al., 2015; Klarkowski, 2016), online surfing or shopping (Mauri et al., 2011; Cipresso et al.,

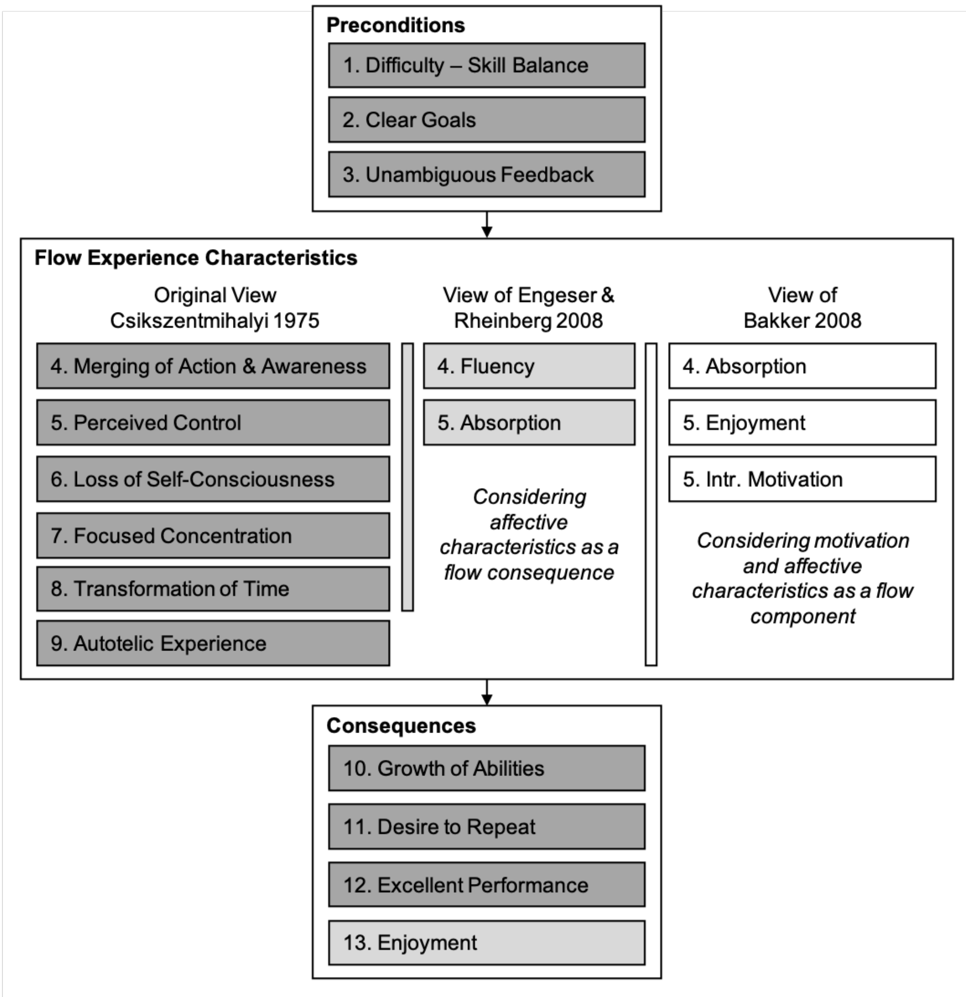

FIGURE 2.1: Flow Theory Components Including Two Alternative Perspectives by Engeser and Rheinberg (2008) and Bakker (2008).

2015), or work (Quinn, 2005; Bakker and Woerkom, 2017). Within the preconditions for flow to occur, the balance between perceived difficulties of the task and perceived abilities of the individual (that both also have to be present at a high level), has been discussed as a primal factor and has spawned theoretic models that aim to describe psychological experience like flow, boredom, anxiety, apathy (and more) in terms of the interplay of these two dimensions (Csikszentmihalyi, 1975; Moneta, 2012; Nakamura and Csikszentmihalyi, 2009; Fong, Zaleski, and Leach, 2015). Figure 2.2 details the historical development of these models. For these models, a few central lines of reasoning need to be highlighted. First, the dimensions of difficulty and skill refer to subjective appraisals of these dimensions. This subjectivity means that in a given situation it is considered less critical what the objective levels of the two dimensions are, but how the individual perceives them to be (Csikszentmihalyi, 1975; Csikszentmihalyi and LeFevre, 1989; Moneta and Csikszentmihalyi, 1996; Nakamura and Csikszentmihalyi, 2009). Second, the balance of the two dimensions is considered to be required present at a sufficiently high level of both difficulty and skill. This requirement means, that flow eliciting situations require the individual to stretch already well-developed abilities. Only in this situation, the individual is not only enjoying the moment but extends capacities (i.e. realises the development of abilities and increases in personal complexity) (Csikszentmihalyi and LeFevre, 1989).

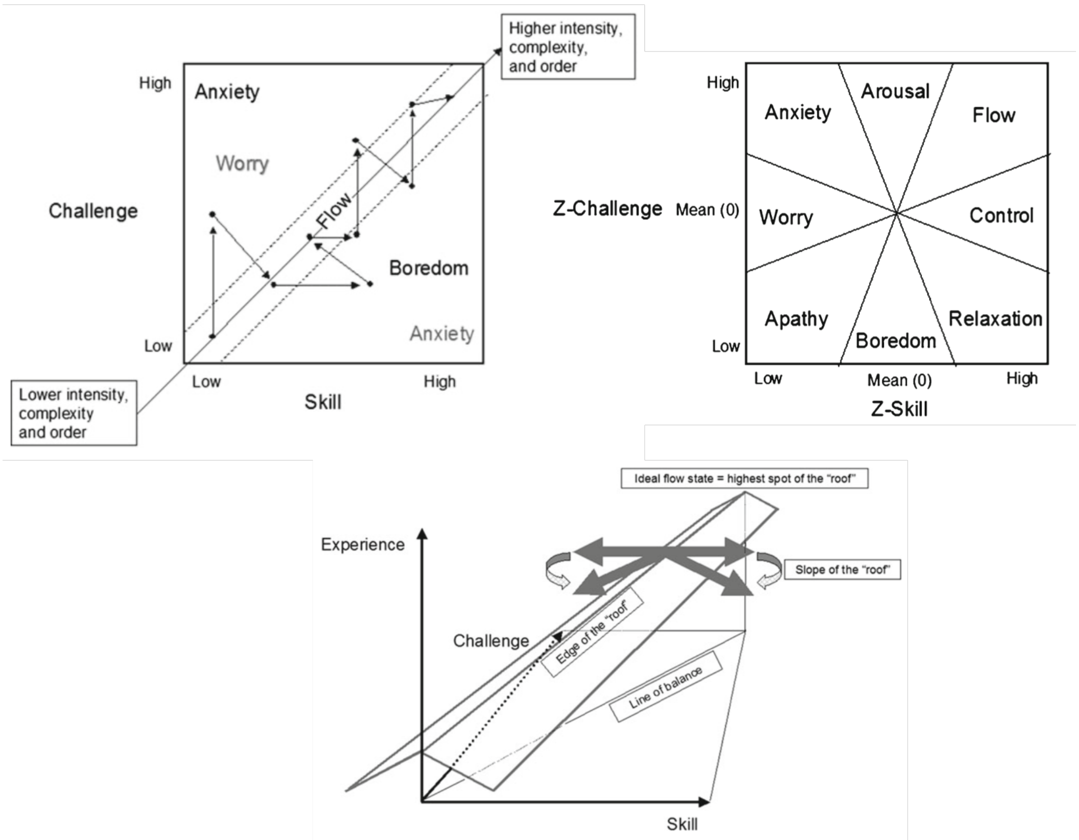

FIGURE 2.2: Flow Experience as a Function of Difficulty and Skill - See Moneta (2012).

This fact is increasingly represented in the refinements of the difficulty-skill models (first by the inclusion of individualised, i.e. z-scored, above-average levels of both dimensions, later as the intercept in the regression model). In its latest iteration, the difficulty-skill models culminated in a three-dimensional model based on regression modelling perspectives (Moneta, 2012). In this approach, flow ("optimal" experience) is explained by the heightened presence (intercept β_0 above 0) of difficulty (β_1) and skill (β_3), that lead to an intensification of flow, together with the required balance of the two dimensions for the potentiality of flow (β_3). The associated regression equation thus becomes:

$$(flow)\ experience = \beta_0 + \beta_1 difficulty + \beta_2 skill + \beta_3 |challenge - skill|$$

The regression model approach was conceived to overcome the limitation of earlier models for the explainability of the different experiential components in the quadrant and channel models. Specifically, studies on the earlier models showed limited potential to explain intensity differences in subjective experience as a product of difficulty and skill balance (i.e. flow intensities), (Ellis, Voelkl, and Morris, 1994; Moneta, 2012), or how difficulty and skill alone contribute to flow experiences. Importantly, the regression model reduces the explanatory focus towards the "optimality" of experience and therefore allows to predict higher levels of flow (optimal experiences), and non-flow (non-optimal experiences) without also having to explain the exact type of other (i.e. non-flow) experience. Therefore the model represents not only a more refined, but also a more parsimonious theoretic account of this central prediction in flow theory. Beyond the sophistication and validity of these difficulty-skill models, it needs to be appraised, that the validity and dominance of this difficulty-skill-balance argument for flow potentiality are also heavily criticised. Some authors argue, that there are diverging empirical results that highlight that some individuals favour balance while some favour slight imbalances (Csikszentmihalyi and LeFevre, 1989; Keller et al., 2011; Moneta and Csikszentmihalyi, 1996; Løvoll and Vittersø, 2014). Also, some research documents the fact that some individuals tend to experience flow more often than others, for reasons that are not yet fully known (Asakawa, 2004; Ullén et al., 2012; Moneta, 2012).

Regardless of its potentiality, the experience of flow is attributed to coincide with excellent performances and heightened creativity (Asakawa, 2004; Csikszentmihalyi, 1975; Nakamura and Csikszentmihalyi, 2009). A large number of studies documents the relationship between flow and improved task performances across various domains like academia, sports, and work (Engeser and Rheinberg, 2008; Stavrou et al., 2007; Schüler, 2007; Brunner and Schueler, 2009; Yotsidi et al., 2018). Thereby, associations between flow and subjective and objective performance variables are detected for short-term (i.e. state) observations and long-term (i.e. longitudinal) measurements. The flow-performance relationship is explained by the flow characteristics that describe a highly functional state (Engeser and Rheinberg, 2008). For example, high concentration during flow and a loss of self-consciousness are supposed to align attentional processes efficiently, and the rewarding experience supposedly increases task motivation, effort, and perseverance (Brunner and Schueler, 2009; Engeser and Rheinberg, 2008). These propositions are further elaborated by neuroscientific work that argues for enhanced neural cooperation of attention processes (Harris, Vine, and Wilson, 2017b; Weber et al., 2009). In particular, it is supposed that during flow, task-related attention is shielded from interference by alerting or orienting attention processes, and that impulse control is improved (Harris, Vine, and Wilson, 2017b; Manzano et al., 2013). In addition, flow experiences have been linked to increases in general well-being (Tse, Nakamura, and Csikszentmihalyi, 2020). High levels of intrinsic satisfaction and enjoyment are ascribed to accompany or follow flow, which is why higher frequencies for flow experience are assumed to improve well-being (higher frequencies of positive emotion experience are generally found as an important contribution to general well-being - see Lyubomirsky, King, and Diener, 2005). The intrinsic enjoyment is considered to be rooted in the experience of competence (self-satisfaction from mastering a difficult task at the edge of one's abilities, but also cognitive efficiency - see Harris, Vine, and Wilson, 2017b). Conversely, flow is considered to feel "good" through the absence of self-conscious and ruminative thoughts (Sadlo, 2016), and the absence of threat (Tozman and Peifer, 2016). In the latter view, flow is attributed to a positive form of stress (similar to the concept of eustress - see Selye, 1980). Thus flow is said to not only have positive emotional qualities but might be shielding from negative emotional experiences too (flow is considered to shield from burnout at work - see Yotsidi et al., 2018). For these cognitive and affective benefits, positive psychology scholars seek to improve understanding of how to foster flow in daily life (de Moura Jr and Bellini, 2019).

2.2 State & Standards of Flow Research

The study of flow experiences is argued to date to be a rigorously researched, empirically supported, and stable concept in the domain of positive psychology (Asakawa, 2004; Nakamura and Csikszentmihalyi, 2009; Engeser and Schiepe-Tiska, 2012).

While a high level of agreement on the definition of flow is present (Engeser and Schiepe-Tiska, 2012), there still exists a lively debate about nuanced aspects like the appropriation of some minor dimensions to the flow construct (some authors include a few different sub-dimensions in their understanding of flow - see Figure 2.1), or the appropriateness of different measurement approaches (Moneta, 2012). Nonetheless, some directions for the extended development of flow research have been put forward by scholars in the field. In their overview of the state of research on flow in the time from 2005-2010, Engeser and Schiepe-Tiska (2012) conclude that flow research is (1) still focused on understanding flow conditions (more so than consequences), and (2) focused on measurement methods (especially with questionnaires). The major research areas at that time were (1) sports, (2) learning in educational settings, (3) game-based learning and media use (human-computer interaction in general), but that research on work contexts is also increasing. Furthermore, as emerging trends, they identify both flow in social contexts and the psychophysiology of flow. The emerging relevance of flow at work is best described by the observed trends of low worker engagement that were outlined in the introduction to this dissertation. As flow is considered a source of improved performances and well-being, it is considered a vehicle that could foster such worker engagement and should, therefore, be understood better. In a similar direction, as many activities, especially in Knowledge Work (KW), are conducted in social contexts, the - until then - somewhat neglected social aspects of flow experiences have come into focus. This trend has continued and is fuelled by propositions, that flow in social interaction is possibly more intense than in isolation (Walker, 2010; Tse et al., 2016; Magyaródi and Oláh, 2017). The importance of neurophysiological flow observation is emphasised by the fact that to date, only obtrusive measures are available (i.e. self-reports), that can be used to gain an impression on the situations in which flow is experienced. Using such report instruments, interrupts flow, which has been called a detriment to flow experience by itself (Moneta, 2012; Engeser and Schiepe-Tiska, 2012). Furthermore, such self-reports do not allow to develop a more precise picture of the time dynamics of flow experiences. This knowledge would be highly valuable to develop interventions that help to foster flow (e.g. adaptive NeuroIS).

However, despite these limitations, self-report instruments currently represent the gold standard as a measure of ground truth for flow experience (Moneta, 2012). As this dissertation follows up on the aforementioned emerging opportunities (flow in KW, in social contexts, and its neurophysiological correlates), it too relies primarily on these self-report instruments as a frame of reference. Therefore, their properties are explicitly reviewed.

The measurement of flow experience is still a central challenge for the theory and empirical research (Engeser and Schiepe-Tiska, 2012; Moneta, 2012). This challenge is reflected by the fact that today, a multitude of approaches exist, albeit with all of them representing self-report instruments. To gain a broader impression of the used practices, especially experimental research (to which the work in this dissertation is most comparable) was analysed for their use of measurement instruments. The observations in this section are, therefore, based on the findings from the two Structured Literature Reviews (SLR) on studies on flow neurophysiology that are described in detail in Section 4.2 and Section 4.3. Based on these findings it can be stated that in flow (neurophysiology) research, the majority of flow research is primarily grounded in self-report assessment, for which multi-item (and multi-dimensional) reports are considered the standard (Moneta, 2012; Keller, 2016). Some earlier approaches of flow self-report assessment like the indirect assessment of flow through the assessment of perceived difficulty and skill alone are still being used by some researchers (e.g. Labonté-Lemoyne et al., 2016; Gaggioli et al., 2013). However, this approach has mostly fallen out of favour due to this indirect and highly reductionist flow measurement approach (Moneta, 2012). With most other research flow is operationalised as a unidimensional, higher-order construct that is reflected by (at least some of) the nine theoretic dimensions outlined in the previous section. This combination means, that flow is typically interpreted as the average of these multiple sub-dimensions. In the aggregation, no distinction is made between the categories of preconditions, dimensions, and consequences. Through this averaging, a valid and reliable self-report indicator is retrieved (after all, high values on the preconditions and outcomes would integrate well with high values on the flow dimensions) (Keller, 2016).

However, while this compensatory multi-dimensional aggregation approach is common to most flow scales, differences exist in the extensiveness and considered sub-dimensions. This variation is primarily explained by different goals of reaching internal validity or feasibility in research designs (e.g. repeated measurement typically requiring shorter questionnaires), or by domain-specific adaptations of flow measurement. In the most traditional reasoning in line with flow theory, the nine-dimensional Flow Short Scale (FSS) is used by some researchers (e.g. Klarkowski, 2017; Shearer, 2016; Kivikangas, 2006) in its long form (4x9 = 36 items) (Jackson and Marsh, 1996) to build empirical analysis on a psychometrically comprehensive basis. The FSS was developed specifically with the goal in mind to provide a multi-item and multi-dimension scale that represents all the nine dimensions outlined in flow theory (Jackson and Marsh, 1996; Martin and Jackson, 2008). However, especially due to the length of the initial 36-item version of the FSS, short scale versions have been developed (i.e. the 9-item Flow Short Scale - sFSS) that comprise short form measures of flow experience built on the strongest loading items of the FSS. Such scales are preferentially used by scholars that employ repeated measures study designs (e.g. Harmat et al., 2015; Manzano et al., 2010). Similarly, alternative scales have been developed like the Flow Kurzskala (FKS) (a 10 item - 2 dimension instrument by Rheinberg and Vollmeyer, 2003; Engeser and Rheinberg, 2008) that applies a more reductionistic approach with the argument that flow can be simplified to comprise the experiential dimensions of fluent action or behaviour and high task absorption. A breadth of experimental flow research has employed the FKS scale, potentially due to this practical sparseness in items and factors (e.g. Wolf et al., 2015; Peifer et al., 2014; Tozman et al., 2015; Peifer et al., 2015; Tozman, Zhang, and Vollmeyer, 2017). In another approach still rather close to flow theory, but with a domain-specific focus, the WOrk-reLated Flow Inventory (WOLF) instrument (a 13 item - 3 dimension instrument by Bakker, 2008) has been developed to capture flow experiences specifically in the work environment. This self-report instrument integrates the dimensions of absorption, enjoyment, and intrinsic motivation. An interesting aspect of this scale is that it is the only instrument that considers flow aggregation using a conjunctive logic. This logic means that flow is only considered to be intense when all three dimensions are present at a high level (each in the 75% quartile). However, this instrument has so far not been utilised in (neurophysiological) flow experiments. Beyond these instruments, a variety of domain-specific instruments exist like the Cognitive Absorption (CA) scale developed in the context of Information Systems (IS) use (Agarwal and Karahanna, 2000), or the Game Experience Questionnaire (GEQ) (long and short versions) that include flow experience in few item manners as a sub-dimension of a larger battery utilised to measure typical experiences during computer game playing (Ijsselsteijn, Poels, and De Kort, 2008). With these measurement instruments, there is an increased distance to the accounts of flow theory. Yet, they are prolifically used in (neurophysiological) flow research, which makes the integration of related work and the establishment of ground truth for flow measurement a difficult venture.

Given this breadth of options for flow measurement, researchers, therefore, currently have to decide on which instrumentation to apply as a metric of ground truth, under the restrictions of the chosen research approach. For this dissertation, that focuses on better understanding how to measure flow in various research setups, the emphasis was placed on selecting a moderate-item, multi-dimensional self-report instrument to provide an anchor for ground truth that is firmly attached to flow theory. As will be described in the experiments in Chapter 5, and Chapter 6, the FKS scale (Engeser and Rheinberg, 2008) was chosen as this reference instrument. The reason for doing so is the sparseness of the scale (that still allows a multi-dimensional interpretation of experiences for more detailed insight) and the previous use in multiple related studies.

To provide a short background of flow theory's development and especially its related research methods, this section briefly outlines some of the historical developments. Given this dissertation's focus on two specific methods - the Experience Sampling Method (ESM) and laboratory induction paradigms - the description of earlier work is kept short. For a more extensive review of the history of flow, the reader is referred to Rich (2013) and Engeser and Schiepe-Tiska (2012). As covered in section 2.1, upon the conception of flow theory, initial accounts of experiences of flow states were derived from interviews with expert performers in fields like sports (e.g. rock climbers), arts (e.g. musicians), and high-risk work (e.g. surgeons) (Csikszentmihalyi, 1975). As these accounts were derived from experiences in extreme situations, and are naturally limited by the interview format (biases and low level of detail), an alternative approach had to be developed to form knowledge about the daily occurrences of flow. The Experience Sampling Method (ESM) was created precisely for this purpose (Csikszentmihalyi and Hunter, 2003; Trull and Ebner-Priemer, 2009). As the name suggests, the ESM works by collecting multiple accounts of daily experience, operationalised in the form of numerous, randomly timed interruptions by pagers (or smartphones), every day over several days to weeks. Through this process, detailed descriptions of experiences in a broad range of activities and circumstances can be collected. This method has been the prime approach to further deepen knowledge about flow experience and its relationship to other constructs of interest on the situational, contextual, and personal level, all within a naturalistic setting (Moller, Meier, and Wall, 2010). While the ESM is seen as a major contribution to both flow and general psychological research (Rich, 2013), and has allowed for important theoretical extensions of the flow concept (for a review see Moneta, 2012), this methodological focus has led to two central limitations. First, nearly all of the knowledge created in the first 30 years of flow study has been correlational (thus limiting the knowledge of flow causes and consequences) (Keller, 2016; Moller, Meier, and Wall, 2010). Second, the focus on observational approaches has limited researchers in their means to deepen knowledge on flow-related phenomena like neurophysiological dynamics effectively.

In the last decade (coinciding with the increased interest and feasibility of neurophysiological approaches) there has been an increased interest in the study of flow in more controlled, experimental flow research approaches (see also Chapter 4). While the experimental study of flow still represents a major challenge (for reasons described below), it should be noted that important advances have been made, specifically centred around one experimental paradigm of Difficulty Manipulation (DM) (Moller, Meier, and Wall, 2010; Keller, 2016). The reason for this focus is likely due to the simplistic appeal of the salient difficulty-skill experience models, that explain flow states as an above-average balance of these two dimensions (see Figure 2.2). While the role of difficulty-skill balance as a single cause for flow experience, is not substantiated (clear goals and feedback are considered in theory as just as highly important) (Moneta, 2012), it has remained as a central factor that guides many flow studies (Fong, Zaleski, and Leach, 2015). To date, the DM approach represents the most established experimental approach in flow research (Moller, Meier, and Wall, 2010; Keller, 2016). For this reason, the experiments conducted in this dissertation are built in reference to it. Due to this central position of the DM approach for flow research in general, and for this dissertation, the approach is discussed in detail next. Afterwards, promising extensions are discussed, that form the basis of the experimental work conducted in this dissertation together with the DM approach.

Chapter 3

Experimental Flow Research

3.1 Best Practice: Difficulty Manipulation

Contents of this section are in part adopted or taken from Knierim et al. (2018a), Knierim et al. (2018b), Knierim, Nadj, and Weinhardt (2019), and Knierim et al. (2019). See Section A.1 for further details.

To observe flow requires adequate means of eliciting the experience. Ideally, such elicitation can take place under controlled conditions to eliminate confounding factors. This means that experimental paradigms are needed that reliably allow the manipulation of the circumstances under which flow emerges. As discussed in the previous chapters, primarily three preconditions are included in flow theory that need to be met, and that should provide opportunities to elicit different flow intensities (balance of difficulty and skill, clear goals, and unambiguous feedback). However, present literature is also discussing important moderating variables that influence the emergence of flow (e.g. autonomy, expertise, task relevance, and social interactions). This chapter reviews the current state of experimental flow research to ground the efforts in this dissertation that are focused on overcoming central limitations in the established paradigms, and, therefore, on intensifying flow experiences under controlled conditions. In the first section, the de-facto experimental standard, the Difficulty Manipulation (DM) paradigm is presented. In the following two sections, the two directions for flow intensification that are primarily pursued in this dissertation are motivated and discussed. These are, first the observation of flow in more naturalistic and autonomous settings (cESM, and AM), and second the observation of flow in social interactions (SCM).

As previously stated, the idea of manipulation of difficulty is rooted in one of flow theory's concepts, that subjective experience is a function of the perceptions of task difficulty and abilities to cope with the task (Moneta, 2012; Csikszentmihalyi, 1975). This model does not only account for the experience of flow (difficulty-skill balance above average levels), but also non-optimal experiences like boredom (higher perceived skill than difficulty), or anxiety (lower perceived skill than difficulty - see Figure 2.2) (Moneta, 2012; Delle Fave, Massimini, and Bassi, 2011). Based in this conception, since the early 2000s (Rheinberg and Vollmeyer, 2003; Moller, Meier, and Wall, 2010), multiple studies have built up an experimental flow elicitation approach that focuses on providing tasks varied only in difficulty, to elicit these proposed experiential contrasts (Rheinberg and Vollmeyer, 2003; Keller and Bless, 2008; Moller, Meier, and Wall, 2010; Keller, 2016). To date, this paradigm (Difficulty Manipulation - DM) has been employed across diverse tasks like computer games (e.g. Tetris, Pacman) (Ewing, Fairclough, and Gilleade, 2016; Barros et al., 2018; Peifer et al., 2014), factual knowledge (Keller et al., 2011), mental arithmetic (Ulrich et al., 2014; Katahira et al., 2018) or chess (Tozman, Zhang, and Vollmeyer, 2017). Also, the DM paradigm has been parametrised for durations of multiple seconds to multiple hours (see Kennedy, Miele, and Metcalfe, 2014 and Ulrich, Keller, and Grön, 2016a for very short, Keller, 2016 and Léger et al., 2014 for very long instances). However, the majority of work utilises ranges of 4-8 minutes per difficulty condition (Keller, 2016). The approach has been strongly focused on gaming tasks given their manipulability, and often given expected lack of performance stressors or other extrinsic expectation confounds (Moller, Meier, and Wall, 2010). It should be pointed out that the manipulation of difficulty, is not unique to flow research. Instead, there are a few fields of research that have previously studied the impact of varying difficulty levels on psychological (and importantly, neurophysiological) experiences. The two main related fields that we found are those of stress research and mental effort or cognitive workload. Both are discussed in Tozman and Peifer (2016) for their relations to flow research. The main differences between the work in these fields pertain to the operationalisation of moderate difficulty conditions and high difficulty conditions. For the former, the specific extension in flow research is the inclusion of calibration and adaptive "optimal" difficulty conditions (Tozman and Peifer, 2016; Keller, 2016; Moller, Meier, and Wall, 2010). For high difficulty conditions in flow research, importance is placed on creating conditions that elicit high demand, without excessive difficulty, to keep participants engaged (Keller, 2016). Overall, the DM paradigm is accepted broadly and is considered a significant advancement to experimental flow research by multiple authors (Moller, Meier, and Wall, 2010; Keller, 2016; Tozman and Peifer, 2016). However, minor and major critiques remain.

In terms of minor limitations of the DM approach, two factors have become apparent. The first minor limitation in experimental flow research is its focus on game tasks, a state that is visible from the SLRs on 20 studies on the PNS and 22 studies on the CNS (see Section 4.2, and Section 4.3). This focus likely came to be due to the simplicity and controllability of game tasks, but also their naturally occurring relation to intrinsically motivated task engagement (Klarkowski, 2017; Rheinberg and Vollmeyer, 2003). Such a focus not only strongly limits the transferability of findings, but is also likely to emphasise common-method biases. It is therefore important that flow DM research (focusing on neurophysiology in particular) starts to employ varied tasks more often and ideally includes multiple tasks for the cross-validation of findings (Barros et al., 2018). The second minor limitation in experimental flow research is that DM is realised relatively inconsistently. Moller, Meier, and Wall (2010) point to the observation that two problems arise in operationalisations that are: (1) controlled and manipulated variables and (2) calibration of adapted difficulty. In detail, they show that while some authors achieve their experimental manipulation using multiple variables (e.g. Keller and Bless, 2008 adjusts both task speed, user control, and Keller, 2016 discusses multiple adjusted variables in different tasks they employed in multiple studies), others do so using only a single variable (e.g. Moller, Meier, and Wall, 2010 using only game speed). The manipulation of multiple variables introduces confounding influences in the first approach, a shortcoming that is not yet well understood. It is for example not understood, how the manipulation of variables that relate to sensory input changes (e.g. the varied number of elements on the screen in the plane battle game task by Berta et al., 2013) might be represented in EEG measurements during flow. A particular operationalisation problem refers to the fact that some scholars provide a moderate difficulty level based on group averages (population/sample baseline). In contrast, others use within-subject calibration of optimal difficulty (individual baseline) (Moller, Meier, and Wall, 2010). The latter is considered to be aligned better with flow theory accounts - given the stressed subjectivity of both the difficulty and skill dimension (Csikszentmihalyi, 1975). The adaptation of difficulty that is sometimes not integrated into flow research, yet considered an essential advancement in flow experiment paradigms, might be creating inconsistent results in flow experiments. Similar issues relate to the difficulty in overload/anxiety experience conditions, that is to be contrasted with flow experience. Specifically, the problem is that some authors focus on providing task conditions in a difficulty that does not exceed an individuals' abilities too much (to keep participants engaged) (Keller, 2016), while others try to contrast flow with excessive difficulties to observe such disengagement (e.g. Ewing, Fairclough, and Gilleade, 2016). The resulting difference is considered to elicit both different self-reported and physiological responses (Keller, 2016; Tozman and Peifer, 2016), and could be a confounding factor in aggregated study results.

Beyond these minor limitations that mostly refer to confounds in the comparability of results, major limitations have been outlined that critique the general efficacy of the DM paradigm to elicit flow experiences in laboratory settings. In general, the role of difficulty-skill balance as a (sole) flow experience determinant has been criticised (see Fong, Zaleski, and Leach, 2015 with a cross-method meta-analytic study review). Specifically, it is noted repeatedly that, the objective difficulty-skill balance could not sufficiently meet the subjective nature of flow experience, and is thus inherently limited in eliciting flow in participant equally likely (Moller, Meier, and Wall, 2010; Keller, 2016; Tozman and Peifer, 2016; Fong, Zaleski, and Leach, 2015; Engeser and Schiepe-Tiska, 2012). Instead, the role and sufficiency of difficulty-skill balance for flow facilitation could be moderated by cultural and individual preferences for slight under- or overload (Løvoll and Vittersø, 2014; Tse et al., 2016). It has for example been reported, that individuals with negative achievement motivation traits like fear of failure, show less intense flow experience responses (i.e. favouring slight overbalance of skills to difficulty) (Engeser and Rheinberg, 2008). Beyond the personal level, it has also been outlined that laboratory flow elicitation might be limited by the neglect of accounting for additional, flow-theory related constructs. Moller, Meier, and Wall (2010) name a variety of factors that should be studied in concert with DM as they expect interaction effects for the elicitation of flow. Specifically, the role of task-relevance (or task instrumentality), but also the role of perceived control (or autonomy) have been stressed as having an important, yet often neglected role for flow (Moller, Meier, and Wall, 2010; Keller, 2016; Tozman and Peifer, 2016; Fong, Zaleski, and Leach, 2015). Such variables could moderate difficulty-manipulation success. In a similar line of criticism, Delle Fave, Massimini, and Bassi (2011) criticise flow laboratory experiments for (1) the expectation that flow will emerge in short trial times, for (2) a typical lack of task interest/self-relevance for study participants, and for (3) the lack of real-life complexity in artificial laboratory settings. The latter is stressed extensively, in the major critique, that flow elicitation in the laboratory might be difficult (not to say impossible) in general, due to the artificial nature of laboratory tasks. This critique is put forward by Hommel (2010), who points to neuroscientific research that finds perceptions of effort to be necessarily higher whenever novel stimuli are presented. This phenomenon occurs in opposition to naturally occurring stimuli for which goal anchors are already constructed that require less effort for goal activation, which in term facilitate goal selection, maintenance and switching in a highly efficient manner in known tasks and environments (Hommel, 2010). Therefore, the experience of effortless attention might be more strongly bound to naturalistic settings than present flow laboratory research is acknowledging. Culminating from these arguments is the question as to whether "real" or deep flow experiences are elicited in the present DM approaches (and to some degree by other artificial laboratory setups as well).

In summary, the main critiques of the present experimental state of flow research are related to the properties of its primary DM paradigm: (1) DM is considered to work well in creating flow experience contrasts. However, it presents the only paradigm that is often operationalised with game tasks, which might have caused a strong influence of common-method bias. (2) Operationalisations of DM instances have not always followed a shared rationale and control of adaptive and high difficulties, and the way they are achieved by manipulating few variables. (3) DM approaches might still be severely limited in eliciting intense flow by weakly accounting for interpersonal variance in difficulty preferences, by weakly accounting for task relevance and task control, and by providing highly artificial scenarios limiting the possibility of effortless attention (Hommel, 2010). Especially as neurophysiological research builds upon the quality of the experimental approaches, extensions of the DM paradigm, and in general, more creative laboratory research are required (Harris, Vine, and Wilson, 2017b). Following up on this requirement, it is a central contribution of this work to explore neurophysiological flow research in alternative, but DM-connected research approaches. Specifically, this work follows up on the recommendations that flow ought to be researched using measurement plurality (not only reported - but also more objective measures like neurophysiological observations) (Engeser and Schiepe-Tiska, 2012; Spurlin and Csikszentmihalyi, 2017), that flow research would benefit from comparing results from multiple research paradigms (i.e. not "just" DM) (Rich, 2013) and multiple tasks (Katahira et al., 2018; Barros et al., 2018), and that flow (neurophysiology) research should strive for increased internal and external validity (Moller, Meier, and Wall, 2010; Hommel, 2010). Specifically, in this work, a variety of paradigms (DM, & cESM, AM, and SCM), tasks (mental arithmetic & scientific writing), and measurement methods (self-reports & neurophysiological) are combined to study flow experiences in the context of KW. The increases in internal and external validity are thereby achieved by adapting the ESM to the laboratory setup and by extending an established solitary DM task to conditions with self-selected optimal difficulties (AM) and the small group level (SCM). These two extensions of flow (neurophysiology) research are discussed next.

3.2 Flow Intensification 1: (Controlled) Experience Sampling

> Contents of this section are in part adopted or taken from Knierim et al. (2018a), and Knierim et al. (2018b). See Section A.1 for further details.

Towards More Naturalistic Flow Research

Within emerging flow (neurophysiology) research, a central focus has been on highly controlled DM game tasks, leaving gaps to understand flow neurophysiology in more unstructured tasks typical to Knowledge Work (KW). Furthermore, it has been criticised whether or not the DM approach can elicit real flow experiences, given the reduced motivation and involvement common in laboratory tasks (Moller, Meier, and Wall, 2010), and given the elicitation of effortful attention from a novel and artificial task (Hommel, 2010). Therefore, current flow (neurophysiology) research could benefit strongly from extending paradigms to improve the elicitation and observation of flow experiences in alternative, controlled approaches. One central aspect of extending approaches is to focus on the observation of flow in more naturalistic settings. The idea is that by studying flow in settings where it might be more typical to emerge (closer to real-world scenarios), both higher internal and external validity might be achieved. To develop such an approach, central assumptions and related research are discussed. The main underlying assumptions for paradigm extension in this direction are three-fold: (1) expertise (high skill levels) together with challenging tasks should elicit higher flow intensities, (2) high task relevance/instrumentality should elicit higher flow intensities, and (3) increased autonomy in how to conduct one's behaviour is required to enable the former two preconditions. The latter requirement means that the selection of challenging and important aspects of the task presents a catalyst for flow intensification. The argument for the necessity of domain expertise for intensified flow experience is outlined by multiple researchers (Hommel, 2010; Ullén et al., 2010; Moneta, 2012; Harris, Vine, and Wilson, 2017b) and forms a critical requirement integrated into flow theory (Csikszentmihalyi, 1975; Nakamura and Csikszentmihalyi, 2009). In fact, in the latest difficulty-skill experience models, the presence of abilities (and difficulties) beyond a moderate threshold is emphasised as a necessary precondition for flow (Moneta, 2012). High levels of expertise are supposed to be required for intense flow experiences due to the facilitation of effortless attention (Hommel, 2010) and highly automated task processing due to reliance on learned behaviours (Harris, Vine, and Wilson, 2017b). In that sense, only when the skills for a task are highly developed, one can get lost in the task because less explicit cognitive effort is required for the acquisition of a skill (Harris, Vine, and Wilson, 2017b; Ullén et al., 2010).

Task relevance or instrumentality is the second factor that has been considered to be important for intense flow to emerge. This assumption is based on the observation, that intense flow appears to be experienced more often by those who pursue a particular endeavour for the sake of mastering it repeatedly (intrinsically motivated behaviour) (Csikszentmihalyi, 1975; Partington, Partington, and Olivier, 2009; Delle Fave, Massimini, and Bassi, 2011). Task relevance is considered to drive flow experience intensities through the expenditure of necessary levels of effort and the channelling of concentration on the primary task that is to be completed (entirely focusing on the task at hand). Both Tozman and Peifer (2016) and Moller, Meier, and Wall (2010) emphasise the role of task relevance. Moller, Meier, and Wall (2010) state that it moderates the difficulty-skill-balance effect, Tozman and Peifer (2016) postulate that increased self-relevance should interact with it (i.e. higher self-relevance given difficulty-skill balance should lead to higher flow). Specifically, Tozman and Peifer (2016) state that challenging experiences can only occur when a situation is relevant to the individual and relevant to reach personal goals. In the previously discussed DM paradigm, task relevance is sometimes introduced by utilising engaging tasks (e.g. games - see Moller, Meier, and Wall, 2010), or extrinsic incentives (e.g. increased ego involvement through social evaluation - see Tozman and Peifer, 2016).

The third factor that is considered to be important for intense flow to emerge is some degree of autonomy in the setup and task of the observed individual. It is increasingly found that autonomy mediates the emergence of flow. For example in the work context, autonomy is found to facilitate flow experience by allowing workers to pursue tasks of personal relevance and with high, but manageable levels of task challenge (Bakker and Woerkom, 2017). In a related manner, Fong, Zaleski, and Leach (2015) in their meta-analysis of flow determinants also report on interest/self-relevance and autonomy/control as central factors at similar importance as difficulty-skill balance. In this sense, autonomy is considered a catalyst of more intense flow by allowing individuals to self-select into optimal boundary conditions. This logic assumes that individuals possess the required knowledge about these situations and themselves and want to create these conditions in a particular situation. Moller, Meier, and Wall (2010) also mentions the role of autonomy in flow research with theoretic and empirical arguments. Autonomy is outlined in Self-Determination Theory as a primary driver of intrinsic motivation (Deci and Ryan, 1985; Ryan and Deci, 2002). Also, Moller, Meier, and Wall (2010) mention an empirical study demonstrating that experiment participants with higher perceived freedom of choice also experience higher task absorption (Mannell and Bradley, 1986).

From these assumptions, it follows that improvements in the experimental elicitation of flow could be achieved through the integration of the requirements for expertise, task relevance and autonomy. Such integration is possible through, for example, a focus on observing domain experts in naturalistic tasks and settings.

Related Naturalistic Paradigms

In general, the idea of naturalistic flow observation is rather established, and both practised in and outside of the laboratory (Delle Fave, Massimini, and Bassi, 2011). On the one hand, traditional flow research has focused on collecting data during daily experience using diary methods like the ESM (Csikszentmihalyi and Hunter, 2003) - albeit not specifically within expert groups. On the other hand, researchers have attempted to employ more naturalistic task experiences by providing more naturalistic stimuli like rather complex, commercially available computer games (Labonté-Lemoyne et al., 2016; De Kock, 2014) and even allow participants to select preferred stimuli out of a range of options and to engage with them repeatedly (Shearer, 2016). The latter approaches are herein termed Engagement (ENG) paradigms because they focus on creating situations in which participants engage with the stimulus for an extended period to create a more naturalistic task experience than is typical with DM approaches. Both the ESM and the ENG approaches come with a particular set of limitations that hinder the integration with the goals of increased internal validity of flow elicitation and neurophysiological study of flow. ESM studies with neurophysiological measures still face severe challenges in terms of participation acceptability (how long are people going to participate, with how many daily interruptions that are to capture flow experience variance), and measurement feasibility (application of neurophysiological measurement is severely limited in daily life, given the high occurrence of measurement artefacts). Notable examples that are attempting to overcome these limitations are recent studies on daily flow experiences that make use of wearable sensors (Gaggioli et al., 2013; Rissler et al., 2018).

However, these studies focus on PNS measures as the daily approach does not yet allow to employ neural measurement. Furthermore, the large amounts of experiential and physiological intra- and inter-personal variances make these ESM-wearable sensors approaches resource intense and inefficient (only a minimal amount of intense flow experiences are likely to be collected in days or weeks). A similar limitation is also present in ENG paradigm approaches. While it could be assumed, that these approaches overcome the short observation time limitation of DM paradigms (Delle

Fave, Massimini, and Bassi, 2011), it could still be contested, that building necessary skills and orientation with a task is not as quickly achieved as the researcher might like. In that regard, it could be considered that those approaches where the repeated presentation of a simple task might be most suited to enable fast skill development and experience of intensified flow. Examples for this type of study are given in (De Kock, 2014; Kramer, 2007), where participants can repeatedly engage with the same task as to master the skill and get accustomed to the situation and the goal anchors, which might make effortless attention more likely to emerge. For this reason, this approach is herein denoted as the Mastery (MAS) paradigm. While the MAS approaches facilitate the controlled, neurophysiological measurement, they are still quasi-experimental (because no variables are manipulated - just continuous observation is performed). Also, they rely on the assumption that skill is built quickly enough to enable intensified flow in controlled settings. This assumption is still questioned, given the proposition, that intense flow experiences require the year-long mastery of complex abilities (Ullén et al., 2010).

For this reason, a particular line of research has attempted to fuse some of the aforementioned approaches (i.e. DM, ENG, MAS) through the study of experts in naturalistic tasks but controlled settings. For example, Tozman, Zhang, and Vollmeyer (2017) have studied the experiences of chess players when placed in matches against less, equally, or more skilled opponents. Thus these authors put the DM paradigm in a naturalistic, yet controlled context while being able to collect PNS based measures. In a related manner, Harmat et al. (2011) and Manzano et al. (2010) observed expert musicians in the laboratory using a variety of physiological measures and asked the participants to play well-mastered or challenging musical pieces. Giving up some more control over the setting, Jaque, Karamanukyan, and Thomson (2015) observed orchestra conductors during their natural sessions of practice using PNS measures. These approaches share the increased likelihood of intensified flow by recruiting expert performers (thus high skill levels and goal anchors are already developed before the observation) and observing them in situations that do not deviate as strongly from their natural settings. The approaches differ with regards to the degree of control/manipulation and with regards to the study domain, which has important implications for the applicable measurement and the design of the study.

Specifically, two aspects should be noted, as they importantly impact the extension of flow (neurophysiology) research in the context of KW. First, situations that require a large amount of physiological movement (e.g. musicians and orchestra conductors) are to date difficult to study using brain imaging techniques due to movement artefacts (Muthukumaraswamy, 2013). Second, all of the described examples provide situations in which participants experience the (repeated) mastery of a well-practised task (e.g. a practised musical piece or a chess game). The KW context provides an advantage for the former aspect. As KW is often completed in sedentary positions in front of computer screens, influences from movement artefacts are not generally as much of a problem, which is why the application of, for example, EEG measurement is feasible. Furthermore, the naturalistic state closely resembles the situation in experimental laboratories, which makes the study of KW in the laboratory relatively naturalistic by default. On the other hand, KW is often highly diverse, rarely with a simple way of solving a problem or repeating a task (Quinn, 2005). Therefore, to observe flow in naturalistic KW settings, a research approach needs to be developed that can balance task diversity and homogeneity to enable neurophysiological investigation. In this dissertation, one such approach is developed and the operationalisation described in detail in Chapter 5.

Controlled Experience Sampling

In summary, in an attempt to increase external validity and naturalistic character of flow laboratory research, the adaptation of the ESM (Csikszentmihalyi and Hunter, 2003) to the laboratory setting is proposed in this chapter. This adaptation signifies a controlled approach prompting individuals to work on a personalised KW task during observation with neurophysiological sensors and through a repeated interruption to "catch flow in the act". This approach of studying expert performances resonates strongly with the origin of flow theory in which the pursuit and occurrence of flow were identified as central drivers for the pursuit of excellence and fulfilling life experiences (Csikszentmihalyi, 1975). Also, the study of flow in naturalistic situations similarly adheres to the tradition of flow research to sample daily activity (Csikszentmihalyi and Hunter, 2003). Furthermore, the strengths of the cESM approach lie in covering the outlined assumptions by observing expert performers in their naturalistic tasks (and environments). In comparison to DM, MAS, or ENG paradigms, observing experts in their natural domain environment not only improves the likelihood of finding high abilities and challenges but also high task relevance. Furthermore, autonomy is typically present to a higher degree when an experimenter does not constrain the observed task. Considering the benefits and drawbacks of the observational cESM approach, they are naturally highly congruent with those inherent to the ESM method. Delle Fave, Massimini, and Bassi (2011) describes these in-depth. They point to the concluding observation, that while low in control and structure, observational approaches are particularly useful for exploratory purposes, which is here argued to be an important step to take, given that flow (neurophysiology) research is still in a nascent stage given the paucity of paradigms. Again, this is what is implied in calls from flow scholars for more creative laboratory research (Harris, Vine, and Wilson, 2017b), in particular such that includes more naturalistic situations (Rich, 2013). Furthermore, Delle Fave, Massimini, and Bassi (2011) highlight the advantages of (1) ecological validity, (2) gathering information on behaviour that can be related to external contingencies, (3) real-time assessment of experience (major advantage), and (4) repeated measurement over time.

However, while the natural observation of domain experts comes with an increase in external validity, it also represents a challenge for traditional laboratory setups (a critical trade-off between researcher control and naturalism for the participant). Specifically, by allowing for a (partially) self-selected task scenario, the variance in observed tasks increases strongly, which might introduce an increased amount of variance from extraneous variables (as compared to the highly controlled DM approaches). Also, by giving up some degree of experimentation control, the researcher might experience difficulties in observing experiential contrasts. It has been remarked, that clarity of goals is an important facilitating factor for flow induction and should be considered an important contextual feature for designing a flow task (Moller, Meier, and Wall, 2010). Similarly, it has been noted that this is a crucial feature integrated into most DM designs, given the simplistic nature of employed tasks (Moller, Meier, and Wall, 2010). However, this is a double-edged sword, as it would be expected that too much task structure could eventually be detrimental to flow induction, by restricting the range of action potential and creativity too broadly (Moller, Meier, and Wall, 2010). The cESM design falls between the extremes of this continuum through the operationalisation of initial goal-setting procedures. Finally, the strongest argument for exploring cESM is that it provides a link from laboratory to domain-specific ESM research (i.e. a linking step when flow neuroscience research will go into the field). This bridging quality could be important for future work that aims to study flow neurophysiology longitudinally - a proposed goal by Rich (2013). cESM offers this property mainly because the naturalistic experience can be structurally altered (controlled) to understand which variables might make flow experiences in the laboratory more or less likely (e.g. through manipulation of environmental stimulation, goal-setting procedures, or other flow preconditions). In this dissertation, the cESM approach is implemented as a research design in the context of KW, to evaluate how the former assumptions are met and how results from this more naturalistic approach compare to results from a more traditional paradigm like DM. Specifically, the approach is explored regarding the impact on flow experience intensities and the volatility/consistency of flow over repeated interruptions on a comparably shorter period (i.e. in comparison to daily ESM interruptions).

For the reasons mentioned above, as one important contribution of this dissertation, the neurophysiological study of flow experiences is conducted at the within-subject level in an established DM task and a novel implementation of a cESM task. Furthermore, bridges between cESM and DM are included in the second experiment of this dissertation in the form of autonomous difficulty conditions (while keeping the majority of the experimental situations and measurement instruments consistent). Lastly, in this second experiment,the intensification of flow experiences and the variety of flow inducing situations is pursued by focusing on a second approach that has received increasing attention over the recent years, namely inclusion of social interaction into experimental tasks. This approach is discussed in the next section.

3.3 Flow Intensification 2: Social Context Manipulation

> Contents of this section are in part adopted or taken from Knierim, Nadj, and Weinhardt (2019) and Knierim et al. (2019). See Section A.1 for further details.

Towards More Social Flow Research

Flow has been centrally conceptualised and studied in the domain of the individual, yet the theory recognises the potential for flow in social interaction (Ryan and Deci, 2002). In recent years, this has given rise to intensified research on flow experience in social contexts (Van den Hout, 2016). These works are particularly fuelled by the propositions, that flow in social interaction is (1) experienced more intensely (compared to flow experience in isolation), (2) has positive impacts on cognitive and affective outcomes for the individual and the social unit, and (3) even represents a qualitatively distinct, under-researched phenomenon (Walker, 2010; Tse et al., 2016; Heyne, Pavlas, and Salas, 2011; Hout, Davis, and Weggeman, 2018). Given that social interaction represents an increasingly occurring phenomenon in the context of KW (Keith et al., 2016; Wuchty, Jones, and Uzzi, 2007), the direction of studying flow in social interaction thus represents both an opportunity for increasing the internal validity of flow experiments (intensifying the experience), and the external validity (studying flow in more natural contexts). For this reason, in this dissertation, an experiment that includes the manipulation of social context SCM was pursued. To provide a foundation for this approach, this chapter summarises the latest related work. Two recent review articles summarise the emerging field of social flow research, extending across qualitative and quantitative studies and discussing conceptualisations of flow experience in social interaction (Hout, Davis, and Weggeman, 2018; Magyaródi and Oláh, 2017). While a consensus appears on the potential positive effects of flow in social interaction is present, the question of its phenomenology (the differentiation of individual and social or group-level flow) is still young (Hout, Davis, and Weggeman, 2018). Some authors agree that the individual-level experience can find aggregation at a higher level of social units (dyads, small groups) incorporating social dynamics like reciprocal individual- and group level interactions (Heyne, Pavlas, and Salas, 2011; Magyaródi and Oláh, 2017; Zumeta et al., 2016). The very least, in this instance, (social) flow experience is connected to social-level phenomena like experiences of collective absorption or a sense of unity, likely driven by implicit coordination and synchronisation processes (Hout, Davis, and Weggeman, 2018; Labonté-Lemoyne et al., 2016; Magyaródi and Oláh, 2017; Walker, 2010). On the front of flow research in social interactions, the majority of work has continued the tradition of survey-based research. Some of this research has focused on evaluating remembered social flow experiences from the past (e.g. Kaye, 2016; Zumeta et al., 2016). Other research has assessed the quality of highly engaging experiences in the field (e.g. musicians during practice - see Gaggioli et al., 2017), students during a seminar project - see Aubé, Brunelle, and Rousseau, 2014; Salanova et al., 2014, or city-wide simulation games - see Admiraal et al., 2011). Given this dissertations interest in the controlled study of flow neurophysiology in small groups, an additional review of related work was conducted, focusing on laboratory experiments of flow in social interaction.

Related Social Flow Research

Google Scholar was searched using the keywords: *"(social OR group OR collective) AND flow"* and the corpus completed through forward-backward search (Webster and Watson, 2002). Integrating with the studies discussed in the two previous reviews (Magyaródi and Oláh, 2017; Hout, Davis, and Weggeman, 2018), 32 empiric articles were identified that focus on flow experience in social contexts. From these, ten studies conducting laboratory experiments were identified.

The results of this literature are documented in Table 3.1. Particular emphasis was placed on understanding the main concepts of interests, the operationalisation of the experiments, and the types of measures being used. The identified studies majorly focused on individual-level flow experience in social contexts (indiv.), in some cases with a conception of additional social-level experience characteristics (coll.). An example for the latter notion is Walker (2010)'s argumentation that individual and social flow should be similar (being rooted in individual flow), yet that social-level factors like interdependent and reciprocal flow dynamics would add an experiential note that alters the experience of the interacting individuals. The focus on (extended) individual flow in social contexts is also visible in the employed self-report instruments. These instruments mostly collect individual flow (that is sometimes aggregated to the social level afterwards, e.g. through the computation of group means and standard deviations - see Heyne, Pavlas, and Salas, 2011). However, in some instances, the social-level flow experience is also conceptualised to be accompanied by experiences like partner unity or synchronisation and collective absorption (e.g. Magyaródi and Oláh, 2017; Walker, 2010; Labonté-Lemoyne et al., 2016).

	Keeler et al. (2015)	Magyaródi and Oláh (2017)	Tse et al. (2016)	Walker (2010) Exp. 1	Walker (2010) Exp. 2	Brom et al. (2014)	Heyne et al. (2011)	Labonté et al. 2016	Keith et al. (2014)	Keith et al. (2016)
Flow Preconditions										
I. Precondts.	x	x	x	x		x	x	x	x	x
Cooperation	x	x	x	x		-	x	x	x	x
Integration	x	x	-	x		x	x	-	x	x
Flow Accompanying Experiences										
Performance	-	x	x	-		-	x	-	x	x
Satisfaction	-	x	x	x		x	-	-	-	-
Growth	x	-	-	x		x	x	-	-	x
Study Design										
Sample Size	4	80	128	30	48	169	135	42	352	376
Group Size	4	1/2	1/2	1/2	2/4	1/2-3	3	2	4-5	4
Context	Mus.	Cogn.	Cogn.	Sport	Sport	Edu.	Cogn.	Gam.	KW, Gam.	KW, Gam.
Environment	F2F	-	F2F	F2F	F2F	F2F, DMC	F2F	F2F	F2F	F2F
Paradigm(s)	IDPM	SCM	DM, SCM	SCM	IDPM	SCM, COM; CPM	DM	MAS	TBM	SCM, IDPM, CPM
Measurement										
Flow Self-Reports	FSS (I.)	FSQ (SU.)	sFSS (I.)	ESF (I.), CS (I.)	ESF (I.), CS (I.)	FKS (I.)	sFSS (I.)	CS (I.)	CA (I.)	CA (I.)
Physiology	Oxy., ACTH	-	-	-	-	Cor.	-	EEG	-	-

Notes: I. = Individual Level; SU. = Social Unit Level; Condts. = Conditions;
Mus. = Music; Cogn. = Cognition; Edu. = Education; Gam. = Gaming;
TBM = Team Building Task Manipulation;
FSQ = Flow Synchronisation Questionnaire; ESF = Experience Sampling Form;
CS = Challenge-Skill Reports; Oxy. = Oxytocin; Cor. = Cortisol.

TABLE 3.1: Controlled Studies on Flow Experiences in Social Interaction.

Analysis of this literature corpus further highlights central variables that are repeatedly discussed as preconditions and corollaries of flow experiences in social interaction. Preconditions are described as the individual level flow preconditions (difficulty-skill balance, clear goals and feedback) by all ten studies. Two additional preconditions outline factors required for a collaborative interaction (not just mere co-presence of individuals during a task - see Walker, 2010 for emphasised discussion of this requirement). The first is a requirement for cooperation that includes factors like member interdependence, coordination of actions (e.g. through communication), and positive, supporting interaction (i.e. cooperation not competition). The second is a requirement for integration that includes well-aligned goals, challenges and abilities (e.g. similar or aligned skills), roles and procedures. Thus, integration of group member interdependence in a task and means for coordination and cooperation can be seen as essential requirements for flow in social interaction. Regarding the experiences accompanying flow in social interaction, both individual and social level variables are reported in three forms. The first is the impact of flow experience on individual and group performance (e.g. productivity or creativity, but also the sharing of information). The second is the impact on satisfaction with oneself (e.g. enjoyment or flow intensity), and the social unit (e.g. satisfaction with a group). The third outcome dimension pertains to individual and social growth, that includes knowledge building, quality of social relationships, and also the facilitation of future social flow experiences through establishment or improvement of interaction structures. For all three dimensions, mostly positive relationships with flow are reported, which highlights why flow is considered a beneficial experience in small group interactions.

Experimental research on flow in social interaction has adopted the DM paradigm (or variations of it where difficulties are continuously kept at a challenging level - e.g. Walker, 2010; Labonté-Lemoyne et al., 2016), yet has also ventured into new directions. These directions primarily pertain to the structured manipulation of interaction forms, for example by comparing the individual to the social flow experience (= SCM), by comparing outcomes in low or high interdependence (= IDPM), or by manipulation of communication (= COM) and cooperation format (= CPM). Lastly, broader manipulations of tasks have been attempted to study task-based outcome differences (e.g. Keith et al., 2014; Keith et al., 2016 provide two different team-building exercises to study flow in small groups). In summary, this shows a focus on traditional flow research paradigms DM with the extension to exploration into social unit level causes and moderators of flow experience outcomes. Another emerging aspect in this literature is a focus on dyads or small groups, which is likely explained by the increasing complexity of interactions with larger groups and the proposition, that smaller groups are more likely to experience a joint state (Walker, 2010; Armstrong, 2008). Furthermore, most of the present research has conducted their experiments in Face-to-face (F2F) settings, perhaps likely for increased ecological validity and related exploration. Similarly, these studies mostly provided equal task roles for each member and fully shared task information.

So far, barely any research has systematically studied neurophysiological phenomena of flow experience in social interaction. Notable exceptions are the articles by Keeler et al. (2015), which studies interpersonal bonding in differently interdependent and challenging scenarios (as indicated by oxytocin & adrenocorticotropic hormone secretion), the study of stress in a diverse range of interaction formats (as indicated by cortisol release) by Brom et al. (2014), and the investigation of neurophysiological state interaction during dyadic gameplay (observed in frontal EEG activity) by Labonté-Lemoyne et al. (2016). This state shows, that while there have been emerging amounts of research on flow neurophysiology on both the PNS and PNS side (see chapter 4), such work has seen almost no extension to the social interaction level.

Digitally-Mediated Social Context Manipulation

In conclusion, while considerable correlational research has been conducted, it can be stated that the controlled study of flow in small groups has only sparsely attracted scholar's attention. In particular, there appears to be a paucity addressing digitally-mediated interactions, that are essential to today's decentralised work environments. As social interaction processes deviate substantially between face-to-face (F2F) and digitally-mediated (DMC) settings (Derks, Fischer, and Bos, 2008; Chanel and Mühl, 2015), the extension of previous work on flow in groups to the digital work context represents an important research gap. Also, while there is an increasing interest to elucidate the underlying neurophysiological processes of the flow experience, there has so far been almost no related research in small group settings. Furthermore, a central observation has been so far, that flow in social interaction may be even more intense than when experiencing flow alone (Walker, 2010). This hypothesis provides an additional opportunity to assess the validity and reliability of previous neurophysiological findings that mostly stem from DM experiments. Given that social interaction represents an increasingly occurring phenomenon in the context of KW (Keith et al., 2016; Wuchty, Jones, and Uzzi, 2007), the approach of studying flow in social interaction thus represents both an opportunity for increasing the internal validity of flow experiments (intensifying the experience), and the external validity (studying flow in more natural contexts). For this reason, in this dissertation, an experiment that includes the manipulation of social context SCM was pursued.

Chapter 4

Flow Neurophysiology

4.1 Fundamentals of (Flow) Psychophysiology Research

To enable the development of adaptive NeuroIS that could in the future be able to facilitate flow experiences, a thorough understanding of the potentials of (wearable sensor-based) neurophysiological data needs to be created. It is for this reason, why the first major research goal in this dissertation focuses on the integration of the present body of knowledge on how neurophysiological data can be used to observe flow experiences (RG1). This chapter discusses the fundamental ideas and challenges of psychophysiological research and reviews the state of neurophysiological flow research in two Structured Literature Reviews (SLR). Thus, a comprehensive integrative effort is put forward that represents the basis for the selection of measurement instrumentation, the development of fitting experimental approaches and the frame of reference for the interpretation of the following experimental results.

Before detailing the state and approach to neurophysiological flow observation, a critical appraisal of the approach is useful to understand its benefits and limitations. On the one hand, researchers are eager to use neurophysiological measures due to their previously detailed benefits (low intrusiveness, automatic and continuous recording, non-participant-manipulated observation). However, their limitations are more rarely considered, and the potential of neurophysiological data is easily over-interpreted. Specifically, studies show that neurophysiological data is often falsely ascribed with superior robustness and explanatory potential (Brouwer et al., 2015). A common misconception in psychophysiological research is, that objectivity and validity of physiological measures surpass the quality of, for example, self-report measurements, or that measures from different domains (physiology, report, behaviour) converge on the same construct (Brouwer et al., 2015; Bridwell et al., 2018; Riedl, Davis, and Hevner, 2014).

However, neurophysiological (e.g. ECG or EEG) measurement is generally subject to the same psychometric considerations as other measurement types that attempt to describe psychological states (Keil et al., 2014). Nevertheless, issues such as task matching, measure reliability, measure specificity, or discriminability are often overlooked (Keil et al., 2014; Cacioppo, Tassinary, and Berntson, 2007; Riedl, Davis, and Hevner, 2014). As with any psychophysiological measure, physiological information needs to be discussed in the form of the quality of the relationship between a psychological and physiological state that can take forms as (1) outcome, (2) marker, (3) concomitant, or (4) invariant (Cacioppo, Tassinary, and Berntson, 2007) (see Figure 4.1). These possibilities mean that the relationship between an independent and a dependent variable can easily be complicated if a third variable is present that shares variance with either variable (e.g. if two subject groups are not well-matched in age, gender, handedness, or time of days of the sampling - see Keil et al., 2014). A clear one-to-one relationship between physiological and report or behaviour variables is rarely the case (Bridwell et al., 2018). It is for these reasons important, that results from neurophysiological studies are compared extensively with related literature, and that data are collected across different measurement scenarios, to establish the reliability of the observed patterns and relationships.

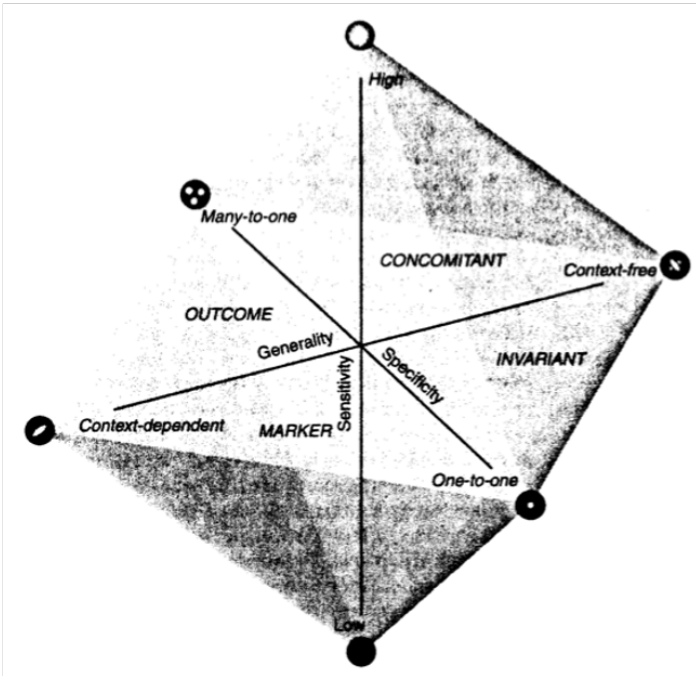

FIGURE 4.1: Taxonomy of Psychophysiological Relationships - See Cacioppo, Tassinary, and Berntson (2007).

Another important, yet often overlooked factor in neurophysiological research is the high degree of variance (both within and between individuals) present in the measured signals (Pivik et al., 1993; Bridwell et al., 2018; Brouwer et al., 2015; Müller-Putz, Riedl, and Wriessnegger, 2015). Highlighting the individuality of, for instance, one's EEG signature is the fact that such signatures are now used as a safety feature in security applications (Armstrong et al., 2015). In this research, the uniqueness of a person's neural signature is used to safeguard access to an Information Technology (IT)-system, as are nowadays fingerprints. Inter-individual variance factors include a myriad of variables like demographic backgrounds (e.g. age, gender, ethnicity) or lifestyle factors (e.g. physical activity, stress, smoking or alcohol consumption - see Valentini and Parati, 2009; Keil et al., 2014). For some measurement methods, additional specific factors have to be accounted for. For example, for cardiac activity, physiological determinants like the circadian cycle, posture, blood pressure or respiratory activity can act as influencing factors (Valentini and Parati, 2009). In another example, for electrical neural activity, factors like handedness, head shape, skull thickness, or cortical structure influence the observed data (Homan, Herman, and Purdy, 1987; Bridwell et al., 2018; Keil et al., 2014; Pivik et al., 1993). A simple example that shows inter-individual differences is age. For cardiac activity, it is known that resting heart rates are lowered on average with increases in age (Tanaka, Monahan, and Seals, 2001). For electrical neural activity, it is known that with age

Delta and Theta frequency power levels are generally reduced with increasing age across the life span (John et al., 1980). Within individuals, physiological signals have been found to vary due to influences from various variables like biochemical, metabolic, circulatory, hormonal, neuroelectric, and behavioural changes (Teplan, 2002; Pivik et al., 1993). These influences mean that the repeated measurement of a single individual can substantially vary simply by the time of day, substance intake, or situational affect (see, e.g. Harmon-Jones, Gable, and Peterson, 2010).

These variability potentials require a variety of considerations. First of all, they require that neurophysiological research attempts to overcome this variance by adequate participant samples, recording rigour, task design, and sophisticated data analysis. Ideally, samples should cover sufficient size and be sufficiently homogeneous to minimise or average out individual differences, as should task designs (e.g. provision of multiple repeated trials) (Müller-Putz, Riedl, and Wriessnegger, 2015; Brouwer et al., 2015; Picton et al., 2000; Pivik et al., 1993). Data analysis should take individual baselines into account (e.g. through the computation of change scores from rest phases or individualise features further), and could even consider computational approaches to isolate individualised features (Bridwell et al., 2018; Brouwer et al., 2015). For instance, Zhang et al. (2019) extract features personalised in the time, frequency, and spatial domain from highly-dimensional EEG data sets.

Finally, it needs to be highlighted, that neurophysiological science is presently at a critical junction of acknowledging the imprecision in the overlap of latent psychological constructs and observed physiological parameters (Bridwell et al., 2018; Riedl, Davis, and Hevner, 2014). The question of how, for example, an observed EEG pattern maps onto constructs from psychology textbooks challenges psychophysiological assessment (Buzsaki, 2006; Bridwell et al., 2018). Similar considerations have been outlined in NeuroIS research (Riedl, Davis, and Hevner, 2014). Specifically, besides the issues of how specific, sensitive, and reliable a particular kind of measure is, the issues of construct validity and convergent validity need to be emphasised as critical and fundamental properties to establish in psychophysiological research. Stated simpler: It is possible, that measures from different domains (physiology, but also different physiological systems, reports, behaviour), may capture different aspects (variance) of a construct (Riedl, Davis, and Hevner, 2014). Therefore, "researchers should not generally expect that different measures in a construct domain are substitutes; rather, in many cases, they may be complements" (Riedl, Davis, and Hevner, 2014, p. xv), or in some cases, they may also be indifferent or divergent. One of the underlying fundamental questions in psychophysiological research is, therefore, whether the data are related representatives of a construct of interest or artefacts of the chosen instruments (Riedl, Davis, and Hevner, 2014). Research indicates that to date, achieving construct validity (convergent and discriminant validity) through multimodal measurement, is often impeded by two phenomena: (1) measures capture only some part of a construct or (2) measures represent multiple constructs (see Figure 4.2) (Riedl, Davis, and Hevner, 2014; Strube and Newman, 2007). The implications of these two phenomena are, for case 1, that real, existing relationships can easily be overlooked because direct modelling might fail due to insufficient overlaps of explained variance from two measures (Riedl, Davis, and Hevner, 2014). For case 2, it is possible that third underlying variables may cause observed relationships, and that theoretic explanations derived from an observed relationship excluding such variables must not necessarily be correct (Riedl, Davis, and Hevner, 2014).

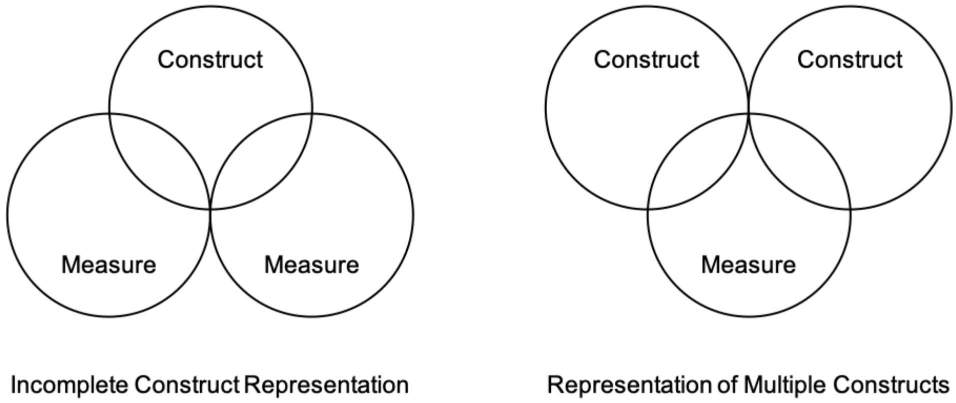

FIGURE 4.2: Construct Validity Threats - See Riedl, Davis, and Hevner (2014) and Strube and Newman (2007).

This confound means, that a central problem common to all measures is, that it is difficult to identify, how much variance of the actual construct of interest is captured, which part of the variance is captured, or if potentially multiple constructs are captured instead. This issue is further complicated, when complex (opposed to rather simple) constructs are investigated. Comparing the convergent and discriminant validity of self-report and physiological (ECG and EEG) measures for three constructs (arousal, engagement, and cognitive load) using a Multi-Trait-Multi-Method approach, Ortiz de Guinea, Titah, and Léger (2013) find that the simpler arousal construct suffers less from mono-method bias than the more complex constructs of workload and engagement (Ortiz de Guinea, Titah, and Léger, 2013; Riedl, Davis, and Hevner, 2014).

This problem of convergent validity is to date very much true for the experience of flow (an experiential composite of six individual constructs). It explains (in measurement theory terms), why so far, no simple, distinct marker of flow has been found. For this reason, following state-of-the-art recommendations on psychophysiological research methodology is of utmost importance in this line of work. These recommendations are, to clearly define the observed state and measure of ground truth, to control for confounding factors, and to define the theoretic connection between state and neurophysiology (Brouwer et al., 2015). Furthermore, it is of high importance for complex state psychophysiology research to include multivariate data and to sceptically interpret the results. These interpretations should explicitly state the underlying assumptions and should carefully distinguish current (i.e. study-specific) observation from generalizable insights (Brouwer et al., 2015; Bridwell et al., 2018; Riedl, Davis, and Hevner, 2014). Besides, the integration of findings from the measurement domains represents an essential aspect of interpreting one's results. For this reason, the extensive literature reviews later in this chapter represent a major contribution of this dissertation and are fundamental pillars for the interpretation of the experimental results.

The field of flow neurophysiology research is rather young. This observation is not only remarked in seminal articles on flow neurophysiology (Peifer, 2012; Harris, Vine, and Wilson, 2017b) but visible when reviewing the publication chronology of the articles reviewed in the course of this dissertation (articles from two SLRs that are described in the following two sections). As can be seen in Figure 4.3, despite the long-established foundation of flow theory (dating back to Csikszentmihalyi's first book in 1975), the majority of theoretic and empiric work on flow neurophysiology has been conducted rather recently. As to why there is such a large gap between the conceptualisation and flow theory, three main reasons are likely: (1) a lack of experimental paradigms to manipulate flow experience in laboratory setups (flow research initially concentrated on interviews and daily diary methods), (2) a lack of theoretical propositions on neurophysiological configurations and patterns during flow experience, and (3) limited availability or feasibility of neurophysiological measurement methods. Remedies for all three problems appeared around the first years of the current century. The Difficulty Manipulation (DM) paradigm was introduced (Rheinberg and Vollmeyer, 2003), first neurophysiological accounts of the flow experience were published (Marr, 2001; Dietrich, 2004), and the first Peripheral Nervous System (PNS) flow study was implemented by a psychology student in 2006 (Kivikangas, 2006). The latter development exemplifies the increasing availability of physiological measurement instruments. Whereas the research on PNS flow physiology slightly pre-dates the Central Nervous System (CNS) work, it appears that especially an interest in CNS observations has risen in the late years of the current decade (from 2014 to 2016).

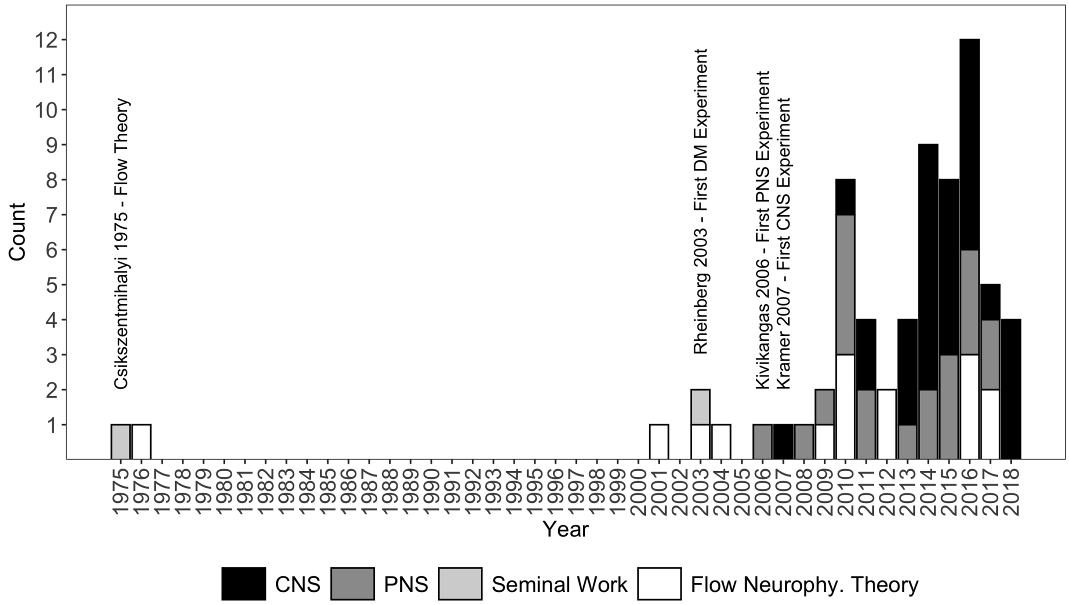

FIGURE 4.3: Overview of the Publications on Flow Neurophysiology as Identified by the Two SLRs in This Dissertation.

Given the complexity of the observation of different aspects of neurophysiological processes, the available, related literature was reviewed extensively to establish a foundation on which to build the herein used research approaches and with which to integrate the results of the experiments in this dissertation. Furthermore, given the apparent youth of the field, the importance of providing integration of knowledge became apparent. The importance of this integration was further emphasised by the results from the bibliometric analysis in the second SLR (focused on the CNS - see section 4.3), that showed a substantial lack of citations of related work. Such a lack of integration of knowledge has likely strongly impeded the progress of identifying neural configurations of flow and the theory-driven development of adaptive NeuroIS. For this reason, as a key, integrative contribution to the research on flow neurophysiology from this dissertation, the following two sections describe in-depth which findings have so far been made regarding how flow experiences are accompanied or manifested by neurophysiological processes.

4.2 Peripheral Nervous System Research

Contents of this section are in part adopted or taken from Knierim et al. (2017c). See Section A.1 for further details.

Flow Observation in the Body

Deriving from the flow experience characteristics (challenge-skill balance, clear goals, unambiguous feedback, autotelic experience, action-awareness merging, sense of control, loss of self-consciousness, transformation of time, and concentration on the task at hand - see Csikszentmihalyi, 1975; Nakamura and Csikszentmihalyi, 2009), rather recently, multiple theoretic propositions have been made about how flow is reflected in PNS and CNS activity. This dissertation integrates findings from both domains, in separate SLR studies. In a first study, the focus was placed on the PNS as related features are of heightened interest in research on adaptive NeuroIS due to high user acceptance of such measurements. Consider, for instance, the increased ubiquity and user acceptance of wrist-worn wearable sensors (e.g. smartwatches) (Seneviratne et al., 2017). Beyond this aspect, there are particularly promising hypotheses regarding the PNS configuration during flow experience, that postulate uncovered potential to detect flow with PNS features alone (Keller, 2016; Harmat et al., 2015).

The central proposition on the PNS side is, that flow represents a state of optimised physiological activation (Peifer, 2012). To investigate this proposition, an emphasis has been put on the observation of the Autonomous Nervous System (ANS), the part of the human nervous system that regulates critical physiological components such as heart rate, smooth muscles (e.g., to control eye movements), and glands (e.g., to release behaviorally relevant hormones). The ANS comprises two major branches, the activating sympathetic branch, and the calming parasympathetic branch (for more background information see Andreassi, 2000 and Section 4.4). Typically, these two branches act antagonistically (Berntson, Quigley, and Lozano, 2007), with sympathetic dominance representing increased physiological arousal and activation, and parasympathetic dominance representing increased relaxation. Knowledge of a present configuration is often derived from observing changes in the cardiovascular system, in particular, the time between adjacent heartbeats, as the heart is related and sensitive to both activities of the sympathetic and the parasympathetic ANS branches (Berntson, Quigley, and Lozano, 2007). Also, observations of hormonal changes are used for this inference (particularly cortisol as a measure of physiological activation - e.g. Peifer et al., 2014; Tozman, Zhang, and Vollmeyer, 2017). It has been repeatedly reported that flow is accompanied by increased (i.e., moderate to high) physiological activation levels (Peifer, 2012; Tozman et al., 2015; Keller et al., 2011; Ulrich, Keller, and Grön, 2016b; Klarkowski, 2016; Bian et al., 2016). Thus, to experience flow, the body must expend some amount of physiological energy (Debus et al., 2014). It follows that without some energy or effort, there is no flow.

However, it has not been established thus far how these increased energy levels are realised, a fact that holds particularly true regarding the interplay of the sympathetic and parasympathetic ANS branches. A set of contesting propositions has so far been brought forward. First, due to increased concentration on a task that is appraised as challenging but not threatening and accompanied by positive affective valence, Peifer (2012) describes flow to be reflected by optimised physiological activation (i.e., moderate peripheral arousal and elevated but moderate sympathetic activation). Second, Keller (2016) postulates that flow is an experience similar to stress resulting from intense mental effort due to high involvement in an activity that requires heightened task difficulty (i.e. high levels of sympathetic activation). Third, Manzano et al. (2010) describe flow physiology as being reflective of positive affect, increased arousal, and increased mental effort, caused by focused attention on a task. In this line of thought, Ullén et al. (2010) follow the concept of effortless attention, arguing that a physiological coping mechanism simultaneously constitutes flow. The latter refers to an increase in the relaxing activity of the parasympathetic branch of the ANS (Harmat et al., 2015; Ullén et al., 2010). Following this thought, they suggest that flow is possibly represented in the body by a rare configuration of the two ANS branches, the state of non-reciprocal co-activation that is considered to emerge when the body requires an increased precision level for calibrating the situational provision of energy in challenging situations (Berntson, Quigley, and Lozano, 2007). Lastly, Léger et al. (2014) propose that high concentration and attention in flow are reflected by a stable, less volatile state of physiological and affective activation. While all of these propositions have received some (early) support, to align the differences in these propositions, more research is needed to consolidate empirical findings into a common understanding. With the novelty of the research conducted, integrations of the emerging knowledge can provide important knowledge building catalysis and provide an answer to the following Research Question (RQ):

- **RQ1:** How can the configuration of the body during flow be described using measures of the PNS?

This work makes two key contributions to flow neurophysiology research and provides a foundation for the theory-driven development of adaptive NeuroIS. First, a systematic review and overview of existing studies that utilise PNS measurements of flow are provided. Such an overview represents a valuable starting point for future research to build on. Second, the present knowledge is integrated, and synthesised knowledge and propositions are provided on how to measure flow using PNS-related instrumentation. Importantly this integration covers findings from different research approaches and thus fosters the identification of meaningful patterns, but also limitations in the present research.

SLR Method

To address the RQ, a Structured Literature Review (SLR) was conducted following the guidelines of Kitchenham et al. (2009) and Webster and Watson (2002). Overall, the review was subdivided into plan, conduct, and report stages (see Figure 4.4). Web of Science and Scopus (Bandara et al., 2015; Hamari, Koivisto, and Sarsa, 2014) were searched with the search string:

"(flow OR cognitive engagement OR cognitive absorption) AND (physiological signal OR psychophysiology OR neurophysiology)"*

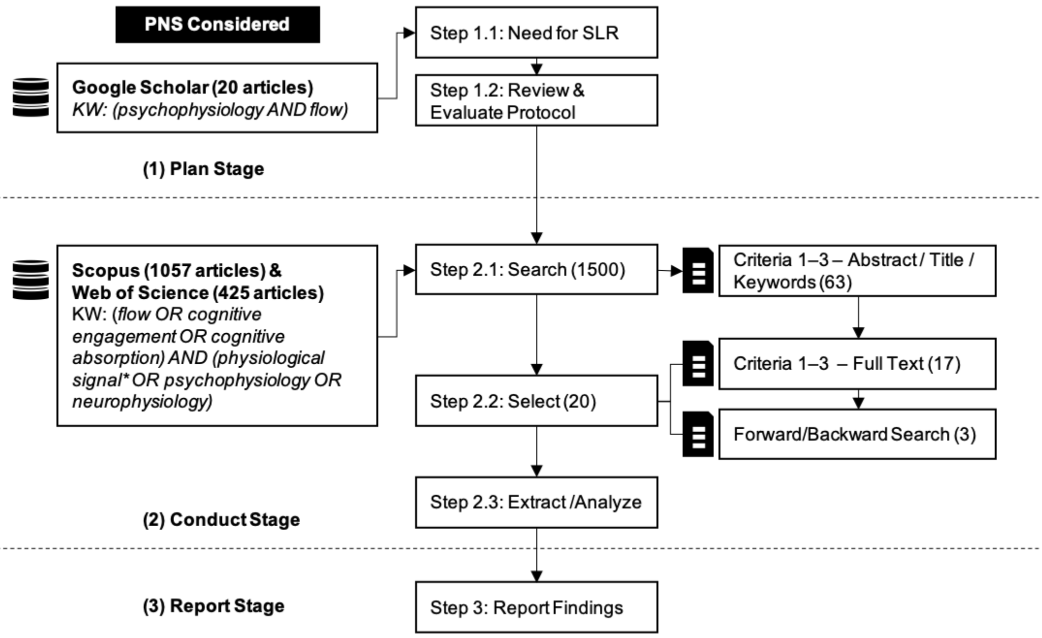

FIGURE 4.4: Stages of the SLR on Flow PNS Observation.

This search string was developed in five steps. First, an exploratory search was conducted using Google Scholar with the search term "psychophysiology AND flow". Second, the first 20 search results were reviewed, and six highly cited studies were identified: Peifer (2012), Manzano et al. (2010), Mauri et al. (2011), Tozman et al. (2015), Kivikangas (2006), and Nacke and Lindley (2008). Third, the full text of these six articles was reviewed, and the terms "neurophysiology" and "physiological signal(s)" were extracted as highly relevant to the research question. Fourth, "cognitive engagement" and "cognitive absorption" were identified as relevant flow derivations. Finally, Boolean operators were used to creating the final search string. To ensure a holistic search, the search was not limited to a specific time period. All studies that met the following criteria were included: The study (1) contains an empirical component, (2) is a peer-reviewed journal article, article in press, in conference proceedings, or book chapter, (3) refers to the psychological phenomenon of flow, (4) focuses on the PNS. The selection criteria were first applied to the abstract, title, and keyword section (excluding 1437 studies). In a further attempt, the criteria were applied to the full text of remaining studies (excluding 46 studies). Finally, a forward and backward search based on the remaining 17 studies was conducted. Thereby, another three relevant studies were identified. Overall the SLR identified 20 relevant studies.

Results

In line with a concept-centric focus (Webster and Watson, 2002), the findings are summarised and split into three main areas: (1) experiment design parameters, (2) theoretical perspectives, and (3) findings on physiological features (see Table 4.1).

To condense the findings and their relationship to flow, a dedicated syntax was developed. As an explanation of its meaning, consider the first study in Table 4.1 by Harmat et al. (2015) as an example. The authors conducted an experiment using the digital game Tetris and evaluated flow with a subset of the Short Version of the Flow Short Scale (sFSS). Insignificant relationships between Heart Rate (HR) and the self-report scale were found (•). The dot thus represents the indication of an absent relationship. Moreover, a positive linear relationship between Thoracic Respiratory Depth (tRD) and the flow self-report scale was found (⟋). A line with two dots (one at the left and one at the right end) thus represents a result from direct modelling of the relationship, and the slope of the line represents the direction of the relationship. As another example, consider the work by Ulrich, Keller, and Grön (2016b). These authors used a mental arithmetic task and found an inverted U-shaped relationship (⌃) between Skin Conductance Level (SCL) and difficulty-manipulated task conditions termed boredom, fit, and overload (B/F/O), in which flow experience was most intense in the fit (F) condition. The line with three dots thus represents the findings across the experiment conditions, and the relationship to flow can be identified from the indirect comparison of the reported flow patterns and the physiological patterns. In the presented studies, conditions denoted as fit (F) or as task (t) are those that showed the highest reported flow intensities. Partial findings like the inverted U-shaped relationship between Low Frequency Heart Rate Variability (LF-HRV) with flow reports in the first half of the experiment by Peifer et al. (2014) are denoted with an asterisk (⌃*).

Within the experimental designs, central variables of interest are sample sizes, flow induction tasks, and dependent variables/measures. In the literature corpus, sample sizes vary strongly across studies ranging from seven to 77 experiment participants. The sample counts include reported, usable observations only. The majority of studies in this SLR (14/20) used games as flow induction tasks. This observation is vital as designing tasks that reliably induce flow states is still a major challenge in flow research (Tozman and Peifer, 2016). In this regard, game paradigms have been criticised as to not sufficiently induce straining experiences (Peifer et al., 2014). Depending on research goals (e.g., in the case of separating flow from stress experiences), this spectrum is critical for flow research. Utile alternatives include high involvement tasks with naturalistic components (e.g. observing chess players like in Tozman, Zhang, and Vollmeyer (2017)). Third, dependent variables differ in two operationalisation formats, that are self-reports (14/20) and experiment conditions (8/20), with some studies utilising both (4/20). Conditions are most often differentiated along the dimension of task difficulty. In total, nine different self-report instruments were used in 15 studies to measure flow.

This category refers to the theoretical propositions in the articles of how flow can be differentiated physiologically (e.g. from stress). Perspectives and studies are ordered in terms of diagnosticity (i.e., how proposed physiology patterns are to isolate flow from other states) (Riedl, Davis, and Hevner, 2014). While increased peripheral physiological arousal is a common denominator in both flow and strain (Peifer, 2012), four central distinguishing patterns are described towards flow: (1) modulation of arousal by relaxing influences (Mdl. Relax), (2) moderate instead of high levels of arousal (Mod. Activ.), (3) stable, less volatile arousal (Sta. Activ.), and (4) the simultaneous presence of arousal and positive affect (Pos. Affect). The abbreviations in parentheses refer to how these perspectives are denoted in Table 4.1. The first perspective is currently the only characterisation that sufficiently distinguishes flow by the phenomenon of non-reciprocal co-activation of sympathetic and parasympathetic branches of the ANS. Three types of studies were derived. The first category (high diagnosticity - sufficiency condition fulfilled) proposes distinct

Articles	N	Context	DV	Mdl. Relx.	Mod. Act.	Sta. Act.	Pos. Aff.	No Dist.	HR	IBI	SDNN	RMSSD	LF-HRV	HF-HRV	LF / HF	Tot. Power	FWHM	Min.PWA	Max.PWA	PRNG	SCL	SD SCL	RR	tRD	aRD	RCY	Cor. Lvl.	Cor. Rea.	ZM	CS	OO	Pup. Dil.
				Theor. View					**Heart**								**Pulse**				**EDA**		**Resp.**				**Hor.**		**Face**			
High Diagnosticity																																
Harmat et al. (2015)	77	Gam.	sFSS	x					•				↘	•									•	↗	•							
Murch et al. (2017)	74	Gambl.	fGEQ14	x										•																		
			R/T WS						•					↘																		
Peifer et al. (2014)	20	Edu.	FKS Ab.	x	x								∧	↗													∧					
			FKS Fl.										∧	•													•					
Tozman et al. (2015)	15	Gam.	B FKS	x	x								•	•																		
			F FKS										↘	•																		
			O FKS										∧	∧																		
			B/F/O WS										↘	↘																		
Manzano et al. (2010)	17	Mus.	sFSS	x			x			↘				↘	↗	↗	•	↗	•	•			•	↗		•			↗	•		
Bian et al. (2016)	34	Gam.	FKS	x	x		x		↗	↘	∧		∧	∧	•								↗	•					•			
Moderate Diagnosticity																																
Léger et al. (2014)	34	Edu.	CA		x	x			↘		↗											↘										
Tozman et al. (2017)	57	Gam.	FKS Ab.		x	x																						∧				
			FKS Fl.																									•				
			B/F/O BS																									∨				
Peifer et al. (2015)	61	Gam.	FKS Ab.		x																						↘					
			FKS Fl.																								↘					
Gaggioli et al. (2013)	15	Life	Ch-Sk.+Pos.		x				↗						↗																	
Kivikangas (2006)	32	Gam.	FSS				x														•								•	↘	•	
Mauri et al. (2011)	30	Soc.-N.	R/T/S WS				x			↘											↗					↗				↗		∨

Experiment Design				Theor. View					Heart								Pulse				EDA		Resp.				Hor.		Face			
Articles	N	Context	DV	Mdl. Relx.	Mod. Act.	Sta. Act.	Pos. Aff.	No Dist.	HR	IBI	SDNN	RMSSD	LF-HRV	HF-HRV	LF / HF	Tot. Power	FWHM	Min.PWA	Max.PWA	PRNG	SCL	SD SCL	RR	tRD	aRD	RCY	Cor. Lvl.	Cor. Rea.	ZM	CS	OO	Pup. Dil.
Low Diagnosticity																																
Keller et al. (2011)	7	Know.	B/F/O WS					x				∨																				
	61	Gam.	B/F/O BS																								↗					
Ulrich et al. (2016)	22	Cogn.	FI					x													∧											
			B/F/O WS																		∧											
Nacke et al. (2010)	29	Gam.	fGEQ36																		↘								↗	•	•	
Drachen et al. (2009; 2010)	16	Gam.	fGEQ14						↘												•											
Nache and Lindley (2008; 2010)	25	Gam.	B/F/I WS																		∧								∧	•	∧	
Klarkowski (2016)	61	Gam.	B/F/O WS																		↗											

Notes: Theor. = Theoretic; Resp. = Respiration; Hor. = Hormones; Mdl. Relx. = Modulation of Activation by Relaxing Influence; Mod. Act. = Moderate Physiological Activation; Sta. Act. = Stable Physiological Activation; Pos. Aff. = Concurrent Positive Affect; No Dist. = No Flow Distinction Proposed; Cor. = Cortisol; Lvl. = Level; Rea. = Reactivity; Pup. = Pupil; Dil. = Dilation; Gam. = Gaming; Gambl. = Gambling; Edu. = Education; Mus. = Music; Soc.-N. = Social Network; Know. = Knowledge; Cogn. = Cognition; Ch-Sk.+Pos. = Combined Indicators Challenge-Skill Balance and Positive Emotion; B/F/O Boredom/Fit/Overload Difficulty Conditions; R/T/S Relaxation/Task/Stress; B/F/I Boredom/Flow/Immersion Conditions; Condition With Highest Flow Intensity is Underlined.

TABLE 4.1: Overview of Flow PNS Physiology Findings.

physiological signatures by investigating arousal modulation by relaxation (6/20). The second category (moderate diagnosticity - necessity condition fulfilled) proposes indicative physiological signatures (6/20). The third category (low diagnosticity) proposes either indistinct physiological signatures (2/20), or do not include hypotheses towards flow physiology specifically (6/20).

This category captures as variables the type of observed physiological parameter (cardiac, pulse, electrodermal, respiration, hormonal, facial muscle, and pupillary reactions) and the resulting physiological findings in terms of derived features (actual metrics). In terms of methods, it can be stated that cardiac features are used most often (12/20), especially in the class of higher diagnostic studies (6/6). This focus is probably due to the property of the cardiovascular system to reflect both sympathetic and parasympathetic activation (Berntson, Quigley, and Lozano, 2007). Therefore, distinguishing flow is enabled by comparison of arousal levels, arousal variability or the isolated activity of parasympathetic activation. Electrodermal Activation (EDA) is the second-most used feature (10/20), albeit mainly in studies with lower diagnosticity (7/8), which is surprising given the theoretical propositions and the property of the EDA metric to be a highly diagnostic indicator of sympathetic ANS branch activation. Electromyography (EMG) measures are used mainly in valence-related studies across classes (7/20). Together, support has been found for all four theoretical propositions, with (1) being mainly related to the sympathetic and parasympathetic ANS activity (HF-HRV, tRD), (2) being most often related to moderate cortisol (CoLe), skin conductance (SCL) and HRV levels (Total HRV and LF-HRV), (3) being related to skin conductance and hormonal level reactivity, and (4) being most often related to increased facial muscle activity (ZM).

Discussion & Future Directions

RQ1 asked, how the configuration of the body during flow can be described using measures of the PNS. This SLR identified four central approaches to the physiological measurement of flow. All include increased levels of arousal, yet vary in their explanation to how arousal states differ from straining experiences such as stress. Of these four, three fulfil only necessity conditions to distinguish flow. The proposition of a non-reciprocal co-activation of the sympathetic and parasympathetic nervous system in flow (Harmat et al., 2015; Ullén et al., 2010) also fulfils sufficiency conditions. For this proposition, however, so far only partial support has been found (Manzano et al., 2010; Bian et al., 2016) and some research has also not found support for this proposition (Harmat et al., 2015). This suggests that more research is required to understand if this proposition has merit, specifically such research that employs unconfounded indicators of activity in both ANS branches with high temporal resolution. The use of for example ECG measures alone is not sufficient for this purpose, as it only allows to infer isolated parasympathetic activity (and includes influences from sympathetic activation in some features) (Malik et al., 1996; Berntson, Quigley, and Lozano, 2007). Therefore, future research that aims to investigate this proposition ought to include multiple measurement methods. For example, combinations of ECG data with impedance cardiography features have been recommended, as the latter can well identify sympathetic activation through observation of the duration of pre-ejection periods of the heart (Tozman and Peifer, 2016). An alternative approach might be to combine ECG and EDA metrics, as the EDA signal is well known as an indicator of sympathetic activity (Boucsein, 2012). Interestingly, these two methods have so far not been utilised together, which is potentially due to a lack of knowledge about these theoretical propositions and integrations of related work. Therefore this SLR might be providing a first, clear overview over a central gap that can be addressed by flow PNS observation. Flow researchers can uniquely contribute to the state of knowledge by advancing the line of research on these diagnostically higher perspectives. Furthermore, an alternative option is highlighted, which is the increase in diagnosticity through a combination of measurement and directions (1)-(4). Support has been found for all directions through different physiological features, which is why it is advisable to integrate multiple propositions in further investigations. An exemplary approach in this direction is reported by Bian et al. (2016). However, as the inclusion of various measurement methods requires additional domain expertise (Brouwer et al., 2015), it is herein recommended that future work focuses on a select set of sparse measures (e.g. 2-3) to enable deepened insight that is critically needed at this stage of research.

Through cross-study integration of the present results, it appears that mainly the aforementioned direction (studying sympathetic and parasympathetic activation patterns) is of the highest value. This recommendation is not to say that others do not have merit, but that the findings and related theoretic accounts do not yet show as clear a path forward. Relating to the time dynamics of PNS measures of flow (the proposed stability of physiological activation during flow), a central challenge is that some identified features (especially HRV metrics) are only robustly usable after aggregations to longer periods (i.e. five minutes). Such a requirement can quickly extend the duration of experiments and make them difficult to realise for participants and experimenters. Relating to the affective dynamics of PNS measures of flow, there is presently still only the observation of facial muscles known as a robust indicator of affective experience (Riedl and Léger, 2016). Results from related studies indicate a mixed picture in terms of which affective experience might be present during flow. As there is still lively debate about whether or not flow is experienced as positively valenced or as a state of neutral affect (due to a lack of self-evaluative thoughts - see, e.g. Engeser and Schiepe-Tiska, 2012), it is unclear how an observation of affect may inform a physiological picture of flow. However, it should be noted, that results have so far not been presented for these measures in the more established experimental flow paradigm of DM. Therefore, future research might want to follow up on this path to narrow the understanding of how affect is expressed through facial muscles. If such a direction is pursued, additional measures of affective orientation, for example in the form of observing asymmetries in frontal brain activation (see Harmon-Jones, Gable, and Peterson, 2010) could be a valuable addition.

Independent of the pursued measurement instruments and theories, a central limitation to the present flow PNS research became apparent, that is a salient focus on game tasks, on DM paradigms, and importantly, on single observation scenarios. While on the one hand, this provides a shared foundation to integrate results, this state of work comes with the central downside of not knowing how transferable findings are to other situations, and how strong identified relationships are in general. Highlighting this state of the research represents an important contribution of this SLR. It is therefore adamant that future research starts to investigate flow physiology in other task domains (e.g. knowledge work or e-Sports that lend themselves to physiological measurement), using other flow elicitation paradigms (e.g. more naturalistic task settings or flow requirement manipulations) and importantly, comparisons of tasks and paradigms for the identification of task-independent flow-physiology relationships. In following these propositions, flow research can benefit from finding means of increased objective validity in flow measurement and also advance constructivist efforts to facilitate flow through adaptive NeuroIS.

Conclusion

This SLR summarised the work from 20 studies on Peripheral Nervous System (PNS) observations during flow. One of the main results is that a need for more cross-task and cross-paradigm research is present in this body of work. In particular, this need is based on the observation of mostly incoherent physiological feature relationships with flow experiences. Cross-situational research can, therefore, help to elaborate on more robust findings on the PNS configurations during flow. The SLR provides a starting point to pursue these directions and therefore contributes to flow neurophysiology research and to the development of theory-driven adaptive NeuroIS. As a particularly promising direction, the observation of physiological activation through cardiac activity patterns emerged. It has been repeatedly found that the flow experience is accompanied by increased (i.e., moderate or high) physiological activation levels. Thus, there is the assertion that to experience flow, the body must expend some amount of physiological energy (Debus et al., 2014). However, it has so far not been established how exactly these increased energy levels manifest. Considering that HRV markers reflect sympathetic and parasympathetic modulation of the heart, several propositions are presently being discussed, that can be well advanced by focusing on cardiac observation across multiple tasks. To reiterate, these propositions refer to how (or if) flow can be differentiated from other experiences (e.g. stress) through PNS measures. Whether or not flow is a state of very high physiological activation (which would be apparent from strongly reduced levels of HRV - e.g. in balanced difficulty compared to hard tasks), of moderate physiological activation (which would be apparent by only moderately reduced HRV levels), or whether or not flow is a state of both high activation and strong calming influences at the same time (as could, for example, be indicated partially by elevated activity of parasympathetic HRV indicators) represents a central question in the current state of flow neurophysiology work. To follow up on this central line of research, it was therefore decided to utilise HRV measures for flow observation that was considered to take place across multiple tasks and experimental paradigms.

4.3 Central Nervous System Research

Contents of this section are in part adopted or taken from Knierim et al. (2018b). See Section A.1 for further details.

Flow Observation in the Brain

While research on the psychology of flow has a rich history, the neural mechanisms behind it have only more recently become subject of theories and empiric studies (Cheron, 2016; Dietrich, 2004; Harris, Vine, and Wilson, 2017b; Peifer, 2012; Weber et al., 2009). The general proposition that has so far emerged is that the corollary of a phenomenologically distinct state as flow should be a unique neurophysiological state (Cheron, 2016; Weber et al., 2009). The present state of flow neuroimaging research can be described in three categories that are neurophysics (functional principles the brain operates by), neuroanatomy (regions and networks of the brain that are involved), and neurocognition (neural activity correlating with cognitive processes). To investigate these propositions, the behaviour of the firing of neuron cells in the brain is observed. In particular, this activity is observed in the cerebrum, the largest part of the brain that contains the cerebral cortex and several subcortical structures, like the hippocampus, basal ganglia, and amygdala (for more background information see Andreassi, 2000 and Section 4.4). Both indirect and direct measures of neuronal activity are used to study flow experiences in the brain. Hemodynamic imaging methods like functional Magnetic Resonance Imaging (fMRI) and functional Near-Infrared Spectroscopy (fNIRS) measure blood flow to brain regions and, therefore, indirectly index in which brain structures neural activity increases (Riedl and Léger, 2016). Electrophysiological imaging methods on the other hand measure direct electrical activity of neuron assemblies in the outer brain regions (the cerebral cortex) (Andreassi, 2000; Cohen, 2017). So far, two central neurophysical mechanisms have been proposed to explain the experience of flow. The first is the reduction of neural activity (reductionist theories - see Dietrich, 2003; Marr, 2001; Peifer, 2012), and the second is the interaction/synchronisation of neural activity (interactionist theories - see Harris, Vine, and Wilson, 2017b; Weber et al., 2009).

The first, central perspective on flow neurophysiology rests on the notion that in flow, task-irrelevant activities are down-regulated for the benefit of the brain to operate without interference at a highly efficient level (Marr, 2001; Peifer, 2012). Arne Dietrich (2003; 2004), proposed a theory based on this reasoning named Transient Hypofrontality Theory (THT). The theory states, that while in flow, activity in frontal brain regions is reduced in favour of concentration of resources in regions of the brain dedicated to processing the task at hand (e.g. from explicit towards implicit information processing in the motor or sensory regions). THT is intuitively plausible, especially so within the sports research context, as learned motor behaviour would show better execution when free from the interference of conscious control (Dietrich et al., 2010; Harris, Vine, and Wilson, 2017b). However, THT has been criticised for being overly simplistic and for neglecting critical concepts related to attentional processes (ease of attentional control, lack of attentional effort, absence of self-referential attention) (Harris, Vine, and Wilson, 2017b; Peifer, 2012; Sadlo, 2016; Weber et al., 2009). This critique has been supported by neuroimaging studies investigating the role of prefrontal brain structures in flow. This research shows a much more nuanced pattern of activation in areas of the Prefrontal Cortex (PFC), specifically reduced activity in the medial PFC (Ulrich, Keller, and Grön, 2016a; Ulrich et al., 2014; Barros et al., 2018), increased activity in dorso- & ventrolateral PFC (Yoshida et al., 2014; Barros et al., 2018), but also no reduction in frontal activity (Harmat et al., 2015). These findings have led researchers to suggest alternative neural patterns while keeping with the paradigm of reduction. One such approach extends the THT by integrating the aforementioned findings of nuanced PFC activity with research on the neurophysiology of self-awareness, particularly with activity in the so-called Default Mode Network (DMN) (Harris, Vine, and Wilson, 2017b; Sadlo, 2016). Activity in the DMN is related to mind-wandering, and thinking about the self, past, and future, and is reduced in goal-directed behaviours (Buckner, Andrews-Hanna, and Schacter, 2008). Conversely, reductions in the DMN could explain the experience of reduced self-awareness, the improvement in performance (as the "inner critic" is silenced, more efficient information processing could take place), but also the rewarding experience (as apparently, it is very pleasing to forget oneself and one's troubles) (Harris, Vine, and Wilson, 2017b; Peifer, 2012; Sadlo, 2016).

Theories from the interactionist view propose the central importance of synchronised interaction of attention and reward networks of the brain as an alternative explanation of flow neurophysiology (Harris, Vine, and Wilson, 2017b; Weber et al., 2009). The first proposition in this direction called Synchronization Theory (ST), conceptualises flow as a qualitatively discrete, emergent phenomenon resulting from the synchronised firing of neuronal networks (Weber et al., 2009). With synchronisation as an energetically cheap principle (Buzsáki and Draguhn, 2004; Siegel, Donner, and Engel, 2012), the resulting energetic optimisation is supposed to explain the perception of effortlessness during flow (Harris, Vine, and Wilson, 2017b). In the interactionist theories, attentional phenomena are at the centre of flow theory. Additional to the previously discussed phenomena of automaticity (PFC, verbal-analytic reasoning & motor region areas), and self-referential attention (DMN), Harris, Vine, and Wilson (2017b) integrate mechanisms that would explain how attentional control, impulse control, and conflict monitoring mechanisms would jointly explain the experiential components of flow (Harris, Vine, and Wilson, 2017b). In summary, they propose that efficient attention and automated action control account for many of the experiential components of flow theory, with attention during flow being considered as more external, less self-conscious, less prone to distraction and more task-directed. The neural underpinnings of these attentional processes would be manifested by a combination of reduced medial prefrontal areas and the DMN (reduced self-awareness), reduced conflict monitoring (reduced activity in the anterior cingulate cortex - see Klasen et al., 2011; Ulrich, Keller, and Grön, 2016b), improved impulse control related to dopamine activity, and considerable activity in networks related to higher-order processing (e.g. the multiple demand network) (Harris, Vine, and Wilson, 2017b).

The borders between the two theoretic perspectives are somewhat dotted. As the research into this area is rather young, it might very well be, that there is even more overlap between some of these propositions that a holistic theory of flow neurophysiology will have to account for. Taken together, it appears that so far, reasoning and evidence mostly document physical, anatomical, and cognitive patterns, that still focus on a lack of interference of task-irrelevant processes and increased neural efficiency during flow experiences. However, reviewing this initial literature on flow neurophysiology theory also highlighted that much of the reasoning is based on studies using hemodynamic imaging (see, e.g. Ulrich et al., 2014; Harmat et al., 2015; Barros et al., 2018), and that emerging EEG literature is only sparsely integrated into this research domain. If however, neural measurement is supposed to be leveraged eventually for adaptive NeuroIS, neuroimaging methods that are usable in different situations (i.e. portable) must be utilised. Presently, only the EEG provides such potential (Blankertz et al., 2016; Lance et al., 2012). Therefore, as previous work has focused on fundamental descriptions of flow neurophysiology, the integration of EEG literature is missing. To provide a foundation for the observation of flow using wearable EEG, it was decided to conduct an SLR to answer the following RQ:

- **RQ2:** How can the configuration of the brain during flow be described using EEG measures?

As will be shown later, the EEG results so far only relate to reductionist theories (i.e. THT-related reasoning) and not to interactionist theories (i.e. ST), as there are no studies that observe neural networks (e.g. using connectivity analyses). Therefore, it was decided to focus in particular on the THT related findings (i.e. frontal brain regions), to consolidate and improve the knowledge in this area by a more nuanced investigation. This work makes two key contributions to flow neurophysiology research and provides a foundation for the theory-driven development of adaptive NeuroIS. First, a systematic review and overview of existing studies that utilise EEG measurements of flow are provided. Such an overview represents a valuable starting point for future research to build on. Second, the present knowledge is integrated and synthesised knowledge, and propositions are provided on how to observe flow using EEG-related instrumentation. Importantly this integration covers findings from different research approaches and thus fosters the identification of meaningful patterns, but also limitations in the present research.

SLR Method

Similar to the previous section, this SLR was conducted according to the guidelines of Kitchenham et al. (2009) and Webster and Watson (2002) (see Figure 4.5). The search strings were developed in five steps. First, an exploratory search on Google Scholar (scanning all 50 returned search pages) was conducted using a search string of synonyms of flow and related terms, combined with often occurring synonyms for psychophysiological research approaches. 103 articles were identified by title and abstract as potentially related to flow psychophysiology (PNS and CNS), of which 67 were confirmed to be after full-text review (excluded articles had no direct relation to flow theory or measurement). At this stage, it was decided to narrow the focus on EEG studies in favour of deepened analysis. Therefore, a preliminary corpus of 18 EEG studies of flow was extracted (using the criteria 1-3 that are outlined below). In the next step, databases for the main search were selected. Twelve databases were compared as to how well the preliminary corpus was represented (WebOfScience, ACM Digital Library, EBSCOhost, Scopus, PsychINFO, IEEExplore, ScienceDirect, SpringerLink, WileyOnline, ProQuest, EmeraldInsight and JSTOR). The top two databases (in terms of the number of retrieved studies from the preliminary corpus), Scopus (13/18) and WebOfScience (9/18) were then chosen to conduct the subsequent literature search. The search string was further refined to capture flow synonyms or related concepts, to focus on capturing EEG-related research, and to exclude studies focusing on alternative contexts of the term flow or studies of other than healthy adult populations. In total, 3.167 articles were returned from the two databases.

In the first decision stage, all studies that met the following (primary) criteria were included: The study (1) is a peer-reviewed journal article, an article in press, in conference proceedings, dissertation, or book chapter, (2) refers to the psychological phenomenon of flow, and (3) the study focuses on psychophysiological measurements, specifically EEG. Duplicate articles were identified in the database results and not additionally reviewed. Afterwards, to reduce the still considerable corpus size, article titles were scanned for their relation to the review context, by an automatic and a manual process. First, all titles not including either one of the main keywords (flow, “optimal experience”, engagement, absorption, immersion, psychophysiologic*, neurophysiologic*, electroencephalo*, EEG) nor more than one of these additional keywords (intrinsic, motiv*, involv*, attent*, creativ*, distract*, chall*, adapt*, learn*, perform*, *physio*, brain) were excluded (1.964). Then, all titles were manually reviewed, and articles excluded if they were found to be unrelated to the review context (1.077). For the remaining sample (126 remaining), the abstracts were screened

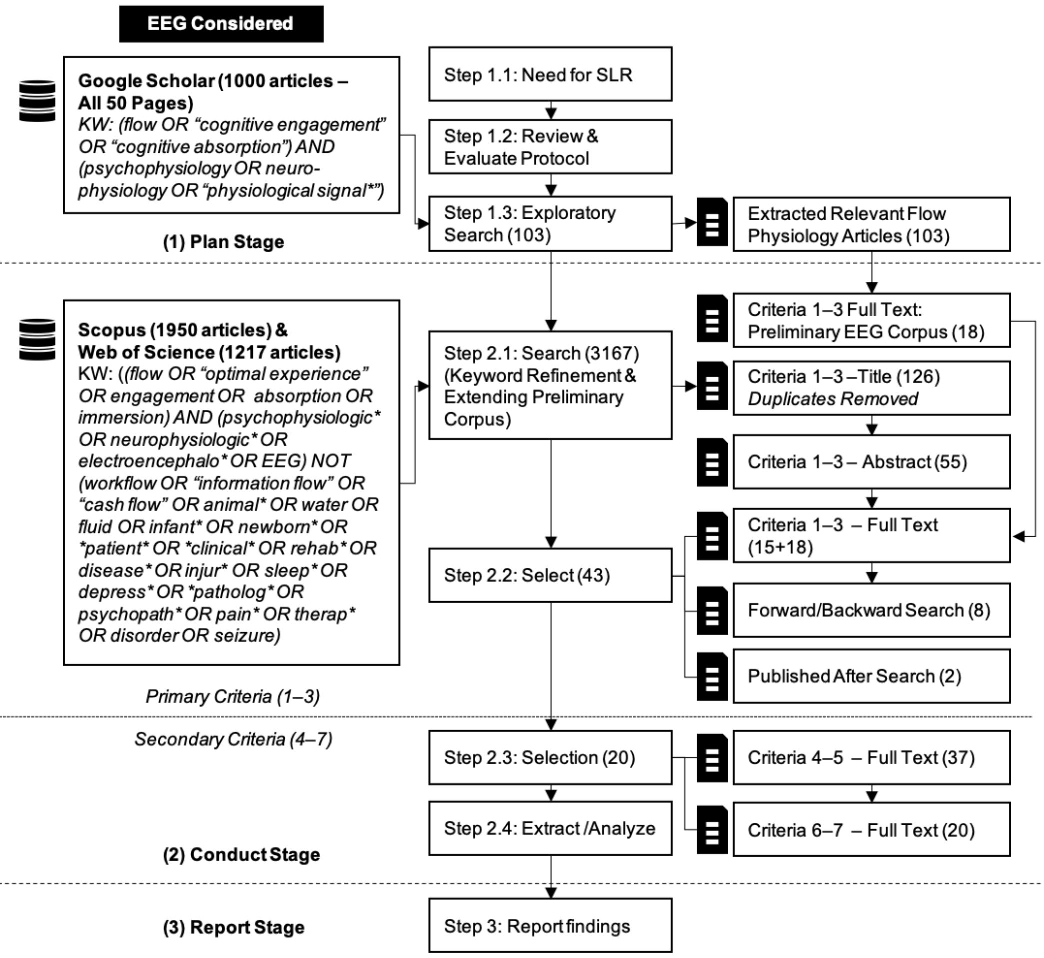

FIGURE 4.5: Stages of the SLR on EEG Observation of Flow.

manually and excluded in adherence to criteria 1-3 as far as applicable (71). For the remaining 55 studies, the full texts were analysed, and all remaining studies excluded using criteria 1-3 (40). The 15 retained articles were combined with the 18 articles from the preliminary corpus and subjected to a backward and forward search process using Google Scholar. This process identified eight additional studies of interest, concluding this first decision stage. After the conclusion of the search, two studies were published that were also added to the corpus. Therefore, for the literature analysis, a broad corpus of 43 studies was retained initially.

During the second and final stage, additional decisions were made that led to further study exclusion. Four (secondary) criteria were considered that should further improve the comparability of the corpus. First, it was decided to focus on studies containing (4) completed empirical analyses (three excluded), and (5) a focus on flow state observation, not trait variables or Brain Computer Interface (BCI) input (three excluded). Second, studies that (6) did not report sufficiently detailed information on measures and data analysis (four excluded as electrode positions were not reported; six excluded as proprietary EEG feature algorithms were used and thus feature extraction information was not reported; one excluded as the role of EEG data in pooled Machine Learning (ML) models is not reported; two excluded as it is not reported how ground truth values were derived; one excluded that does not report main flow-related condition comparisons), and (7) did not use any other, but physiological measures to provide a ground truth observation of flow were removed due to the heightened danger of ungrounded physiological inference and circular reasoning (Richter and Slade, 2017; Brouwer et al., 2015) (three excluded). The search led to a corpus of 20 publications (22 studies). Articles discussing flow neurophysiology from theoretical perspectives or using other neuroimaging methods were also retained for bibliometric analysis.

Results

Following a concept-centric focus (Webster and Watson, 2002), the findings are summarised in three areas: (1) bibliometric status of the body of work, (2) theoretical perspectives and research methods, and (3) findings on EEG features.

A short bibliometric analysis (Aria and Cuccurullo, 2017; Börner, Chen, and Boyack, 2005; Zupic and Čater, 2015) was included to comprehend the state of EEG flow research better. In particular, the integrative character of EEG research within itself, but also with other neuroimaging studies and flow neuro-theory was investigated. References of all articles in the final EEG literature corpus were extracted once manually and once automatically using the programming library CERMINE (Tkaczyk et al., 2015). Bibliometric results are reported in the form of an historiograph in Figure 4.6 that shows the publication time course and centrality measures (in-degree and out-degree) (Börner, Chen, and Boyack, 2005). It can be seen in the historiograph that the majority of EEG studies (75%) has been published in the last five years, demonstrating the youth and emergence of this approach. The results of the reference analysis (within the EEG corpus) show a strongly fragmented research field. Many, and even recent, articles show degree counts without or very little citation of other EEG studies in this corpus (in-degree of <= 1 for 90%, out-degree of <= 1 for 80% of studies). Other neuroimaging work is referenced even less, only by five studies at all, and central tendencies for all studies are close to zero (min = 0, mean = 0.35, median = 0, max = 2, out-degrees of <= 1 for 90% of studies). The integration metrics for the EEG studies with theoretic accounts are slightly higher (min = 0, mean = 0.65, median = 0, max = 4, out-degrees of <= 1 for 85% of studies), yet the majority of citations falls to three articles. It is therefore observed that the integration with theoretical work on flow neurophysiology is also mostly absent in studies of EEG patterns during flow. Altogether, this points to a shortcoming of the present literature in terms of integration of reasoning and evidence, highlighting the need for this SLR and in general, more integrative work in flow EEG research.

As the first part of the corpus analysis, utilised theories, research methods, and analysis foci were inspected to provide context for the comparability of study results. Table 4.2 provides an overview of these results. Four publications (only 20% of publications) refer to major flow neurophysiology theories. This distribution shows that little explicit EEG study of major theories is present. Of the major theories, only Transient Hypofrontality Theory (THT), but not Synchronization Theory (ST) is investigated. Furthermore, 11 studies do not include any explicit theory or hypotheses (55% of publications). This further highlights a problematic lack of theoretic integration. A possible cause for this distribution is the high number of mainly explorative efforts, but also that some studies only peripherally investigate flow experiences. Within the alternative hypotheses, one is close to a major theory (NE = Neural Efficiency). In contrast, others are more distant to major theories and cluster with either

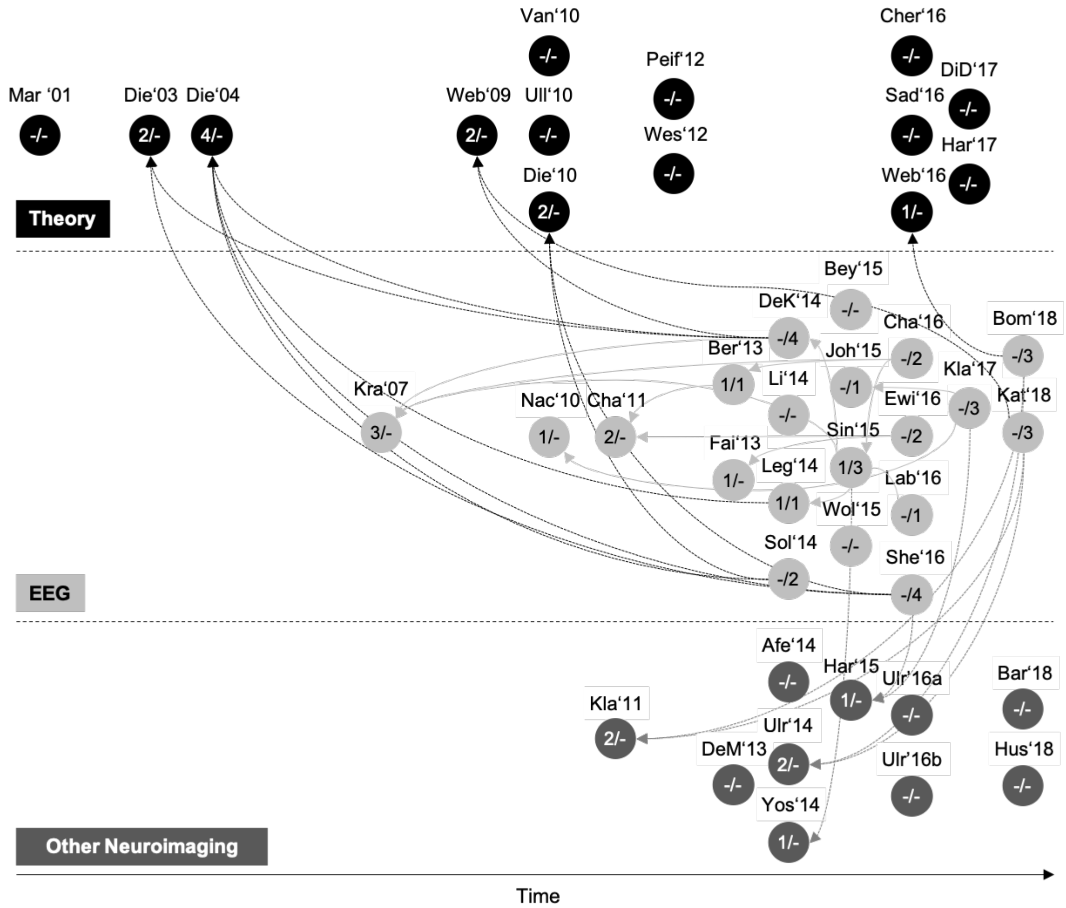

FIGURE 4.6: Historiograph of Three Flow Neurophysiology Study Categories. Arrows Show Referencing. Values In the Nodes Show In-/Out-Degrees. Referencing Was Only Evaluated for the EEG Studies. The Degree Counts and Reference Mappings are Summarised Again in the Appendix Table A.1.

attention (FA = Focused Task Attention), arousal or emotion concepts (AP = Approach Motivation, IA = Increased Arousal, RE = Increased Relaxation). The distribution of alternative hypotheses is similar. NE hypotheses are combined with all arousal and affect concepts, which shows an increased diversity in theoretic opinions. This diversity is further apparent in the presence of antagonistic predictions, especially relating arousal/relaxation states (IA vs RE). Regarding research designs and measures, the majority of studies (72.73%) uses acceptable, albeit relatively small sample sizes (mean = 26.45, median = 21). Also, the majority of studies (86.36%) uses high quality (i.e. research-grade) EEG systems with multiple electrodes on a gel or saline basis. In terms of establishing ground truth of flow elicitation, self-report instruments are almost always used (86.36%). This focus shows that this is still likely considered the benchmark (see also Moneta, 2012). Some studies (40.91%) include behavioural metrics, always operationalised with some type of task performance measure (either positive = score, negative = errors, and efficiency = moves). All studies are conducted in laboratory environments. Most (54.55%) use the DM paradigm. Given the similarity of the MAS to the DM paradigm (both focus on an adapted difficulty as flow entrance requirement), it can be said that difficulty adaptation is the central experimental approach. In terms of analysis foci, the majority of studies (95.45%) performs traditional Statistical Inferential Modelling (SIM). Yet, some studies also

Reference	Theory View	Design N	Design Par.	Context	Ground	Measures Sys.	Measures FOI	Measures ROI
Statistical Inferential Modelling (SIM)								
Bombeke et al. (2018)	FA	18	MAS	Gam.	Rep.,Per.	RG	ERP,FB	P
De Kock (2014)	THT	20	MAS	Gam.	Rep.+Per.	RG	FB	WS
Katahira et al. (2018)	-	16	DM	Cogn.	Rep.,Per.	RG	FB	WS
Kramer (2007)	NE, IA	10	MAS	Gam.	Phy.,Per.	RG	FB	T
Labonté et al. (2016)	NE, RE	42	MAS	Gam.	Rep.	RG	FB	F
Soltész et al. (2014)	THT	20	DM	Gam.	Per.	RG	FB	WS
Wolf et al. (2015)	NE, AP	29	MAS	Sports	Ref.	RG	FB	T
Ewing et al. (2016)	IA	20	DM	Gam.	Rep.,Per.	RG	FB	WS
Fairclough (2013)	IA, AP	20	DM	Cogn.	Phy.,Per.	RG	FB	F+C
Léger et al. (2014)	THT	36	DM	Edu.	Rep.,Phy.	RG	FB	F
Klarkowski (2017)	-	50	DM	Gam.	Rep.,Phy.	IG	FB	F+O
Beyer et al. (2015)	-	28	ENG	Gam.	Rep.	RG	FB	O
Johnson et al. (2015)	-	55	ENG	Gam.	Rep.	CG	FB	F
Nacke et al. (2010)	-	27	ENG	Gam.	Rep.,Phy.	RG	FB	WS
Li et al. (2014) - Exp. 1	-	44	DM	Gam.	Rep.,Phy.	IG	FB	F
Li et al. (2014) - Exp. 2	-	44	DM	Gam.	Rep.,Phy.	IG	FB	F
Chatterjee et al. (2016)	-	20	DM	Cogn.	Rep.,Per.	CG	Hjorth	F
Sinha et al. (2015)	-	16	DM	Cogn.	Rep.,Per.	CG	FB	F
Machine Learning (ML)								
Berta et al. (2013)	-	22	DM	Gam.	Rep.,Phy.	CG	FB	F+T
Chanel et al. (2011)	-	14	DM	Gam.	Rep.,Phy.	RG	FB	WS
Shearer (2016) - Exp. 1	THT	27	ENG	Gam.	Rep.,Phy.	RG	FB,Raw	F+P
Shearer (2016) - Exp. 2	THT	4	ENG	Gam.	Rep.,Phy.	RG	FB,Raw	F+P

Notes: HFA = High and Focused Task Attention; NE = Neural Efficiency;
APP = Approach Motivation; IA = Increased Arousal; RE = Increased Relaxation;
Par. = Paradigm; Gam. = Gaming; Cogn. = Cognition; Edu. = Education;
Rep. = Self-Report; Phy. = Physiology; Per. = Performance;
RG = Research Grade (Many & Gel-based Electrodes); IG = Intermediate Grade
(Some & Possibly Wet Electrodes); CG = Consumer Grade (Few & Dry Electrodes);
F = Frontal; T = Temporal; P = Parietal; O = Occipital; WS = Whole Scalp.

TABLE 4.2: EEG SLR Overview of Extracted Concepts.

(and sometimes focally) utilise ML methods (including feature selection and state classification) (18.18%). All these studies focus on temporally aggregated findings during the observation (e.g. the full five minutes of a task condition pooled). On the other hand, a small percentage of studies (13.64%) analyses the Time Dynamics (TD) EEG patterns within such conditions.

Regarding the (EEG) results extraction, a particular effort was made in not only the inclusion of positive (significant), but also negative (insignificant) and potential findings (e.g. mixed results, feature selection results, or descriptive comparisons of central tendency measures). This returned a high amount of collected data (>1000 unique results, i.e. unique combinations for region x feature x method, for the 22 studies). Findings were split into homogenous sub-categories to increase the comparability of results. These sub-categories are first of all based on employed methods: (1) SIM Findings - 880 findings, and (2) ML Findings - 50 findings. Second of all, studies observing EEG data changes over time during a task were collected separately as (3) TD Findings - 148 findings. Given the large number of reported findings, it was in this stage decided to narrow the focus on studying a restricted feature range. First, it was decided to focus on findings that relate to frontal brain regions to consolidate THT-related findings. Additionally, it was decided to focus on frequency-domain features that represent the majority of reported observations. Findings from feature selection procedures of ML studies that inform which features might have diagnostic value to separate flow intensities were also included in the following results. To emphasise the feasibility of this approach of focusing on frontal cortex frequency power changes, three additional summaries were used. In an overview of the 21 studies employing SIM or feature selection methods, six different types of features are reported. Table 4.3 shows that the two largest clusters of utilised features in all studies are for frequency-domain features. Frontal brain frequency band changes in flow experiences have been reported often. The frequency-domain features include band amplitude and power findings, their ratios and lateral asymmetries. An overview of the types of features is presented in Table 4.4. The following ranges were used to categorise reports from different frequency ranges: Very Low Frequency (VLF) (0,5-2 Hz), Delta (0,5-4 Hz), Theta (3,5-8,5 Hz), Alpha (7-14 Hz), Beta (10-35 Hz), Gamma (30-50 Hz), Very High Frequency (VHF) (30-100 Hz), Broad (1-40 Hz). The majority of findings cluster in the simple frequency band features, particularly in the ranges from Theta to Beta.

Feature Class	Feature	n-Studies	n-Paradigms	n-Methods	n-Tasks
Time-Domain	ERP	1	1	2	1
	Hjorth	1	1	1	1
	Freq. Band	20	3	6	5
Freq.-Domain	Ratio	4	3	3	2
	Asymmetry	2	2	2	2
	Coherence	4	3	3	2

TABLE 4.3: Diversity of Observed EEG Feature Types.

Freq. Band	VLF	Delta	Theta	Alpha	Beta	Gamma	VHF	Broad
Occurrence	2	8	17	16	14	6	3	3

TABLE 4.4: Counts of Reported Frequency Band Ranges.

To reiterate, the study of frontal regions has been preferred often based on the THT account of flow neurophysiology (Dietrich, 2004). So far, for THT's central hypothesis of frontal activity reduction during flow, little support has been found in fMRI (Ulrich et al., 2014) and fNIRS (Harmat et al., 2015; Barros et al., 2018) imaging studies. Instead, it appears parts of the PFC, specifically lateral parts, are highly active during flow, yet the medial PFC shows activity decreases during flow, and underload/boredom conditions show a more general PFC reduction (Harris, Vine, and Wilson, 2017b; Barros et al., 2018). Frontal activity has also been reported in most of the related flow EEG studies, with repeated results supporting the region as a location of interest. While three of these studies (Chanel et al., 2011; Berta et al., 2013; Sinha et al., 2015) report on the relevance of frontal activity for the ML-based classification of flow states, nine other studies describe activity in more detail. The findings on frontal frequency band patterns with flow experiences are ordered by the three major frequency bands (Theta, Alpha, and Beta) and are reported in Table 4.5. The notation for the findings is the same as described in Section 4.2 for the PNS SLR. One difference is that here, findings are reported concerning their observation on the scalp. This detail means that the results are reported from left to right hemisphere with different stages of distance to the midline. For instance, a finding listed furthest to the right indicates that the observation was made on electrodes on the lateral outside of the scalp (e.g. electrode F8). In addition, the table shows the used frequency band ranges (in Hertz), the context of the experimental task and the type of employed analysis (indirect comparison of conditions are denoted by separated letters - e.g. B/F/O; direct comparisons using correlation or regression analyses are indicated by flow self-report instrument that was used). A novel element in this table (in comparison to the PNS SLR) is the x relationship, which indicates that a potential to differentiate the experimental conditions was found (e.g. from Feature Selection (FS) algorithms), but that the exact relationship to these conditions or flow is not reported. Whenever mixed results are reported (e.g. in the form of x/ •), this means that multiple models were tested and some reported differentiation for conditions, while others did not.

Aggregating the results of these studies that primarily focused on frequency band activity across difficulty-manipulated conditions, numerous results can be integrated. First, one of the more robust findings is an elevated level of frontal Theta band activity in flow. More precisely, frontal Theta levels appear to be elevated in flow conditions (e.g. moderate or balanced difficulty conditions), when compared to easy difficulty conditions. Furthermore, either similarity between flow and hard (difficulty) conditions (Soltész et al., 2014; Katahira et al., 2018) are found or decreases from flow to hard conditions indicating an inverted U-shaped relationship between frontal Theta activity and task demands (Ewing, Fairclough, and Gilleade, 2016; Fairclough et al., 2013). Support for frontal Theta to differentiate situations of lower and higher flow has also been noted in ML research on flow classification (Chanel et al., 2011; Sinha et al., 2015). In contrast to this, a minority of findings shows no relationship between frontal Theta and difficulty condition comparisons (Klarkowski, 2017; Chanel et al., 2011), between frontal Theta in quasi-experimental condition comparisons (lower vs higher flow intensity trials) (De Kock, 2014), and correlation or regression analyses using self-reports and frontal Theta relationship analysis (Katahira et al., 2018; Klarkowski, 2017). Also, one instance of a negative relationship between frontal Theta and flow self-reports is found (Li et al., 2014).

Second, repeated observations have been made for frontal Alpha activity, albeit with even higher diversity. The Alpha frequency band and the diversity of findings are particularly relevant to review, as Alpha has been recommended as the prime candidate to study areas of (prefrontal) downregulation (Cheron, 2016; Harris, Vine, and Wilson, 2017b). This proposition is based on the understanding that Alpha is an inhibitory oscillatory rhythm that indicates cortical idling (Buzsaki, 2006). Therefore, increased Alpha power in a particular region indicates a reduction of neural activity. In the present corpus, one study finds increased Alpha power with higher flow self-reports (Léger et al., 2014). This finding is supported in its direction by a quasi-experimental comparison of difficulty conditions, in which frontal Alpha

Reference	Context	DV	Hz	LH (Lat. to Med.)			Mid./All	RH (Med. to Lat.)		
Theta Frequency Range (4-8 Hz)										
Katahira et al. (2018)	Cogn.	B/F/O WS	4-7		⌐		⌐		⌐	
		sFSS			⟋		·		·	
De Kock (2014)	Gam.	fL/fH WS	4-7		·					
Ewing et al. (2016)	Gam.	B/F/O WS	i(4-7)				^			
Fairclough (2013)	Cogn.	B/F/O WS	i(4-7)				^			
Sinha et al. (2015)	Cogn.	fL/fH WS (FS)	4-7,5			x				
Berta et al. (2013)	Gam.	B/F/O WS (FS)	4-8	·						·
Chanel et al. (2011)	Gam.	B/F/O WS (FS)	4-8	·	x / ·	x / ·	x / ·	x / ·	x / ·	x / ·
		B/F/O WS		·	·	·	x	·	x	·
Klarkowski (2017)	Gam.	B/F/O WS / FSS	4-8						·	
Li et al. (2014) Exp. 1	Gam.	GEQ	4-8		⟍					
Li et al. (2014) Exp. 2	Gam.	GEQ	4-8		⟍					
Soltész et al. (2014)	Gam.	B/F/O WS	4-8		⟋		⌐		⌐	
Alpha Frequency Range (7-13 Hz)										
Ewing et al. (2016)	Gam.	B/F/O WS	7,5-10				·			
Soltész et al. (2014)	Gam.	B/F/O WS	8-11		·		·		·	
Ewing et al. (2016)	Gam.	B/F/O WS	10,5-13				⟍			
Soltész et al. (2014)	Gam.	B/F/O WS	11-13		·		·		·	
Katahira et al. (2018)	Cogn.	B/F/O WS	10-13		⌐		⌐		⌐	
		sFSS			·		·		·	
Berta et al. (2013)	Gam.	B/F/O WS (FS)	8-12	x						x
Chanel et al. (2011)	Gam.	B/F/O WS (FS)	8-12	x / ·	x / ·	x / ·	x / ·	x / ·	·	x / ·
		B/F/O WS		·	·	·	·	·	·	·
De Kock (2014)	Gam.	fL/fH WS	8-12		·					
Labonté et al. (2016)	Gam.	B/A/F/O WS	8-12						^	
Léger et al. (2014)	Edu.	CA	8-12		⟋					

Reference	Context	DV	Hz	LH (Lat. to Med.)			Mid./All	RH (Med. to Lat.)		
Klarkowski (2017)	Gam.	B/F/O WS / FSS	8-13						•	
Beta Frequency Range (12-35 Hz)										
Berta et al. (2013)	Gam.	B/F/O WS (FS)	12-15	x						x
De Kock (2014)	Gam.	fL/fH WS	12-15		•					
Léger et al. (2014)	Edu.	CA	12-22		↘					
Berta et al. (2013)	Gam.	B/F/O WS (FS)	15-20	x						x
De Kock (2014)	Gam.	fL/fH WS	15-20		↗					
Sinha et al. (2015)	Cogn.	fL/fH WS (FS)	16-20			x				
Soltész et al. (2014)	Gam.	B/F/O WS	13-25	•			•		•	
Berta et al. (2013)	Gam.	B/F/O WS (FS)	20-30	x						x
De Kock (2014)	Gam.	fL/fH WS (WS)	20-30		•					
Soltész et al. (2014)	Gam.	B/F/O WS	25-35	•			•		•	
Chanel et al. (2011)	Gam.	B/F/O WS (FS)	12-30	x / •	x / •	x / •	x / •	•	x / •	x / •
		B/F/O WS		•	•	•	•	•	•	•
Klarkowski (2017)	Gam.	B/F/O WS	13-30						↖	
		FSS							•	
Katahira et al. (2018)	Cogn.	B/F/O WS	14-30	•			•		•	

Notes: Most Publications Report Freq. Band Power, Katahira et al. (2018) Report Freq. Band Amplitudes;
Mid. = Midline; Lat. = Lateral; Med. = Medial; Gam. = Gaming; Edu. = Education; Cogn. = Cognition;
B/A/F/O = Boredom/Apathy/Fit(Flow)/Overload Difficulty Conditions;
fL/fH = Lower Flow / Higher Flow Group Split; Condition With Highest Flow Intensity is Underlined.

TABLE 4.5: Overview of Flow EEG Findings.

power becomes maximal at right frontal sites during increased and balanced difficulty situations (Labonté-Lemoyne et al., 2016). In contrast, within the DM group comparison studies, findings of decreases in Alpha activity with increasing task difficulty are found (Ewing, Fairclough, and Gilleade, 2016). Katahira et al. (2018) report the inverse relationship, but use amplitudes as the unit of analysis (squaring the results would also instead indicate a reduction of frontal Alpha with difficulty increases). ML research also finds frontal Alpha activity to be a difficulty-differentiating feature (Berta et al., 2013; Chanel et al., 2011). However, some studies find no relationship of frontal Alpha power with flow, with different situations, for example through flow self-report correlation and regression analysis (Katahira et al., 2018; Léger et al., 2014; Klarkowski, 2017), or quasi-experimental condition comparison (lower vs higher flow intensity trials) (De Kock, 2014). Also, multiple studies find no relationship of frontal Alpha in difficulty condition comparison (Soltész et al., 2014; Ewing, Fairclough, and Gilleade, 2016; Chanel et al., 2011; Klarkowski, 2017). The latter finding appears to occur more often in this contrast to the studies mentioned earlier when the broad Alpha band or lower Alpha band components are observed. Overall it can be noted that a variety of frontal Alpha findings have emerged that do not show a clear relationship with flow so far.

Third, some observations have also been made regarding frontal Beta band activity in flow, with ML reports demonstrating differentiation potential alone for broad Beta (Chanel et al., 2011) and Beta sub-bands (Berta et al., 2013; Sinha et al., 2015). Beyond this aspect, results are mixed. One study finds that left frontal beta band reductions correlated with higher flow self-reports (Léger et al., 2014). On the other hand, DM and quasi-DM studies point to increases in frontal Beta at both left and right frontal sites (Klarkowski, 2017; De Kock, 2014). However, the slightly most repeated finding is a lack of relationships between flow and frontal Beta in difficulty condition comparisons (Chanel et al., 2011; Soltész et al., 2014; Katahira et al., 2018). The meaning of these findings, together with how these frequency ranges have been observed in more fundamental neuroscientific studies, is discussed in the next section.

Discussion & Future Directions

RQ2 asked how the brain's configuration during flow can be described using EEG measures. This SLR highlights five aspects of the present state of flow EEG research, specifically the study of frontal brain regions to investigate THT-related hypotheses.

Overall, a low degree of integration of the present EEG work on flow with itself and with other neuroimaging and neuro-theoretic research is found. This fragmentation is not only visible by the apparent lack of citations, which is to some degree understandable given the youth of the field, but also by the lack of inclusion of theoretic accounts on flow neurophysiology. Instead, several articles follow the motivation to explore neurophysiological patterns during increased flow experience primarily, even some that have been published rather recently (i.e. even amongst the very latest published studies). This lack of integration has likely impeded the advancement of more refined observation of diagnostic EEG patterns during flow and might be a central reason for why a relatively high diversity of findings is present, despite somewhat comparable paradigms and analyses. It is, therefore, an essential recommendation for future research to take into account such previous results when interpreting how novel findings are likely to integrate with such previous work. It presently seems of utmost importance that more detail is added to future flow research in terms of explanations for observed patterns (in light of employed paradigms, tasks, analyses and pre-processing steps) and terms of considerations of the robustness of reported findings. To aid in this endeavour, the utilisation of cross-task or cross-paradigm studies appears as a highly valuable approach for the identification of flow-related EEG patterns. As was the case for the work in the PNS SLR, there is currently a strong focus present on game tasks and DM paradigms, that can be used as an anchor for new work, but that should be overcome importantly for the assessment of the robustness of identified relationships.

Despite the reasonably comparable and (for the most part) very adequately operationalised experimental approaches, so far, barely any highly diagnostic findings have emerged that would help to identify a unique configuration of flow (in frontal regions), that can be detected by the EEG. Such a finding is not necessarily something that would have to occur, given that little is known and theorised about what the neural configuration during flow is. However, the present (mostly explorative) body of work can be appraised in terms of how diagnostic emerging results are. The findings that could be classified as having higher diagnosticity (e.g. showing maxima or minima in situations with higher flow) are either related to direct linear modelling with self-report constructs, or given for example with the frontal Theta maxima during balanced difficulty conditions (Ewing, Fairclough, and Gilleade, 2016; Fairclough et al., 2013). Whether or not frontal Theta is a sign of flow is yet to be clarified, especially as frontal Theta power has so far more strongly been related to changes in task difficulty (Borghini et al., 2014; Silvestrini, 2017). Moderate Theta levels suggest some degree of long-range cortico-cortical communication, particularly emerging from frontal regions (where midline Theta increases are typically reported) as the Anterior Cingulate Cortex (ACC) recruits information from other brain regions to meet task demands (Borghini et al., 2014; Silvestrini, 2017). A possible explanation for the maxima in (Ewing, Fairclough, and Gilleade, 2016; Fairclough et al., 2013) could be that the task difficulty was too high so that disengagement occurred. However, such disengagement must not result in lower Theta due to quitting the task altogether. It is, therefore, possible that maximal frontal Theta activation could be a sign of high attentional task engagement, that peaks during flow experiences.

Regarding the diagnosticity of linear modelling of flow with self-reports, caution appears warranted. In the presented corpus, analysis methods using direct linear modelling of flow self-reports and EEG features appear to often return quite different results from those studies that compare differences across difficulty conditions. The explanation for this complication is two-fold. Either there is a different part of the variance of the flow construct that is observed through either method (Riedl, Davis, and Hevner, 2014; Léger et al., 2014), or the patterns observed across manipulated difficulty are not as strongly directly linked to flow intensities (they could relate more to difficulty or other, third variables). In either case, it is not yet sufficiently clear, how to integrate these diverging findings. Importantly, it should be considered whether or not a focus on direct linear models is adequate. In contrast, in PNS flow studies, non-linear models have increasingly shown utility in detecting direct relationships between physiology and flow self-reports (e.g. Tozman et al., 2015; Bian et al., 2016). Given these questions, future research should, therefore, focus on reconciling the presence of relationships with more detail and sophistication in experiment designs and employed analyses. For example, the use of more steps in DM paradigms or the combination of DM paradigms with other experimental approaches could help in clarifying flow relationships to frontal Theta power changes. In addition, the use of non-linear analyses appears as a useful approach to deepen and consolidate results from indirect analyses (i.e. difficulty condition comparisons) with direct models.

Concerning THT, there is so far no clear evidence, as to whether or not it is observable by the EEG. In particular, observations of frontal Alpha activity patterns have returned mixed results. Frontal Alpha observation is recommended to identify cortical idling during flow and could indicate prefrontal downregulation (Cheron, 2016; Harris, Vine, and Wilson, 2017b). Across studies, experiment tasks, and methods of analysis, it is not yet possible to conclude whether or not a reduction in prefrontal cortex areas is visible by the EEG in terms of increased Alpha power. Furthermore, the findings of increased frontal Theta power during increased flow would appear to indicate an increased (or at least not decreased) utilisation of prefrontal structures like the ACC - as this is a well-documented finding in working memory research (Silvestrini, 2017; Klimesch, 1999; Borghini et al., 2014). In this regard, while a general reduction of activity in prefrontal structures also seems unlikely in terms of EEG findings, more refined approaches are required to understand the frontal cortex configuration during flow experiences. Specifically, it appears that so far, some of the more refined analyses of frontal electrodes and especially the distribution of power in a refined Alpha band spectrum have not yet been integrated well into the line of flow EEG research. Often frequency band ranges are extracted using generalised, broad ranges (e.g. Theta 4-7.5 Hz or Alpha 7.5-12.5 Hz), despite the evidence, that such generalised ranges can mask frequency sub-band specific changes (Klimesch, 1999), and despite more established segmentation in EEG research on meditation (Hinterberger, Kamei, and Walach, 2011). Importantly, evidence from laboratory experiments has highlighted that Alpha band components can show different and even sometimes opposing patterns (Klimesch, 1999). For example, by segmentation of personalised Alpha bands into three 2 Hz wide subcomponents, lower Alpha bands (Lo1 and Lo2) have been found to relate to general attentional demands and alertness over the whole scalp (Klimesch, 1999). The upper Alpha has been found to react to changes in task-specific processes in topographically restricted regions (Klimesch, 1999). As flow is not only repeatedly associated with cognitive demands in the form of working memory recruitment (i.e. Theta range activity), but also often in relation with attentional processes (Harris, Vine, and Wilson, 2017b), it would seem of high interest to employ Alpha band segmentation to not only identify regions of reduced neural activity, but perhaps even identify global changes in attentional demands, and task-specific pattern changes. Furthermore, given the findings of opposing activation in medial and lateral PFC sites through hemodynamic imaging (Ulrich et al., 2014; Ulrich, Keller, and Grön, 2016b; Barros et al., 2018), it would appear that a more nuanced observation of Alpha power changes at different frontal electrode positions could deliver useful additional insights.

Lastly, higher frequency findings (i.e. Beta) on frontal locations have been mixed. As to whether or not, a particular pattern would be expected here only a few propositions have been put forward. Léger et al. (2014), for example, expect a reduced frontal Beta activation as a sign of reduced arousal during flow. As in other related flow research, moderate physiological arousal levels have been proposed to be a (moderately diagnostic) correlate of flow experiences (Peifer, 2012) (see Section 4.2), such an hypothesis seems somewhat justified. On the other hand, Beta power (especially when segmented in lower and higher parts) is known to react positively to changes in mental workload through increased task difficulties (Michels et al., 2010). It is therefore similarly plausible that flow would at least be represented by elevated frontal Beta levels. However, such a finding might also be more expressed in central and posterior brain regions (Michels et al., 2010). From the present state of research, it is not yet clear, if a Beta change is common during flow experiences, or if it perhaps depends on electrode positions and Beta band segmentation. For example, fronto-lateral Beta increases are found by Klarkowski (2017) in a broad Beta band, and by De Kock (2014) in a mid-range Beta band. The reasons for the large diversity in frontal Beta changes during flow need to be better understood.

Beyond these main advancements that can be made to flow EEG research, it should also be mentioned that a variety of additional observations would be interesting to follow up on. The analyses and discussions above have focused on analyses conducted with temporally aggregated EEG data (e.g. means of windows of a few minutes of observed EEG data). Especially as the EEG has the main advantage to study direct activity at a very high temporal resolution, studying time dynamics of flow would seem like a unique approach to learn about the dynamics of flow onset, protection and offset. In more fundamental neuroscientific work, it has, for example, been found that the study of time dynamics can help to more robustly identify instances of approach or avoidance motivation (Allen and Cohen, 2010). Approach motivation (observed through what is called Frontal Alpha Asymmetry - see Harmon-Jones, Gable, and Peterson, 2010; Smith et al., 2017) has been suggested to be related to flow (Labonté-Lemoyne et al., 2016). This association is based in the understanding that flow is an experiential state that is characterised by high intrinsic motivation and a desire to repeat an experience. It would, therefore, be plausible to assume that during flow, individuals experience a high level of approach motivation.

Conclusion

This SLR summarised the work from 22 studies on EEG observation during flow. A central result is that more detailed spectral, spatial and temporal analyses across tasks and paradigms are needed to consolidate this highly fragmented state of work. The SLR provides a starting point to pursue these directions and therefore contributes to flow neurophysiology research and to the development of theory-driven adaptive NeuroIS. Altogether, it can be stated, that there is still a long way to go to uncover how EEG measures can describe the flow experience (see, e.g. Katahira et al., 2018; Soltész et al., 2014). A focus on broad regions or frequency bands is likely to oversimplify the neural state during flow. However, it can aid in better understanding what occurs in the brain in terms of broader cognitive and physiological processes. More comprehensive, detailed, and feature-integrative studies are needed. To aid in the endeavour of isolating and consolidating EEG patterns of flow, two general directions with similar value should be highlighted that are replication and exploration. Many present findings would greatly benefit from additional, comparable data to consolidate the understanding of the neurological basis of the flow state, and replication would provide anchors for new explorations. Therefore, flow EEG research should more strongly increase in detail than in diversity. It is for these reasons why the propositions heavily feature advanced analyses methods, whose possibilities should be substantiated by refined designs and measurement. In terms of research designs, refinements should strive to improve internal and external validity, reliability and diagnosticity. To improve internal validity, the elicitation of more intense flow experiences should be considered.

Beyond these factors and especially given the current lack of highly diagnostic EEG findings, the inclusion of additional measurement methods has been highlighted (Katahira et al., 2018; Harris, Vine, and Wilson, 2017b; Cheron, 2016). Co-registered hemodynamic imaging and EEG methods could provide more refined insight into the neural state during flow. However, these are also relatively complicated and costly approaches. In contrast, ANS modulation of cardiac activity is considered a valuable extension for the assessment of neural activity during task performance and stressful experiences (Thayer et al., 2009; Thayer et al., 2012). The inclusion of ECG measures to EEG study of flow could, therefore, provide a simple, yet informative addition (Harris, Vine, and Wilson, 2017b). Lastly, advanced analysis methods (in detail and quality) represent the central direction for future work. Increasingly detailed analyses should be realised first in feature extraction processes. Improvements should occur in the study of effects in frequency band splits, and individualised features (e.g. individualised Theta and Alpha) that have been reported to represent an often occurring confound in EEG research (Klimesch, 1999; Ewing, Fairclough, and Gilleade, 2016; Hinterberger, Kamei, and Walach, 2011). Afterwards, improved localisation of signal sources by application of spatial filtering methods (a significant improvement in EEG methods in the past decade - see Cohen, 2014; Blankertz et al., 2016) could provide necessary additional detail.

The field of flow EEG research appears still ripe with opportunity. The main contribution of this SLR is the documentation and integration of a highly fragmented body of work. In this dissertation, this basis is extended by two experiments using two different tasks with four different paradigms in total. Using both ECG and EEG measurement and following up on the proposition of refined frontal power investigation (by studying more frontal locations and frequency sub-bands) further consolidation of the state of knowledge on flow neurophysiology is provided. Before the experimental results are presented, the following section details the main neurophysiological measurement approaches to provide an essential background about their properties.

4.4 ECG & EEG Methodology

Selection of Measurement Approaches

ECG and EEG measures were considered to represent likely candidates for the investigation and continuous detection of flow experiences due to their low costs, high portability and high temporal resolution. These characteristics make them both valuable for fundamental research and its eventual transfer into real-world applications (i.e. flow-facilitating adaptive NeuroIS). Furthermore, both measurement systems are not only highly developed for research purposes (the EEG for instance, has been employed in research for almost 100 years - see Buzsaki, 2006), but are also increasingly available for the utilization in field studies and daily applications. ECG sensors are readily integrated into chest belts that can be worn throughout the day (Baig, Gholamhosseini, and Connolly, 2013; Weippert et al., 2010), and have for these reasons already been used in field studies on emotion and flow experiences (Wilhelm et al., 2006; Gaggioli et al., 2013). Moreover, ECG sensors have been directly integrated into textiles to enable continuous and comfortable cardiac activity observation (Yoo et al., 2009). The maturity of the technology has even allowed scholars to develop their own low-cost wearable ECG devices that are used in field studies after having been sent to participants' homes by mail (Courtemanche et al., 2020). Similarly, EEG sensors have seen increasing potential to bridge laboratory and field research. The development of dry-electrode EEG systems has greatly facilitated the sensor usability, and related feasibility studies have repeatedly demonstrated acceptable data collection qualities (Guger et al., 2012; Mullen et al., 2015). Subsequently, publications in prestigious academic journals have appeared that leverage such technology for EEG research in laboratory and field settings (see, e.g. (Krigolson et al., 2017; Barham et al., 2017; Bobrov et al., 2011)). Wearable EEG sensors have been used to observe mental workload levels in software engineers at work (Kosti et al., 2018) or to provide neurofeedback for emotion regulation (Ramirez et al., 2015).

Furthermore, the results from the SLRs further emphasise the relevance of observing neural and cardiac activities. Thus, the focus on these two measurement domains was considered to be well-grounded. The focus was kept on two measures for additional reasons. Psychophysiological research requires high levels of domain knowledge on neurophysiological dynamics to enable the derivation of substantial insights (Brouwer et al., 2015). Therefore, focusing on two measures is already a challenging endeavour. However, as both ECG and EEG observe electrical activity, at least this shared functional principle allows for methodological overlaps (e.g. in terms of signal processing and feature extraction). Besides, neuronal and cardiac activities have been jointly researched in the related, but more general fields of cognitive-affective regulation research (Thayer et al., 2009; Mather and Thayer, 2018) and mental workload research (Borghini et al., 2014; Silvestrini, 2017). Also, in flow research specifically, the integration of frontal neural activity and cardiac reactivity has been considered due to shared physiological regulation mechanisms (Barros et al., 2018; Harris, Vine, and Wilson, 2017b). Therefore, the integration of both neural and cardiac measures was considered to provide complementary insights.

Altogether, combining ECG and EEG measures for the description and continuous detection of flow experiences was considered promising, feasible, and timely in terms of how such instrumentation could become implemented in adaptive NeuroIS in the near future. To provide a critical understanding of the underlying processes of these data domains, the following two sections briefly describe the underlying physiological systems, the general principle of the measurement instruments, and the features extracted from the signal that are primarily investigated in this dissertation.

A Primer on the Autonomous Nervous System (ANS), Electrocardiography (ECG), and Heart Rate Variability (HRV)

The Autonomous Nervous System (ANS) & Cardiac Activity

The human nervous system consists of different parts. The following descriptions are based on the work by Andreassi (2000). A primary distinction is made between the Central Nervous System (CNS) (brain and spinal cord) and the Peripheral Nervous System (PNS) (all tissue except for the CNS). The PNS comprises two subsystems that are the somatic and autonomic nervous system. The somatic system consists of cranial and spinal nerves to and from the sensory organs, muscles, joints, and skin. The main functions of the somatic nervous system are movement control and the transmission of sensory information (e.g., vision, temperature, touch). The Autonomous Nervous System (ANS) acts as the coordinator of critical bodily activities, including digestion, body temperature, and blood pressure. The term "autonomic" implies that these functions are mostly operating outside of volitional control. The ANS is further subdivided into two branches, the excitatory sympathetic nervous system and the inhibitory parasympathetic nervous system. Whereas the former is related to increased physiological activation due to the mobilisation of energy resources in emergency and stress situations (“fight-or-flight”), the latter is related to physiological deactivation, or relaxation during return to safer circumstances (“rest-and-digest”). These ANS responses are correlated with a number of physiological changes. Thus. sympathetic activation correlates with: pupil dilation, skin conductance elevation, airway relaxation, heartbeat acceleration, heart rate variability decrease, glucose release, and muscle tension. Parasympathetic activation results in reverse effects: pupil contraction, skin conductance reduction, airway constriction, heartbeat slowing, heart rate variability increase, halted glucose release, and muscle relaxation.

For research on flow experiences, the influences of the ANS on cardiac activity have been of central interest (see Section 4.2). The cardiovascular system is controlled by the two branches of the ANS, through its connection to the vasculature and the heart (Berntson, Quigley, and Lozano, 2007; Milnor, 1990). The heart represents the pump in the cardiovascular system. The main heart muscle consists of two nodes (sinoatrial and atrioventricular node), two atria and two ventricles. The atria and ventricles are separated by the heart valve (see Figure 4.7). The heart prompts blood circulation in blood vessels through muscle contraction (systole) and relaxation (diastole) in alternation (Milnor, 1990). Each cardiac cycle (i.e., heartbeat) begins with the depolarization of the sinoatrial node (the heart’s pacemaker), causing a contraction of the atria. The electrical signal spreads, the atrioventricular node depolarizes causing a contraction of the ventricles. After the stimulation of the ventricles, a refractory period follows. Then, after repolarisation of the ventricles, the heart is ready to repeat the cycle. This cycle of events allows the heart to perform the pumping action. The unit of analysis in cardiac activity observation is the time between two heartbeats that forms the basis of measures like Heart Rate (HR) and Heart Rate Variability (HRV) (Malik et al., 1996). This heartbeat sequence can, for example, be retrieved through ECG.

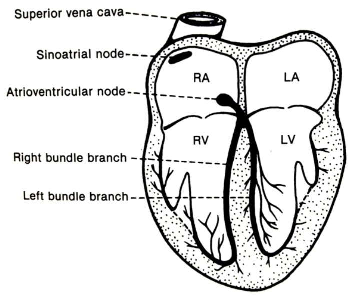

FIGURE 4.7: The Human Heart - See Milnor (1990).

Electrocardiography (ECG)

The ECG records electrical potentials in millivolts. For a recording, electrodes are placed on prescribed locations on the skin (Berntson, Quigley, and Lozano, 2007). A clinical ECG consists of 12 leads but three- and five-lead ECGs, which simplify the setup process, are also available (Petty, 2015). In this dissertation, to utilize wearable sensors, a three-lead configuration (termed Lead II) was chosen (see Figure 4.8).

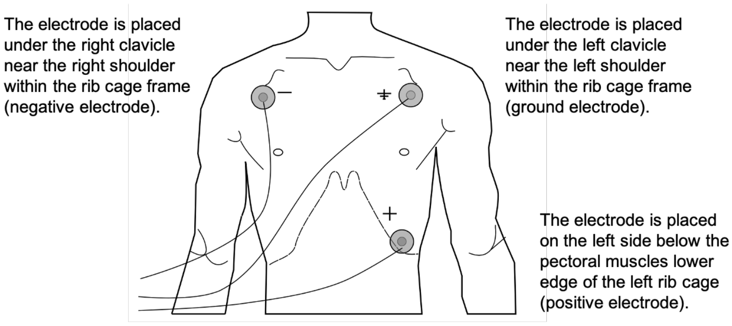

FIGURE 4.8: ECG Lead II Placement Standard Based on a Three-Electrode System as Used in the Present Work - See Fortin-Cote et al. (2019).

Based on a heartbeat, an electrical impulse travels through the heart. During the migration of the electrical stimulation wave, a variety of potentials differing in size and direction arise (Milnor, 1990). Figure 4.9 shows a conceptual illustration of an ECG-signal of a heart in normal sinus rhythm. The characteristic electrical signal is segmented in the so-called PQRST-cycle that is further divided into the initial P-wave, the QRS-complex, and the final T-wave. Following the depolarization of the sinoatrial node, the contractions in the right and left atria produce the first electrical impulse referred to as P-wave. A flat line follows as the electrical impulse moves to the ventricles. The right and left ventricles produce the next impulse referred to as QRS-complex. The final impulse is referred to as the T-wave, representing electrical recovery or return to a resting state for the ventricles. To extract HR and HRV measures from an ECG recording, detecting the QRS-complex is essential for measuring the time interval between the succeeding heartbeats (Malik et al., 1996).

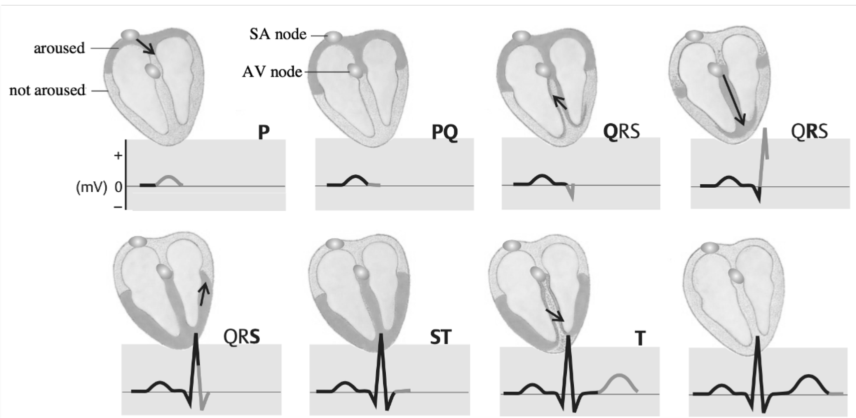

FIGURE 4.9: Segments of an ECG Recording - See Kerner and Brückel (2011).

Heart Rate Variability (HRV)

HRV is a function of ANS activity. The sympathetic part, among others, activates the release of stress hormones, increases the heart's contraction rate and force (cardiac output), but decreases HRV. Conversely, the parasympathetic part reduces the heart rate, but increases HRV, to restore homeostasis after a state of physiological activation. Importantly, this interplay between the sympathetic and parasympathetic parts of the ANS makes it possible that the heart instantaneously responds to different situations and needs (Malik et al., 1996; Berntson, Quigley, and Lozano, 2007). A variety of indicators extracted from ECG recordings exist that make the assessment of changes in heart rates possible (Malik et al., 1996; Shaffer and Ginsberg, 2017). Most of these indicators are calculated either in the time-domain or the frequency-domain of the signal. The most relevant signal aspect is the time interval between subsequent peaks in the QRS-complex, known as RR-interval or Inter-Beat Interval (IBI). RR-intervals of normal signals are known as NN-intervals.

Time-domain methods are typically based on the difference between succeeding NN-intervals (Malik et al., 1996). These features comprise, for instance, the Standard Deviation (SD) of the difference between adjacent NN-intervals (SDNN), the root mean square difference between adjacent NN-intervals (RMSSD) or the percentage of adjacent NN-intervals that do not differ more than 50ms (PNN50). Additional time-domain features (e.g. geometric or non-linear) are increasingly studied (Shaffer and Ginsberg, 2017). However, they have so far not seen application in flow research, which is why they are not detailed here further. The interpretation of statistical time-domain measures depends on the time window over which the measures are calculated (Malik et al., 1996; Shaffer and Ginsberg, 2017). Statistical time-domain measures are usually calculated over five minutes (short- term) to 24 hours (long-term) (Malik et al., 1996). The lower bound of five-minute recordings, therefore, represents an essential requirement in experimental flow physiology research (see, e.g. Tozman et al., 2015; Harmat et al., 2015). According to Shaffer and Ginsberg (2017), both the sympathetic and the parasympathetic nervous system impact the Standard Deviation of NN-Intervals (SDNN). By contrast, the Root Mean Square of Successive Differences (RMSSD) is predominantly influenced by the activity of the parasympathetic nervous system (Berntson, Quigley, and Lozano, 2007).

Frequency-domain methods rely on Power Spectral Density (PSD) analysis of the NN-interval sequence (Malik et al., 1996). The PSD provides information on how power distributes over frequencies. The frequency components depend on the length of the NN-interval sequence. For short-term recordings of two to five minutes, the frequency components are divided between very low (VLF: <0.04 Hz), low (LF: 0.04-0.15 Hz), and high frequency (HF: 0.15-0.4 Hz) components. The LF-HRV component is influenced by both the activity of the sympathetic nervous system and the parasympathetic nervous system (Berntson, Quigley, and Lozano, 2007). The High Frequency Heart Rate Variability (HF-HRV) component is only influenced by the parasympathetic nervous system. Conversely, HF-HRV metrics have been found to correlate positively with the time-domain feature RMSSD (Malik et al., 1996; Berntson, Quigley, and Lozano, 2007). Time- and frequency-domain measures present advantages and disadvantages. Time-domain measures do not rely on methods to obtain a PSD estimate and can be calculated more easily from the NN-interval sequence. However, frequency-domain measures may be the better choice for studying the context-dependent activity of the sympathetic and parasympathetic nervous systems. Also, frequency-domain measures have shown higher robustness for shorter observations (Berntson, Quigley, and Lozano, 2007).

Finally, it needs to be appraised that for an analysis of HR and HRV measures, it is recommended to take influential factors into account (Andreassi, 2000; Valentini and Parati, 2009). Influences occur from the proximal environment (e.g. temperature, or social interaction), but also internal variables like the consumption of recreational drugs, posture, or exercise (Andreassi, 2000; Valentini and Parati, 2009). For these reasons, the experiments presented in this dissertation have attempted to control for such issues by participant screening and requirements like the abstinence of recreational drugs before the experiment. Descriptions of these screening criteria are outlined in the experiment procedures in Chapters 5 and 6.

A Primer on Neuroanatomy and Electroencephalography (EEG)

Neuroanatomy

At the smallest unit of analysis, operations of the brain are coordinated by billions (estimated 50 to 100 billion) of brain cells called neurons (Andreassi, 2000). Neurons are excited by exogenous (e.g. sensory input) and endogenous stimuli (activity from other neurons) (Andreassi, 2000; Blankertz et al., 2016). As a critical point in stimulation is reached, the neuron will discharge and transmit the electrical impulse over to other cells. The concomitant effects are called excitatory or inhibitory post-synaptic potentials (Andreassi, 2000) due to their resulting increases and decreases of neuronal activity levels (Andreassi, 2000; Buzsaki, 2006). These post-synaptic potentials of neuronal activity can be recorded at the scalp level (Andreassi, 2000; Cohen, 2014; Cohen, 2017; Buzsaki, 2006). To better understand the process through which neurons coordinate operations in the brain, an overview of the brain's anatomy is provided.

From an overarching point of view, the major sections of the brain can be described as the cerebrum, cerebellum, and brain stem. The cerebrum comprises the largest part of the brain and contains two halves (hemispheres) (Andreassi, 2000). The cerebellum sits on top of the posterior part of the brain stem and transmits information to and from the spinal cord and to other parts of the brain. The brain stem sits at the posterior ventral part of the brain and structurally connects with the spinal cord. The cerebrum contains the cerebral cortex (outer layer) and several subcortical structures (including the hippocampus, basal ganglia, and amygdala). The outer layer, that is the cerebral cortex is the outer grey matter (where the cell bodies sit) covering the surface of the cerebrum. The cortical surface controls basic sensory and motor functions. By orientation on central fissures, the cerebral cortex can further be divided into four lobes: frontal, parietal, occipital, and temporal (Andreassi, 2000). Lobes and cerebral cortex are visualized in Figure 4.10. The frontal lobe is considered to

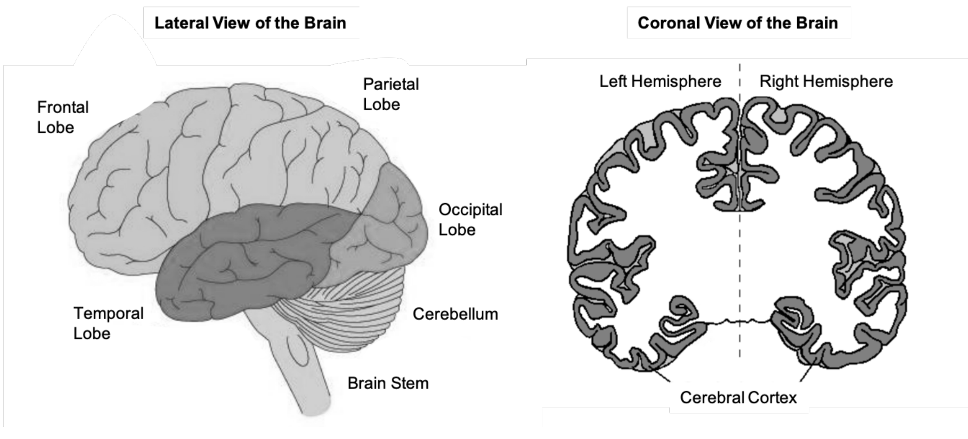

FIGURE 4.10: Main Structures of the Cerebrum, Specifically Lobes and Cortex - See Ferrez and Milan (2007).

be centrally involved with cognitive processes like planning, decision-making and problem-solving (Andreassi, 2000; Carlén, 2017). At the temporal regions, auditory and language-processing related functions are situated. At central and parietal regions lie the primary motor area and the somatosensory cortex that handle input signals from bodily sensations and regulate the execution of motor behaviour. At the occipital lobe, the primary visual cortex is responsible for processing visual input (Andreassi, 2000). While this basic description of the brain's anatomy and related functionality is supposed to give an overview, it should be noted that it drastically simplifies the complexity of function and interaction. Contemporary neuroscience has outlined an understanding of the brain as a complex system in which the interaction of multiple subcomponents is more likely to explain function (Buzsaki, 2006). Therefore, investigation of brain processes needs to focus not only on location to study function, but interaction in time and space. In this dissertation, the emphasis is placed on the observation of time dynamics through the observation of oscillatory, electrical patterns of neuron assemblies (Cohen, 2017; Buzsaki, 2006; Andreassi, 2000).

Oscillatory patterns describe a rhythmic activity in time, in this case electrical discharge of neurons. Functionally speaking, neural oscillation describes a form of self-organization of individual neurons and the brain as a whole (Buzsáki and Draguhn, 2004). Numerous studies have revealed that both individual neurons and assemblies of neurons resonate and oscillate at multiple frequencies at the same time (Buzsáki and Draguhn, 2004). The likely reason for the oscillatory resonance of neurons is argued to reside in two main properties: (1) the precise timing allowed from (variable) rhythmic behaviour allows a simple encoding of information, (2) oscillation is an energetically cheap principle with sufficient variability to achieve coordination, thus explaining the ability of the brain to synchronize neural activity across local and distant neuronal networks (Buzsáki and Draguhn, 2004). While slower rhythms are said to synchronize large spatial domains, faster rhythms bind together smaller assemblies (Buzsáki and Draguhn, 2004; Siegel, Donner, and Engel, 2012). Electroencephalography (EEG) allows to observe this rhythmic electrical activity of neuron assemblies in the brain, providing promising means for continuous user state detection (Blankertz et al., 2016; Wascher et al., 2019).

Electroencephalography (EEG)

Within neuroscientific research methods, electromagnetic observation (in particular EEG) has in the last century become a central tool for the study of neural activity due to its ability to cost-effectively, noninvasively, and directly measure the electrical activity of neuron assemblies at a high temporal resolution of millisecond changes (Cohen, 2017; Andreassi, 2000; Pizzagalli, 2007). Furthermore, information from the EEG has been shown to link to perception, cognition, affect, and motor action reliably through various types of features (Buzsáki and Draguhn, 2004; Pizzagalli, 2007; Cohen, 2017; Bridwell et al., 2018). The EEG measures synchronously oscillating electrical discharges of neuron assemblies, predominantly in the cerebral cortex (Pizzagalli, 2007; Andreassi, 2000). This means that the EEG only records some of the electrical activity of the brain (i.e. not of deep subcortical regions) (Müller-Putz, Riedl, and Wriessnegger, 2015). The EEG is recorded from multiple electrodes typically positioned symmetrically across the head surface based on an internationally standardized placement system (the 10-20, or 10-5 system for >75 electrode sites - see Figure 4.11). Electrical activity measured at the scalp by EEG is always a relative value, specifically, the difference in electrical activity between two or more electrodes

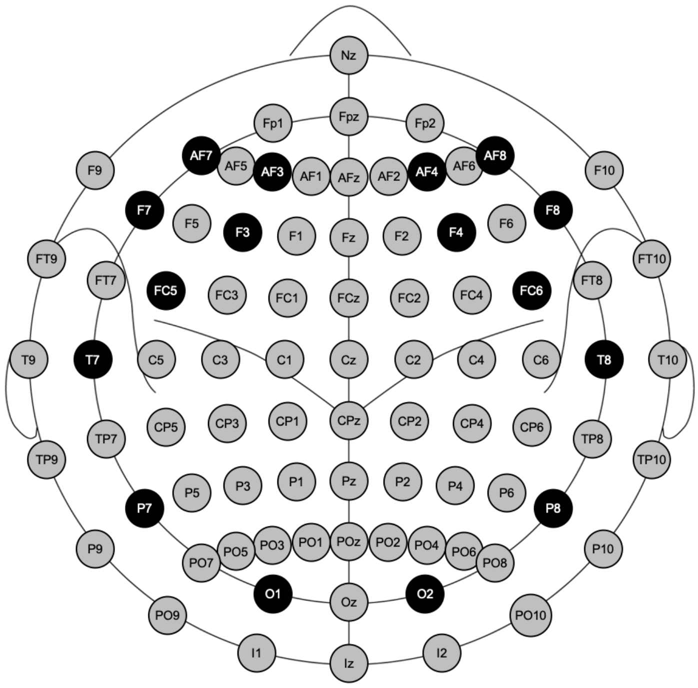

FIGURE 4.11: Electrode Positions in EEG Research - The Extended International 10-20 System by Oostenveld and Praamstra (2001). Black Locations Are Available in the Emotiv Epoc+ EEG Headset Which is Used in This Dissertation.

(Cohen, 2014; Pizzagalli, 2007). The placement of reference electrodes can occur in multiple forms and plays an important role in the analysis of EEG waveforms, as they can exacerbate, dampen or regionally distort the relative EEG readings (Müller-Putz, Riedl, and Wriessnegger, 2015; Pizzagalli, 2007; Cohen, 2014).

Generally, features from the EEG are derived from the time-domain (i.e. voltage changes in time) and the frequency-domain (changes in energy of sinusoidal oscillations) (Cohen, 2014; Bridwell et al., 2018; Keil et al., 2014). Features in both dimensions have been strongly linked to aspects of cognition (Bridwell et al., 2018; Cohen, 2017; Siegel, Donner, and Engel, 2012; Pizzagalli, 2007), yet show stark contrasts in the way they are recorded and extracted from the EEG signal. In the time-domain, two types of ongoing EEG activity are considered that are spontaneous (or continuous) EEG and Event Related Potential (ERP) (Müller-Putz, Riedl, and Wriessnegger, 2015). Spontaneous EEG is the measure of omnipresent, ongoing neural activity in the brain that fluctuates with typical (peak-to-peak) amplitudes under 75 microvolts (Müller-Putz, Riedl, and Wriessnegger, 2015). ERP features as their name suggests are phasic reactions to externally introduced stimuli (e.g. visual, auditory, somatosensory, or olfactory) that manifest in either positive or negative voltage changes visible at a particular point in time after the stimulus introduction (Müller-Putz, Riedl, and Wriessnegger, 2015; Picton et al., 2000; Cohen, 2014). In contrast, frequency-domain features are more similar in their form of recording and analysis to ECG features. In this dissertation, they are primarily used for this similarity and their previous use in flow neurophysiology research (see Section 4.3).

EEG Frequency Power Features

In the frequency domain, sinusoidal oscillations are commonly described in ranges describing their oscillatory speed (Buzsáki and Draguhn, 2004; Andreassi, 2000; Müller-Putz, Riedl, and Wriessnegger, 2015; Cohen, 2014; Pizzagalli, 2007). The function of these frequencies is still subject to much debate (see, e.g. Cohen, 2017). However, it is known that five major frequency bands differ reliably with relations to specific cognitive aspects, and also with topographical distributions and temporal characteristics (Cohen, 2017). A schematic extraction of these frequencies from a continuous EEG signal is shown in Figure 4.12. Delta waves (1-4 Hz) are considered to be an inhibitory mechanism and are typically associated with sleep states and some neurological pathologies (e.g. unconsciousness or coma) (Pizzagalli, 2007; Müller-Putz, Riedl, and Wriessnegger, 2015). Theta waves (4-8 Hz) are associated with some sleep states, meditation, and drowsiness (with a widespread scalp distribution), but also with increased mental workload (over frontal midline positions) (Müller-Putz, Riedl, and Wriessnegger, 2015; Silvestrini, 2017; Borghini et al., 2014; Pizzagalli, 2007). The Anterior Cingulate Cortex (ACC), a neural structure situated just behind the medial Prefrontal Cortex (PFC) is considered as the generator of this frontal Theta activity (Pizzagalli, 2007). Alpha waves (8-13 Hz) are found to relate to relaxed wakefulness and are majorly implicated as inhibitory processes, specifically as a marker of cortical idling (Müller-Putz, Riedl, and Wriessnegger, 2015). Alpha rhythms are most prominent over posterior regions (Pizzagalli, 2007). Beta waves (13-25 Hz) have been related to mental states such as active concentration, task engagement, excitement, attention, or vigilance and typically present with a symmetrical fronto-central distribution (Müller-Putz, Riedl, and Wriessnegger, 2015; Pizzagalli, 2007). Beta activity is therefore concerned to be primarily an excitatory mechanism (Müller-Putz, Riedl, and Wriessnegger, 2015). Beta band activity typically replaces Alpha band activity during cognitive activity (Pizzagalli2007). Beta band activity has also repeatedly been found related to increases in mental workload over frontal, temporal and posterior sites (Michels et al., 2010). Gamma waves (25-200Hz, but mostly not much higher than 40 Hz) are often associated with arousal and perceptual binding (Müller-Putz, Riedl, and Wriessnegger, 2015; Pizzagalli, 2007).

These frequencies can occur both competitively and simultaneously within the same brain structure or over the whole scalp. One of the major drawbacks of EEG observations is the quality of the spatial resolution and the aggregation of synchronous activity (Cohen, 2017). The main issue is the problem of source localization as it has been found that resulting electrical activity at the cortical surface can theoretically be traced back to an infinite number of possible sources, an observation termed the “inverse-problem” (Pizzagalli, 2007). Even though recent developments with larger electrode arrays propose that the EEG may eventually approach the spatial resolution of other neuroimaging techniques (Pizzagalli, 2007; Cohen, 2017), at the moment, EEG readings come with the grain of salt that sources of neuronal activity are difficult to localize from the recorded data. To overcome the inverse problem, it is of major importance to postulate physiologically and anatomically sound assumptions about putative EEG sources (Pizzagalli, 2007). Therefore, neuroanatomical reasoning on flow experience (i.e. considering the neuroscientific theoretical perspectives outlined in Section 4.3) is of high importance in this research. Lastly, similar to ECG observations, EEG research must control for possible confounding influences. For example, gender, age or handedness are known to influence anatomical development and can, therefore, confound sample comparability if such factors are not sufficiently

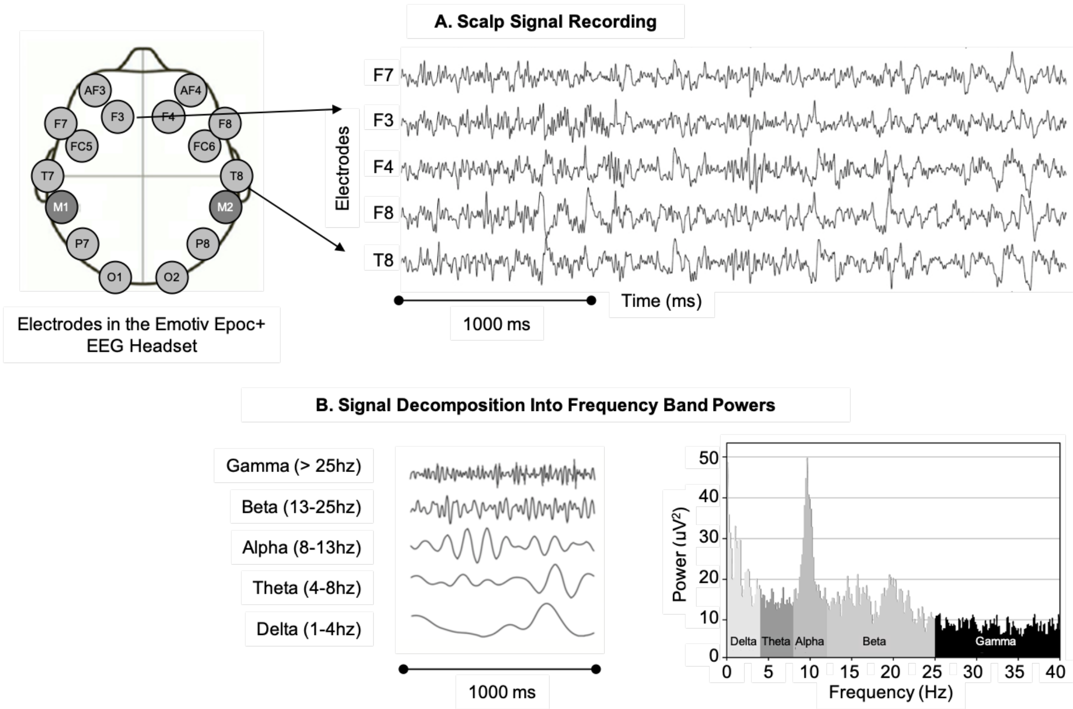

FIGURE 4.12: EEG Signal Decomposition Into Frequency Bands - See Buzsaki (2006) and Cohen (2014).

varied or controlled (Picton et al., 2000). Also, substance intake (e.g. recreational drugs) and behavioural factors like movement (of body and especially facial muscles) can lead to measurement artefacts (Cohen, 2014; Pizzagalli, 2007). Therefore, not only participant screening but also rigorous signal processing must be applied to derive substantial results. The procedures implemented in the experiments in this dissertation are outlined in Chapters 5 and 6 and Appendix Section A.4.

In concluding this background section, it can be summarised that the combination ECG and EEG measures for the description and continuous detection of flow experiences was considered promising, feasible, and timely in terms of how such instrumentation could become implemented in adaptive NeuroIS in the near future. Given the similarity in the electrophysiological measurement principles, these two methods were chosen to facilitate the endeavour that is the detection of flow experiences using neurophysiological measures across various tasks. The next chapters will now describe the procedures and results from the experiments that build on this theoretical, empirical and methodological foundation.

Chapter 5

Experiment 1 - Difficulty Manipulation & Experience Sampling

Contents of this section are in part adopted or taken from Knierim et al. (2018a) and Knierim et al. (2018b). See Section A.1 for further details.

5.1 Exploring Flow in More Naturalistic Knowledge Work

Building on the integration of previous flow research in the earlier chapters, the herein described experiment integrates paradigms and measurement approaches into a first experiment. This experiment lays the groundwork for the second and third research goal of this dissertation, to identify how limitations of present experimental paradigms can be overcome for the intensification of flow experiences in the laboratory (Research Goal (RG)2), and to identify which neurophysiological patterns of flow can be identified with wearable sensors across different measurement scenarios (RG3). Through this experiment, vital empiric evidence is collected that can aid in the future development of flow-facilitating adaptive NeuroIS.

While flow facilitation in real-world settings is still a substantial challenge due to the complex requirements (e.g. absence of distractions, the structure of the task, challenge of the task, physiological and psychological state of the individual) (Ceja and Navarro, 2012; Spurlin and Csikszentmihalyi, 2017), investigations under laboratory conditions have demonstrated sufficient flow manipulability for fundamental studies to build upon (Moller, Meier, and Wall, 2010; Keller, 2016). Nonetheless, a focus on artificial laboratory setups is also highlighted as a particular shortcoming in experimental flow research (see Chapter 3). This focus is especially visible for flow neurophysiology research (see Chapter 4). At present most of the neurophysiological research is conducted in highly controlled game tasks, leaving gaps to understand neurophysiological configurations during flow in primarily cognitive, and more unstructured tasks that are typical in Knowledge Work (KW) (Quinn, 2005; Moller, Meier, and Wall, 2010). Only a few experiments have so far employed mainly cognitive tasks for flow elicitation such as performing mental arithmetic (Ulrich et al., 2014) or matching invoices (Rissler et al., 2018) with adapted difficulties. These approaches require reproduction to ascertain their utility for laboratory flow research closer to the KW context. Furthermore, while there have been serious advancements in fields as affective computing (Picard, 2003), neuro-ergonomics (Parasuraman and Wilson, 2008), and brain-computer interfaces that keep extending the applicability of real-time physiological measurement and neuroimaging to in situ phenomena including attention, operator workload and engagement (Blankertz et al., 2016; Kosti et al., 2018), the study of neurophysiological correlates of flow in more naturalistic scenarios (i.e. closer to the real world) is still sparse (see Chapter 4).

As KW demand is estimated to increase strongly (see Chapter 1), Experiment 1 aims to build a bridge from controlled experimental setups towards more naturalistic settings by adapting an original flow research method, namely the Experience Sampling Method (ESM) (Csikszentmihalyi and Hunter, 2003) (see Chapter 3) to a laboratory setting. The ESM was developed to overcome interview limitations (e.g. recollection bias) and catch flow closer to its occurrence through repeated interruption (Csikszentmihalyi and Hunter, 2003). Only more recently, experimental flow induction has been developed with the primary paradigm of Difficulty Manipulation (DM) (Moneta, 2012). While the DM approach has been deemed useful to elicit contrasts, it has also been criticised as to whether deep flow experiences are elicited, given the low involvement often present in experiment tasks, and given the often artificial task nature (Moller, Meier, and Wall, 2010; Delle Fave, Massimini, and Bassi, 2011). This artificiality prevents participants from capitalising on high levels of pre-developed expertise (Ullén et al., 2010), and might require attentional processes incompatible with flow experience (Hommel, 2010). This shortcoming has created calls for more creative laboratory research on flow experiences (Harris, Vine, and Wilson, 2017b).

Therefore, by inviting knowledge workers to continue an ongoing, personal work project in a laboratory environment, a naturalistic task is combined with the controllability of task-external factors in the laboratory. In this controlled Experience Sampling (cESM) approach individuals can work on a personalised, yet comparably structured knowledge work task while being observed using neurophysiological sensors and being interrupted multiple times to "catch flow in the act". This cESM approach shares similarities to flow research with expert musicians (Harmat et al., 2011; Manzano et al., 2010) or chess players (Tozman, Zhang, and Vollmeyer, 2017) that perform their naturalistic task within the area of their expertise. However, as KW tasks are not generally similarly segmentable in this manner, the paradigm focuses on the continuation of a larger piece of work and utilises the repeated interruption of this work to observe changes in flow. Such naturalistic (i.e. closer to real-world) approaches also share an important distinction in terms of expertise and intrinsic motivation to approaches that have focused on Engagement (ENG) paradigms where participants are, for instance, asked to play an (unknown) game and report their experience afterwards (e.g. Labonté-Lemoyne et al., 2016; Shearer, 2016). Importantly, naturalistic approaches for the KW context have not yet been investigated previously. By analysing experience across interruptions, and by comparing them to a standard flow induction approach, this experiment aims to answer the main Research Question (RQ) of how well the cESM approach can elicit flow. More specifically, the RQs to be answered in Experiment 1 are:

- **RQ3:** Is the laboratory-based flow elicitation using a mental arithmetic DM task reproducible?
- **RQ4:** Is flow elicitation in the laboratory intensified by a more naturalistic task setting (through the cESM approach)?
- **RQ5:** Which neurophysiological correlates of flow can be observed across different cognitive task scenarios with wearable sensor systems?

To summarise, in Experiment 1, the experience of flow is observed in two cognitive tasks (mental arithmetic and scientific writing) across two experimental paradigms, one established as reference (DM), and one newly developed to provide increased naturalism in the controlled laboratory environment (cESM). Together this work contributes to the literature on flow experience, and the development of adaptive NeuroIS by (1) advancing the understanding of flow elicitation in laboratory settings, by (2) extending flow neurophysiology research to the KW context, and by (3) delivering insights into flow neurophysiology (specifically changes in HRV and frontal EEG powers) across tasks.

5.2 Experiment Basis

As discussed in the chapter on experimental paradigms for flow elicitation (see Chapter 3), the manipulation of a task's difficulty is used as a primary means to create flow experience contrasts in the laboratory. DM has been found to work sufficiently well across a variety of tasks (Keller, 2016; Moller, Meier, and Wall, 2010). It has therefore been employed primarily with digital games (e.g. Tetris - see Ewing, Fairclough, and Gilleade, 2016, Space Invaders - see Rheinberg and Vollmeyer, 2003, Racing Games - see Tozman et al., 2015, or Pacman - see Harmat et al., 2015), but also with expert musicians (Harmat et al., 2011) or chess players (Tozman, Zhang, and Vollmeyer, 2017), and with primarily cognitive tasks like Tower of London Reasoning (Chatterjee, Sinha, and Sinha, 2016), the n-back working memory task (Fairclough et al., 2013), or knowledge trivia (Keller et al., 2011). With a focus on KW settings, it was for this dissertation decided to utilise a simple cognitive task DM instance. Specifically, a mental arithmetic task has been used which has been created by Ulrich et al. (2014), Ulrich, Keller, and Grön (2016b), and Ulrich, Keller, and Grön (2016a). While the instantiation has undergone some alterations in all experiments of this dissertation, the reference to the initial design by Ulrich et al. (2014) represents the anchor for the herein derived results. For this reason, and to ease the comparability of the design changes, the initial design is outlined here in detail.

In the original mental arithmetic DM design by Ulrich et al. (2014), two or more numbers have to be mentally summed and entered. In every trial, the result had to be a three-digit number. Participants were asked to enter the result as "accurately and fast as possible" (Ulrich et al., 2014, p. 195). The time for each trial was 18 seconds, and a break of 4 seconds between trials was used. During this break, the expression "xxx + x" was shown. Once an answer was entered, the break expression was shown. If an answer was not provided, the trial was aborted after the 18 seconds, and the break expression was shown.

The trials were delivered in three conditions. These conditions were configured to deliver very easy ("Boredom" condition) equations non-adaptively, or dynamically-adapted, moderately difficult equations ("Flow" condition) or hard equations ("Overload" condition). In the "Boredom" (B) condition, only two numbers were shown, with the first summand randomly drawn from an interval of [100, 109], and the second is randomly drawn from an interval of [1, 9]. However, it was ascertained that a result could only be in the interval of [101, 110] so that no mental carries would have to occur. In the "Flow" condition, the difficulty would adapt dynamically based on the performance of the participant. Therefore, two or more numbers were shown. The higher the level, the more summands were shown. For a level increase, either, another single-digit summand was added or a second digit would be added to a single-digit summand if one was available. Level decreases used this logic inversely. The results of the last two trials (sliding window) were used to evaluate the task performance. If two out of these two trials were solved correctly/incorrectly, difficulty was increased/decreased by one level. The starting level for the flow condition was estimated during a five-minute calibration phase using the “Flow” condition logic. It is assumed (as it is not reported in Ulrich et al., 2014) that the initial level of the calibration condition used a two number, double-digit difficulty, where the result had to be a three-digit number (e.g. 65 + 73). The average level of the last 25% of trials was used to compute the starting level of the “Flow” condition. The “Overload” (O) condition operated with a similar adaptation logic as the “Flow” (F) condition. The differences were though that the initial difficulty level for the “Overload” condition was set to be three levels higher than the starting level of the “Flow” condition and could never fall below this starting level. Also, the difficulty was increased/decreased in the “Overload” condition, when three out of five/four out of five trials (sliding window) were correct/incorrect. A central goal of this approach was to keep the difficulty at a level too high for participants to master, but also to prevent permanent frustration and disengagement. Examples of equations by condition and level difference are shown in Table 5.1.

Lastly, an initial introduction phase using the “Boredom” treatment logic for five minutes was included in the experiment by (Ulrich et al., 2014). In the progress of this dissertation, it was found in pre-testing, that the length and difficulty of this initial phase could be critical to elicit a more boring task experience in the actual “Boredom” condition that follows later during an experiment. Shorter introduction phases appeared to retain an element of relaxation and excitement about being able to solve many simple tasks in quick succession during “Boredom” conditions. Each difficulty condition (besides introduction and calibration phase) lasted for 184 seconds and was presented three times in a block design that used two sequences (“R-B-F-O-F-R-O-B-O-B-F-R” and “R-B-O-F-O-R-F-B-F-B-O-R”). The “Rest” (R) condition denotes a 25-second eyes-open resting phase (white screen). After each condition, self-reports about flow experiences were collected, as were task preferences once at the end of the experiment.

As Ulrich et al. (2014) presented their task to participants in an fMRI scanner, the input modalities were peculiar in the sense that a trackball was used to control the on-screen keyboard to type in the numbers or to correct a mistake (see Figure 5.1). In the experiments of this dissertation, desktop computers with keyboards and mouses were used instead. Participants could immediately see their input which had the default value of “000”. Mistakes could be corrected within the trial time window. No feedback about the correctness of trials was shown to the participants. These design elements were replicated exactly for the experiments in this dissertation.

This concludes the description of the mental arithmetic task design using a DM paradigm by Ulrich et al. (2014). Support for the feasibility of the task to induce varying levels of flow according to the flow theory notion of challenge-skill balance/imbalance has been found in some additional attempts (Ulrich, Keller, and Grön, 2016b; Ulrich, Keller, and Grön, 2016a). However, as this task is considered as a reference in the experiments of this work, its reproducibility represents a cornerstone for the evaluation of the flow intensification efforts in this dissertation.

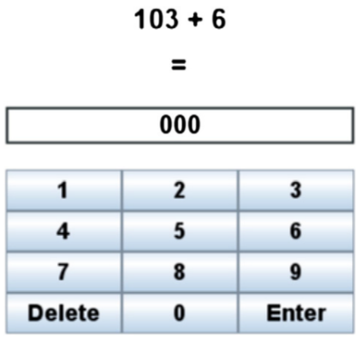

FIGURE 5.1: Mental Arithmetic Task from Ulrich et al. (2014). Excerpt from a Screenshot During the "Boredom" Condition.

5.3 Experiment Design & Preparation

5.3.1 Materials

In this experiment, each participant worked on: (1) solving arithmetic equations manipulated in difficulty, and on (2) writing a research project report both within the same session.

Mental Arithmetic DM Task

The mental arithmetic task was chosen as a reference to a validated DM task (Ulrich et al., 2014). Replicating the design by Ulrich et al. (2014), participants sum two or more numbers, depending on the active and dynamically adjusted difficulty level. Slight adjustments were made to the design and procedure due to pre-test findings and due to the requirement of including two tasks into the experiment. First, the task difficulties were found to be too high in pre-tests. This complication might be caused by utilising a different student sample. While the participants in Ulrich et al. (2014) were most likely psychology students, the students from the Karlsruhe Institute of Technology (KIT) experiment pool predominantly pursue study majors in business and engineering sciences and were considered to have a different affinity or ability level for arithmetic tasks. Therefore, the Low Difficulty (EASY) condition was further simplified by only drawing equations in one of three forms (101 + 1, + 2, or + 3). Also, for the Calibrated Optimal Difficulty (CAL) condition, the difficulty adaptation logic was changed so that difficulty increased/decreased when three (not two) sequential responses were correct/incorrect. This alteration was included to slow down the up-levelling process that was found to lead to overly high difficulties too soon, which might be a consequence of another alteration, the extension of the condition duration to five minutes (instead of three minutes in Ulrich et al., 2014). The longer duration was required for the observation of HRV features for which the five-minute window is an essential basis (Malik et al., 1996). In contrast, to keep the overall duration of the experiment at bay, the introduction phase (three minutes) and the calibration phase (four minutes) were shortened. Also, the difficulty conditions (EASY, CAL, HARD) were only presented once (not three times as in Ulrich et al., 2014), for the sake of keeping the experiment duration acceptable to participants. As a final alteration, eyes open resting phases (60 seconds) were included after each difficulty condition. These washout phases were included to reduce the risk of carry-over effects between the conditions. The rest of the approach is the same as in Ulrich et al. (2014) and as outlined in Section 5.2. A screen-shot of the math task is shown in Figure 5.2, and a simplification of the task difficulties is shown in Table 5.1, with the changes highlighted in comparison to the design by Ulrich et al. (2014).

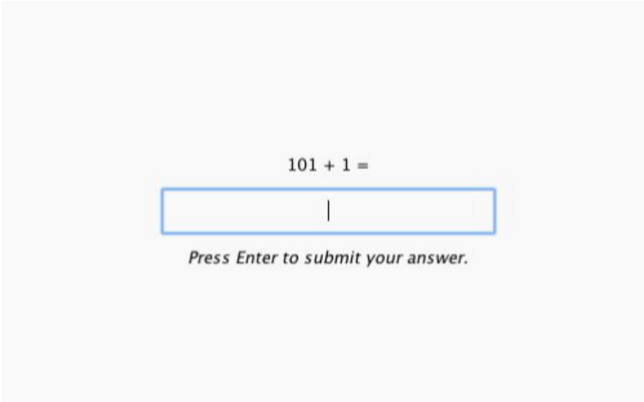

FIGURE 5.2: Mental Arithmetic Task in Experiment 1. Excerpt from a Screenshot During the EASY Condition.

Condition	Level	Ulrich et al. 2014	Experiment 1
EASY	0	103 + 6	101 + 2
CAL	1	65 + 73	
CAL	2	58 + 91 + **4**	
HARD	15	72 + 12 + 32 + 67 + 29 + 58 + 63 + 14 + 45	
HARD	16	19 + 46 + 55 + 26 + 73 + 49 + 57 + 10 + 34 + **5**	

Notes: Additional Digits with Level Increase are Highlighted in Bold.

TABLE 5.1: Mental Arithmetic Task Difficulties in Experiment 1.

Writing cESM Task

To provide an approach for flow intensification, a more naturalistic paradigm was developed for the KW context. This approach combines expertise and intrinsic motivation (through the continuation of a personally relevant project) with the options to control the task environment in the laboratory. Similar approaches have been employed with expert musicians, that were invited to laboratory settings and asked to repeatedly perform a musical piece that they find challenging (Manzano et al., 2010). However, as KW tasks are not generally similarly segmentable in this manner, the paradigm focuses on the continuation of a larger piece of work and utilises repeated interruption of this work to observe changes in flow. For the similarity of this more naturalistic approach to the traditional flow field research method (ESM), the approach is termed cESM. As a task instance for this paradigm, scientific writing was chosen due to its nature of being a challenging and frequent task for students (exemplary future knowledge workers). Also, writing (scientific or literary) has previously been related to engaging experiences in general and flow in particular (Csikszentmihalyi, 1996; Erhard et al., 2014; Galluch, Grover, and Thatcher, 2015). Participants brought their own thesis project to work on for a session of 23 minutes

(3· 7minutes for writing + 2 minutes of initial orientation - time for goal setting and survey interruptions not included). A similar, three-stage logic of (1) orientation, (2) calibration, and (3) action was employed to design the writing task aligned to the mental arithmetic task. The structural similarity is shown in Figure 5.6.

In the first stage (introduction/orientation), participants were given time to inspect the state of their document for two minutes, to facilitate the entry in the writing process that followed afterwards. In pre-tests, it was found that participants were well-prepared in advance to continue working on their thesis, being very familiar with their open tasks. Therefore, the short duration of two minutes was considered sufficient for this stage.

In the second stage (calibration), the preconditions for flow experience were operationalised by a structured goal-setting process. Specifically, participants were asked to define a challenging yet achievable goal, to create optimal difficulty circumstances for the following writing session. To standardise the goal-setting approach, the SMART mnemonic was used (Doran, 1981). The SMART goal setting format has been used extensively in previous literature, for example in academic planning (Day and Tosey, 2011; Lawlor, 2012; Moeller, Theiler, and Wu, 1988) and clinical rehabilitation (Bovend'Eerdt, Botell, and Wade, 2009; Bowman et al., 2015). The SMART letter operationalisations were chosen following guidelines by Bovend'Eerdt, Botell, and Wade (2009) and Bowman et al. (2015) and adopted to the present setting. The letters, therefore, herein refer to a goal that is specific (S), measurable (M), achievable (A), relevant (R) and time-bound (T). In addition to the SMART letters, participants are asked to provide a more generic, overall goal for their session. This approach was also considered to facilitate flow experiences. Setting a specific goal (S) (i.e. that is less abstract) has been found to facilitate high-quality writing outcomes (Flower and Hayes, 1981), and should provide the flow prerequisite of having goal clarity. In addition, deriving a goal attainment measure (M) was considered to help fulfil the second flow prerequisite of unambiguous feedback. Lastly, the focus on a relevant (R) and achievable (A) goal, was considered to enhance the optimality of a task challenge further. The choice to ask for the provision of "challenging" and "relevant" goal parts was made to ascertain flow prerequisites and general guidelines for proper goal setting are met (as outlined above). Furthermore, following guidelines on general pitfalls of goal setting (Latham and Locke, 2006), participants were told that they would not have to fear evaluation of goal attainment. Examples of what an expected goal would look like were provided to overcome limitations of individuals' abilities to goal-setting (Latham and Locke, 2006) and to increase consistency between participants. The feasibility of this approach was pre-tested, and testers reported no difficulties in defining their writing session goals according to the provided schema. Figure 5.3 shows a screen-shot of the goal-setting stage.

Afterwards, in the third stage (action), the writing session began. The thesis writing software was standardised to Microsoft Word in full-screen mode. This setup was chosen to provide a comparable yet familiar task environment for participants. The experiment software loaded the participants' documents automatically and switched between surveys and Microsoft Word autonomously. Participants were informed that their writing process was saved before each interruption. For a screen-shot of the task, see Figure 5.4. Participants were informed that they would be interrupted at some point during their session to report their experience during the writing process.

FIGURE 5.3: Goal-Setting Stage of the Writing Task in Experiment 1. Screenshot from a Participant With Completed Goals.

However, they were not informed about the timing and frequency of these interruptions. This approach was chosen to facilitate task concentration (the anticipation of a survey interruption should not detract participants). Also, the decision was made in reference to traditional ESM designs, where participants are asked to complete surveys at random times (but once per defined interval - e.g. 2 hours) to "catch flow in the act" (Csikszentmihalyi and Hunter, 2003; Gaggioli et al., 2013). After each interruption, participants were informed about how much time they had approximately left to complete their task (i.e. "more than half of the time" and "about a third of the time"). This information was included to enable flexible adaptation of task goals and processing if required. However, no information about actual time was displayed on the screens or in the experiment environment (participants were asked to remove their watches) to facilitate concentration on the task. Such recommendations are generally outlined for laboratory flow research (Moller, Meier, and Wall, 2010). Assuming that some time is required to re-enter the writing task after an interruption, and to keep the overall task duration in balance with minima (some time is needed to complete a thesis writing task), and maxima (the overall experiment duration needs to be contained) requirements, three writing periods of seven minutes were included in the design. The constant interruption frequency was selected to initially assess flow experience fluctuation, without additional variance from interruption timings. After the experiment, participants received a copy of their saved thesis progress.

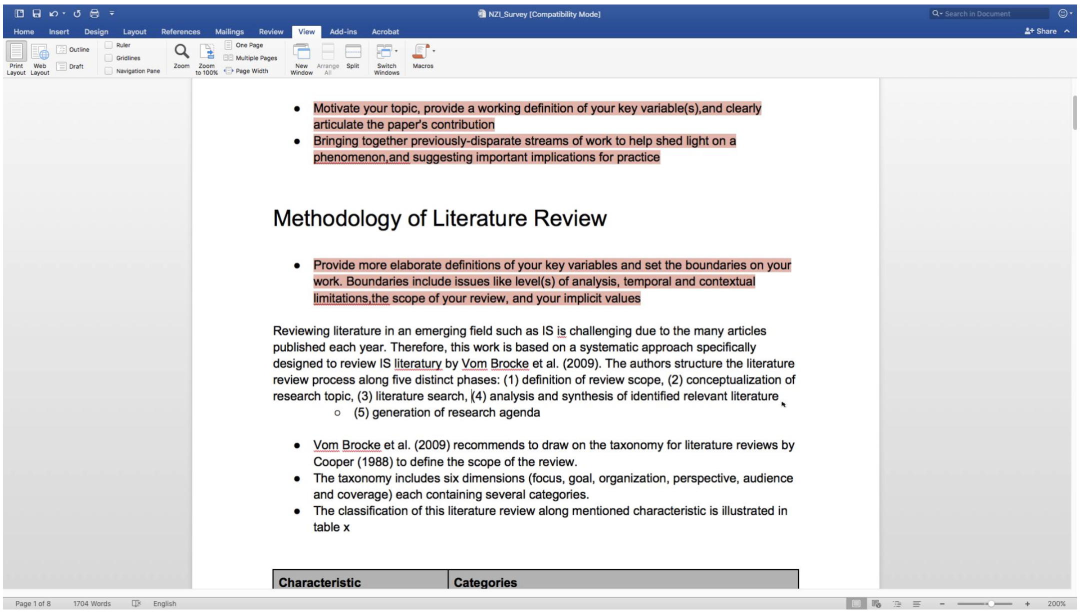

FIGURE 5.4: Writing Task in Experiment 1. Screenshot from a Participant During a Writing Period.

5.3.2 Procedure & Sampling

Experiment 1 was conducted in a laboratory setting with air-conditioned booths, one participant at a time. Each participant completed both the mental arithmetic and the writing task (within-subject design). The task order was randomised. Also, the three mental arithmetic task conditions were ordered randomly, which resulted in a total count of 12 procedure variations ($2 \cdot 3!$ combinations). All variations were executed once. At the start of the experiment, participants were welcomed, informed about the upcoming procedure and measurements. Afterwards, participants were asked to sign a consent form for their participation. Next, participants were guided to their booth in the laboratory and fitted with the physiological sensors on the head (EEG) and chest (ECG), and the signal quality adequacy was checked. Following, participants completed a first survey collecting demographic information and some initial state variables. Figure 5.5 shows the setup in the laboratory booth. To complete this preparation stage, participants then completed eyes-open and eyes-closed baseline phases in which they were asked to "let their mind wander to wherever it takes them", to keep their eyes focused on a black fixation cross on a white screen (in the eyes-open phase), and to avoid unnecessary movements. The same message and fixation cross were shown for the washout screens before each math task condition and between math and writing task. During the task stages, participants responded to surveys after each condition (round surveys), after each task (mental arithmetic or writing - task surveys), and at the end of the experiment. After the last survey

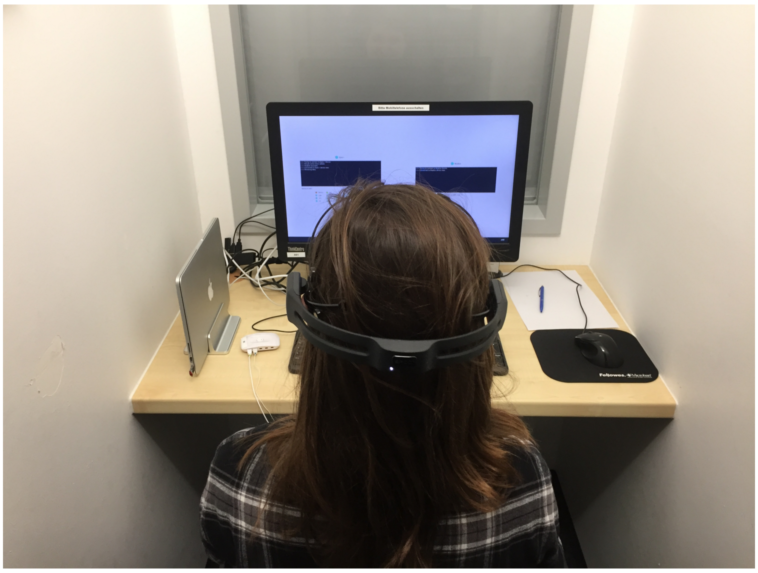

FIGURE 5.5: Photo of the Laboratory Booth Setup in Experiment 1.

was completed, sensors were removed, and participants were debriefed. Figure 5.6 outlines the complete procedure.

Twelve students (three female) participated voluntarily and were recruited from the circle of students working on a thesis project under the supervision of colleagues at the KIT. Table 5.2 summarises additional information on the sample. Participants were screened for being generally healthy, not taking any mind-altering medication, having full eye-sight (with or without correction), and abstinence of the consumption of alcohol, marijuana or other recreational drugs in the past 24 hours before the experiment. Furthermore, participants were asked to arrive at the laboratory with washed hair and not to use hair gel, hairspray, or similar products. In the recruitment survey, participants reported average thesis challenge levels of 4.3 (SD: 0.98), on average normal, but rather low levels of domain-general flow proneness (compared to the large sample used in Ullén et al., 2012), and on average rather low domain-specific flow proneness for writing tasks. The latter variable might be explained by the lower level of writing expertise, as the majority of thesis projects was still on the bachelor level. Paired Wilcoxon signed rank comparisons showed no difference in preference for math or writing tasks.

Variable	Counts / Distributions
N	12
Age (Mean / Median)	24,83 / 24
Gender (Female / Male)	3 / 9
Handedness (R / L / Ambi)	10 / 1 / 1
English Abilities $\geq$ B1	100%
Thesis Type	7 Bachelor Thesis / 4 Master Thesis / 1 Seminar Thesis (Master Level)
Thesis Challenge (Mean / Median - 7p)	4,33 / 4

TABLE 5.2: Sample Description for Experiment 1.

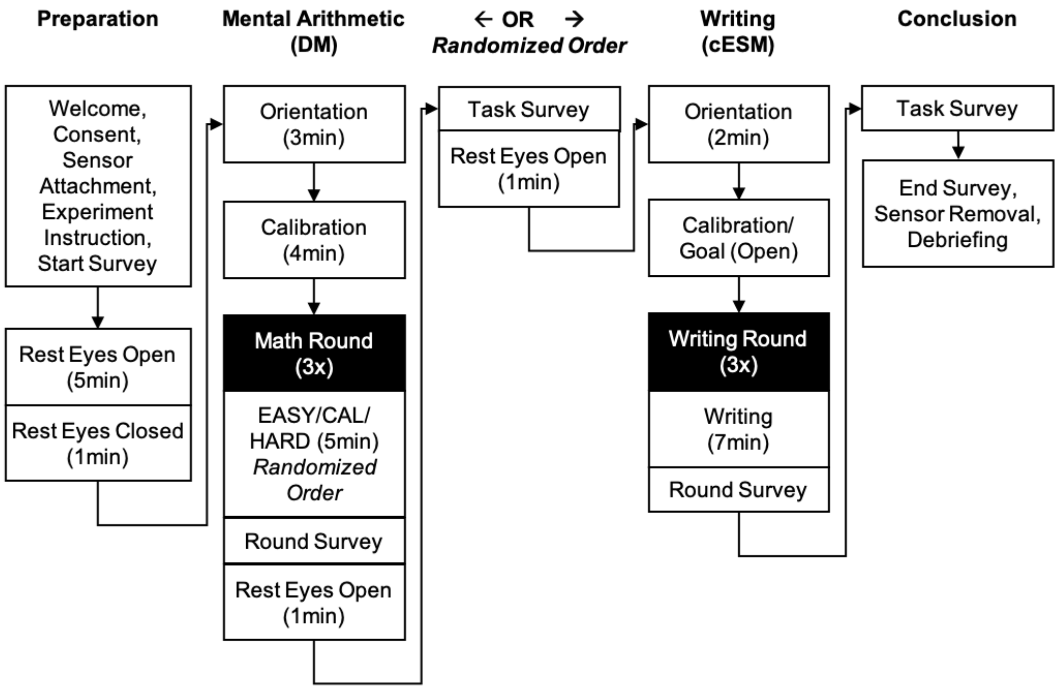

FIGURE 5.6: Procedure of Experiment 1 for Each Participant (Task Type and Difficulty as Within-Subject Factors).

5.3.3 Measures

Demographic information about age, gender, handedness, study majors, language abilities and thesis challenge levels (single item from Engeser and Rheinberg (2008)) were collected through a survey preceding the experiment session. This survey also collected information on thesis challenge levels. Afterwards, self-reported perceptions of experiences were collected at three levels: (1) after the mental arithmetic task conditions or after the writing task interruptions (both herein termed "rounds"), (2) after a task (mental arithmetic or writing - herein termed "tasks"), and (3) at the end of the experiment (herein termed "end"). Round questionnaires contained scales on flow and task demand (ten-item FKS and one additional task demand question all by Engeser and Rheinberg, 2008, stress (five-item construct by Tams et al., 2014, and affect (single question arousal & valence Self Assessment Manikin (SAM) scales by Bradley and Lang, 1994), amongst others. Task surveys included scales on task importance (Engeser and Rheinberg, 2008). End surveys included scales on mental arithmetic task preferences (Ulrich et al., 2014), and domain-general flow proneness (Ullén et al., 2012). Almost all questions used 7-point Likert scales (SAM arousal and valence used 9-point). Tables A.2, A.3, A.4, A.6, and A.8 in Appendix A.3 provide an overview of all measured variables, including the item operationalisations.

Neurophysiological data were collected using two electrophysiological methods. Particular emphasis was placed on the utilization of wearable sensors. The selected sensors were chosen, balancing the trade-off of acceptable signal quality and wearability. ECG data were collected in Lead II configuration using gelled electrodes. ECG data were sampled at a rate of 1000 Hz, and with a 12-bit resolution using a Biosignalsplux signal hub. EEG data were collected with an Emotiv Epoc+ headset. This 14-channel wireless headset uses saline-based electrodes, collecting data at a sampling rate of 256 Hz and with a 16-bit resolution. Electrode sites are AF3, F3, F7,

FC5, T7, P7, O1, O2, P8, T8, FC6, F8, F4, AF4 (according to the international 10-20 system - see Oostenveld and Praamstra, 2001). Two reference electrodes, the Common Mode Sense (CMS) and Driven Right Leg (DRL) were placed on the left and right mastoids (M1 and M2). While the headset comes with limitations in terms of data quality (primarily due to a non-gelled electrode contact basis), it has been found to deliver adequate data for the present type of study (e.g. a frequency power analysis, not an event-related potential study) (Barham et al., 2017). Also, the Epoc+ headset has been used in previous studies related to the KW context (Kosti et al., 2018), and related to flow experiences (Klarkowski, 2017). Before the application of the headset, the felt-pad electrodes were moistened with a standard 0,9%-NaCl saline solution. After the application, acceptable contact qualities were controlled for all electrode sites using the proprietary impedance information supplied by the manufacturer's application programming interface.

5.3.4 Data Processing

Based on the experiment design, round variables were measured at six points in time (repeated measures with six cells - three math conditions, three writing interruptions), with the variables nested in the tasks (mental arithmetic or writing) and difficulty manipulation or interruption number. All data for this experiment were processed along with the following schema: First, outliers in the data were identified and removed. The metric of ≥ 2 SD from the construct mean was chosen as a compromise between removing too much data and retaining severe outliers. Then, with the focus of conducting mean level comparisons, assumptions for parametric analyses of variance were tested. Normal distribution (Shapiro-Wilk test) and variance homogeneity (Fligner-Killen test) were violated for many samples, prompting the use of non-parametric tests. The choice for variance-based mean comparisons is rooted in their prevalence as a best practice in neurophysiological studies (see Chapter 4) and to account for the challenge of non-overlapping variance between reported data and observed neurophysiological processes (see Section 4.1). As instances of non-parametric tests (alternatives to the here designated one-way repeated-measures Analysis of Variance (ANOVA) and pairwise t-Tests), Friedman tests for main effects and Wilcoxon signed-rank tests were chosen. The details per data domain (report, or neurophysiology) are outlined in the following paragraphs.

For the self-report data, given the small sample, internal consistency was assessed for multi-item constructs by inspection of average item-total correlations, and Cronbach's Alpha values. One item was removed from the flow construct, as substantial improvements in reliability indicators were indicated. Table 5.3 shows all indicator values and documents acceptable internal consistency of the used constructs.

Variable (Items Retained)	Cronbach's Alpha	Avg. Item-Total Cor.
Flow (9/10)	0.81 / 0.94	0.63 / 0.83
Stress (5/5)	0.84 / 0.82	0.82 / 0.81

Notes: First Value = Mental Arithmetic Task, Second Value = Writing Task; Numbers in Parentheses = Retained/Measured Items.

TABLE 5.3: Latent Variable Internal Consistencies in Experiment 1.

ECG data were processed following the guidelines of Malik et al. (1996). Data from all participants were included. ECG data were primarily processed using the Python toolboxes BioSppy (Carreiras et al., 2015) and NeuroKit (Makowski, 2016) to derive time-series data of adjacent heartbeat intervals (RR-intervals). Afterwards, based on the RR-interval data, HRV features were computed in the same toolbox and cross-validated using the R toolbox RHRV (Martinez et al., 2017). In terms of extracted features, HRV metrics that have been central to previous flow physiology research (see Chapter 4.2) were selected for analysis. Similar to related research (Harmat et al., 2015; Tozman et al., 2015; Keller et al., 2011) change scores were used in the analysis (ΔHRV = HRV$_{task}$ - HRV$_{baseline}$) of five-minute window time-domain (SDNN, and RMSSD) and frequency-domain (LF-HRV, and HF-HRV) features preceding each survey. For the definition of these metrics, see Malik et al. (1996). The complete ECG processing pipeline is summarised in Appendix A.4 Table A.9.

EEG data were processed following the guidelines of Cohen (2014). Appendix A.4 summarises the complete feature extraction pipeline. Data were processed for a homogenised sub-sample (three female participants were excluded) to lower the impact of gender-based variability (Picton et al., 2000). Also, two data sets were excluded due to recording failures. The retained sample comprised seven right-handed males. Data preparation, feature extraction, and analysis were conducted in R, signal processing and artefact removal in EEGLab (Ver. 14.1.1) (Delorme and Makeig, 2004). Initially, experiment phases of interest were extracted (eyes-open baseline, all three math task conditions, all three writing task rounds) and channels centred through mean subtraction. Afterwards, the extracted data were loaded into EEGLab where a 0.5-45 Hz bandpass and a 50 Hz notch filter were applied. Signal data were then inspected for artefact removal. First, channels that had failed to collect data were removed. Then, paroxysmal artefacts were removed manually. Afterwards, using the infomax algorithm, an Independent Component Analysis (ICA) was performed to remove data components related to eye blinks and sideway saccades. Next, frequency bands were extracted for the frontal electrodes (AF3, F3, F7, FC5, FC6, F8, F4, AF4) similar to Ewing, Fairclough, and Gilleade (2016) based on 2s long epochs with 50% overlap and tapered using a Hann windowing function. Average band power (μV^2) was extracted using the Fast Fourier Transformation (FFT). Only artefact-free and complete epochs were used (epochs containing more than 95% of required samples, i.e. $> 2s \cdot 256$ Hz = 512 samples). Extracted frequency bands are Theta (4-8 Hz), Alpha (8-12 Hz), and Beta (12-30 Hz). Also, for the Alpha and Beta band, additional subsegments were extracted that are LoAlpha (8-10 Hz), HiAlpha (10-12 Hz), LoBeta (12-15 Hz), MidBeta (15-20 Hz), and HiBeta (20-30 Hz). Afterwards, frequency powers were normalised (natural logarithm transformation). Electrodes were pooled by computing the mean for three regions of interest (ROI) that are all frontal sites (AF3, F3, F7, FC5, FC6, F8, F4, AF4), left frontal sites (AF3, F3, F7, FC5), and right frontal sites (FC6, F8, F4, AF4). Next, feature epochs were aggregated temporally by computing the median over each experiment phase. Median use was preferred as a way of conservative data interpretation, taking care of potential outliers (Cohen, 2014). Finally, to facilitate comparisons between experiment phases, change scores were computed by subtracting the eyes open baseline phase mean from each experiment phase (e.g. ΔTheta = Theta$_{task}$ - Theta$_{baseline}$). For additional analysis of shorter periods, the same procedure was repeated on 30 second long epochs within each condition (round). The window length of 30 seconds was chosen based on the report by (Soltész et al., 2014), who argue that at the start of phases, temporal differences could occur in this interval already.

5.4 Results

5.4.1 Manipulation Checks

Difficulty Manipulation

In the first step, as a difficulty manipulation check, the perceived difficulty variable was inspected (similar to Keller et al., 2011; Tozman et al., 2015). Friedman tests showed a highly significant main effect with large effect size, indicating a high degree of variation in perceived difficulties throughout the experiment (see Table 5.4). Follow-up pairwise Wilcoxon signed rank tests further showed a pattern supporting manipulation success (see Figure 5.7). Specifically, significant differences were found between all mental arithmetic task conditions, showing stepwise increasing difficulty from EASY to High Difficulty (HARD) conditions, confirming manipulation success. The perceived difficulty level in the CAL condition (mean of 5.18) suggests that the task was potentially still slightly too difficult for an optimal demand level. Aside from one trend level difference between writing sample 1 and 3, no differences were found within the writing task, indicating a rather consistent task difficulty during the writing phase, with a potential decrease towards the task end. The perceived difficulty levels in the writing samples lay consistently between the mental arithmetic EASY and HARD condition, and possibly below the CAL condition, indicated at trend level for the first two writing rounds and the 0.05 significance level for the last writing round. The absolute levels of perceived difficulty during the writing task (means of 3.73 to 4.42) suggest a task demand level close to optimality.

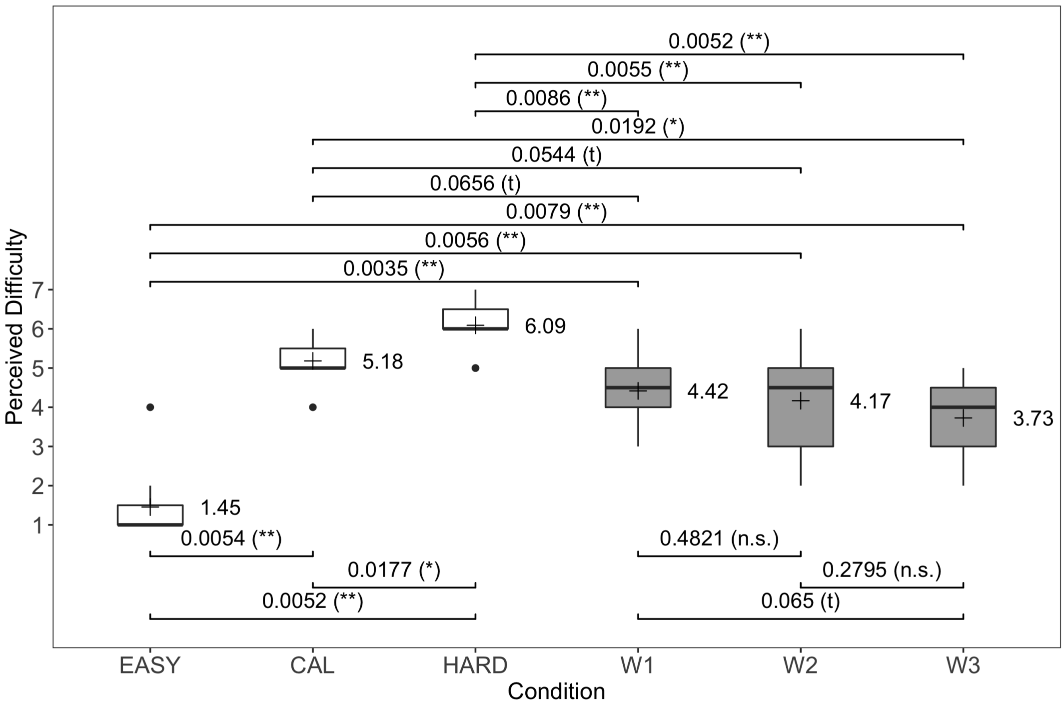

FIGURE 5.7: Perceived Difficulty Reports in Experiment 1. Crosshairs and Numbers Next to Them Represent Means.

Goal Setting Usefulness

To evaluate the potentiality of flow experience in the writing task beyond the calibration of difficulties, participants were asked at the end of the experiment about how well the goal-setting procedure worked to set clear goals for the writing task. Two questions were created for this purpose, one asking about the helpfulness of the goal-setting procedure. The first asked about the helpfulness directly: "How helpful was the goal to guide your actions in writing?". The second asked about its utility indirectly by asking about the need to adjust the self-set goals throughout the writing task: "How much did you have to adjust your goal as you moved forward in writing?". Histograms with the responses (7-point Likert scales) are shown in Figure 5.8. The results show that most participants found the procedure helpful (mean = 5.17), and did not have to adjust their goal in substantial ways (mean = 3.08). However, it should be noted that these findings are not unanimous, which means that there might still be room for improvement in this goal-setting procedure. Together with the manipulation of difficulty, these findings are taken as support that flow preconditions were likely instantiated sufficiently for the writing task.

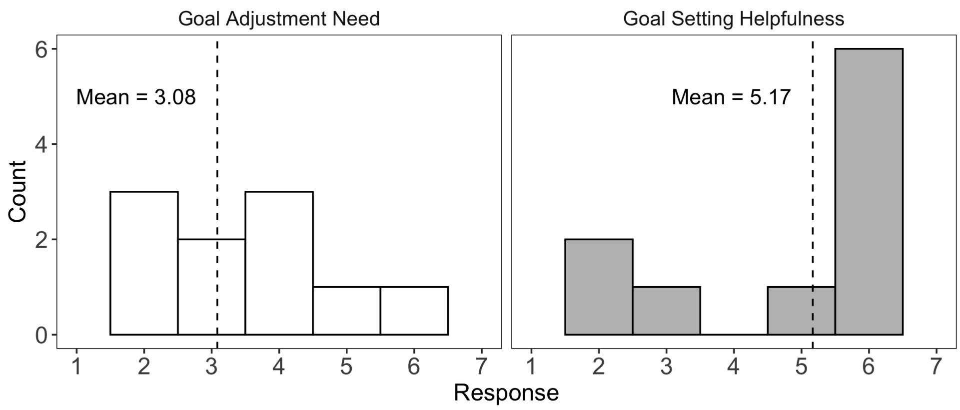

FIGURE 5.8: Goal-Setting Usefulness for the Writing in Experiment 1.

Task Importance

To complete the comparison of the two tasks in terms of flow experience likeliness, participants indicated the level of importance they put on both tasks. No significant differences were found in a paired Wilcoxon Signed Rank test (W = 20.5, p = 0.491, mental arithmetic mean = 3.82, writing mean = 4.00). This finding means that at least in terms of task importance flow experience ought to be equally likely in both tasks. Assuming comparability of task importance and successful manipulation (in particular for the mental arithmetic task), variables related to flow experience were investigated further to assess variation of flow in line with theoretic expectations.

5.4.2 Flow & Related Experiences

Flow

Flow experience was assessed as a construct comprised of perceptions regarding fluent action and absorption in the task (using the FKS scale by Engeser and Rheinberg (2008) and Rheinberg and Vollmeyer (2003)). Significant main effects are found with a moderate effect size (see Table 5.4). Comparisons of the flow reports (see

Figure 5.9) indicate significant differences between the mental arithmetic CAL and HARD condition only. These findings mean that flow reports only partially support theoretical predictions of maximised flow during optimally difficult task conditions (an inverted U-shape). Within the writing task, there were no significant differences, indicating a consistent flow experience level in writing. Additional support for this observation is also found in the within-subject range of flow reports across all writing rounds (mean range = 1.13, SD = 0.62). Across both tasks, repeated, significant differences are found for flow between the mental arithmetic EASY and HARD conditions with the writing rounds. Also, there is a trend level indication of higher flow in the first writing sample than in the mental arithmetic CAL condition. Therefore, flow was reported at least as high in writing as in the mental arithmetic CAL condition, and potentially higher in the first round of writing.

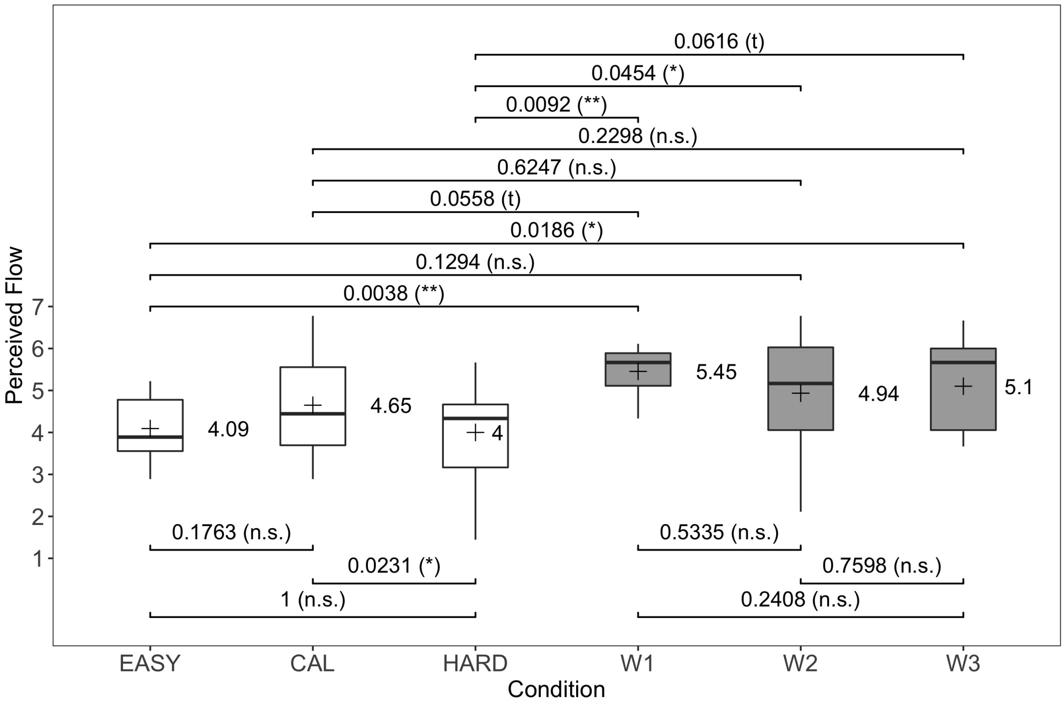

FIGURE 5.9: Flow Reports in Experiment 1. Crosshairs and Numbers Next to Them Represent Means.

Together with the perceived difficulty findings, the results suggest that within the mental arithmetic task, a maximum of flow experience was most likely during the CAL task condition. The lack of separation between the EASY and CAL condition is potentially caused by high levels of perceived fluency during the very easy task. Similar complications have been reported in related work (Peifer et al., 2014; Peifer et al., 2015; Tozman, Zhang, and Vollmeyer, 2017), and some work has therefore also included participant action slowing-mechanisms during these easy task conditions (Keller et al., 2011; Harmat et al., 2015). Nevertheless, together with the results from the other variables below (perceived stress and neurophysiological data), it is considered that flow was most likely at its highest intensity in the CAL condition within the mental arithmetic task and in the first writing round across all measurements.

Dependent Variable (DV)	Test Statistic	P-Value	Effect Size
Perceived Difficulty	$\chi^2(5) = 31.1914$	<0.0001	0.7798 (L)
Perceived Flow	$\chi^2(5) = 22.5397$	0.0004	0.4098 (M)
Perceived Stress	$\chi^2(5) = 32.9056$	<0.0001	0.6581 (L)
Notes: Reported Effect Sizes are Kendall's W (Kendall and Smith, 1939), Interpreted as: 0.1 - < 0.3 (Small = S), 0.3 - < 0.5 (Moderate = M) and >= 0.5 (Large = L) Effect Sizes - See (Tomczak and Tomczak, 2014).			

TABLE 5.4: Friedman Tests for Self-Reports in Experiment 1.

Stress

Perceptions of stress were collected and evaluated to assess the presence of motivated performances (see Tozman et al., 2015). Significant main effects with large effect sizes are found for the perceptions of stress. Post-hoc comparison of stress reports (see Figure 5.10) revealed significant differences between all three mental arithmetic task conditions, with stepwise increasing stress levels from EASY to HARD. Within the writing task, the stress levels did not differ significantly. Stress levels were consistently lower in the writing task than in the mental arithmetic CAL and HARD conditions. The results altogether indicate low stress perceptions during the writing and easy mental arithmetic task conditions.

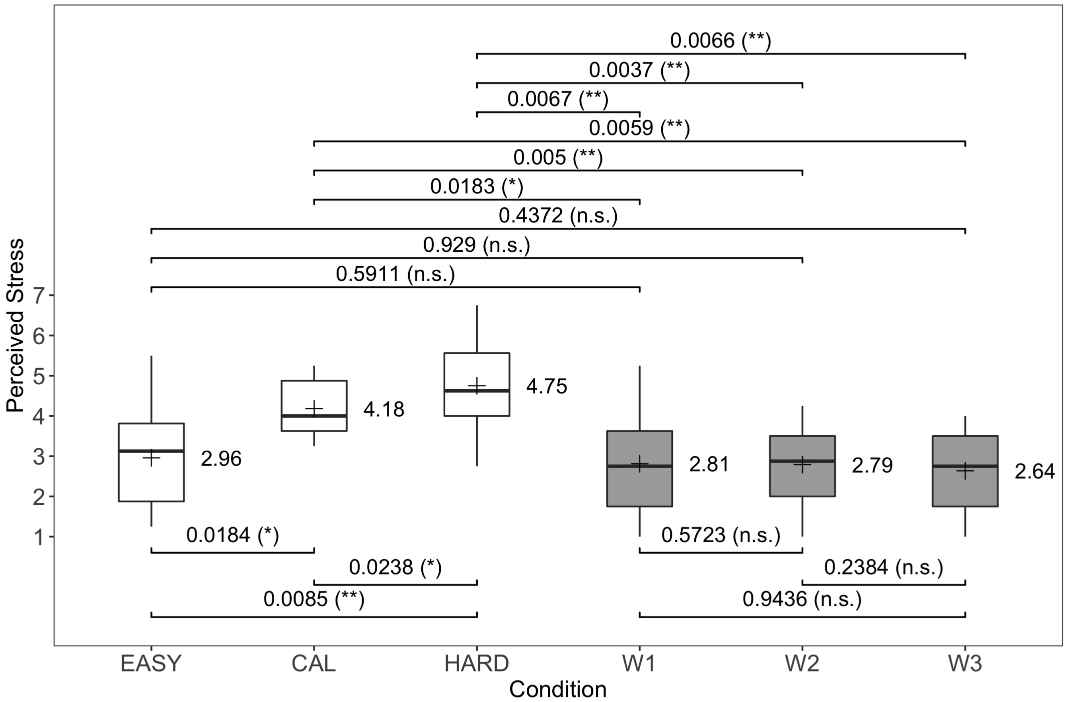

FIGURE 5.10: Stress Reports in Experiment 1. Crosshairs and Numbers Next to Them Represent Means.

These results show both expected and surprising results. On the one hand, increased arousal and stress have been reported to occur with increasing difficulties (see, e.g. Tozman et al., 2015; Klarkowski, 2016), which is why the patterns for the mental arithmetic task are in line with related literature. Interestingly, however, stress perceptions were low in all rounds of the writing tasks, despite greater perceived difficulties than in the mental arithmetic EASY condition, and despite similarly elevated flow levels as in the mental arithmetic CAL condition. The role of these findings is discussed further after the presentation of the neurophysiological results, that extend the picture further on these dynamics of difficulty, flow, and stress.

5.4.3 Neurophysiological Results

ECG Results

First, variations in heart rate variability HRV were assessed to follow up on examining how the cardiac system operates during flow, both within the traditional DM paradigm and across paradigms and tasks. Friedman tests of main effects across the sampling points were only significant for RMSSD and HF-HRV, in both cases indicating small effect sizes (see Table 5.5). RMSSD and HF-HRV are considered two typically positively correlated HRV indicators of parasympathetic ANS branch activity (Berntson, Quigley, and Lozano, 2007). Both HRV metrics show similar patterns in post-hoc test results (see Figure 5.11 and Figure 5.12). For the RMSSD feature, Wilcoxon tests of mean differences show a significantly higher level in the mental arithmetic CAL condition than in the HARD condition. This finding was not corroborated by the HF-HRV feature, which indicated similar HRV levels across mental arithmetic task conditions, albeit with a stepwise decrease in means from EASY to HARD. No significant difference across sampling points was found for all the writing task rounds, in both features. However, importantly, task comparisons showed significantly lower HRV levels in both RMSSD and HF-HRV in comparison to both mental arithmetic CAL and HARD conditions. HF-HRV was also significantly lower in all writing task conditions compared to the mental arithmetic EASY condition, indicating a consistently stronger withdrawal of parasympathetic cardiac modulation during the writing task. Together these results indicate that flow is likely related to increased physiological activation (moderate or low HRV as indicated by parasympathetic cardiac modulation features).

Dependent Variable (DV)	Test Statistic	P-Value	Effect Size
ΔRMSSD	$\chi^2(5) = 13.2000$	0.0216	0.2640 (S)
ΔSDNN	$\chi^2(5) = 4.5064$	0.4790	-
ΔPNN50	$\chi^2(5) = 7.4156$	0.1915	-
ΔHF-HRV	$\chi^2(5) = 12.2987$	0.0309	0.2236 (S)
ΔLF-HRV	$\chi^2(5) = 4.1429$	0.5290	-

Notes: Reported Effect Sizes are Kendall's W (Kendall and Smith, 1939).

TABLE 5.5: Friedman Tests for HRV Features in Experiment 1.

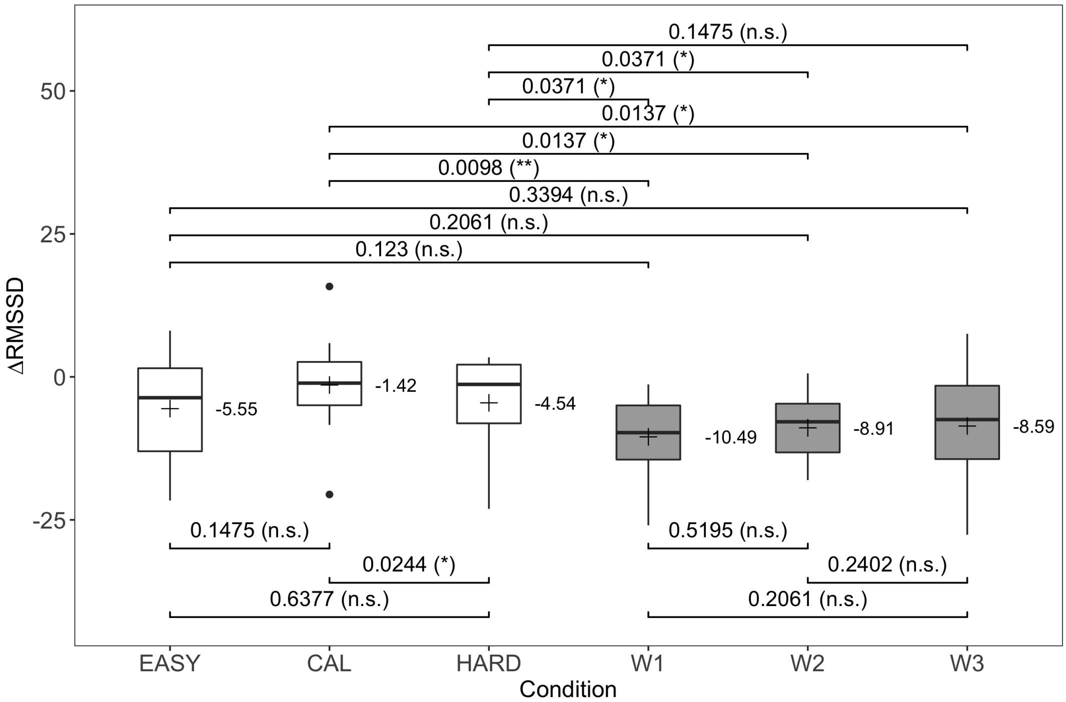

FIGURE 5.11: ΔRMSSD Results in Experiment 1. Crosshairs and Numbers Next to Them Represent Means.

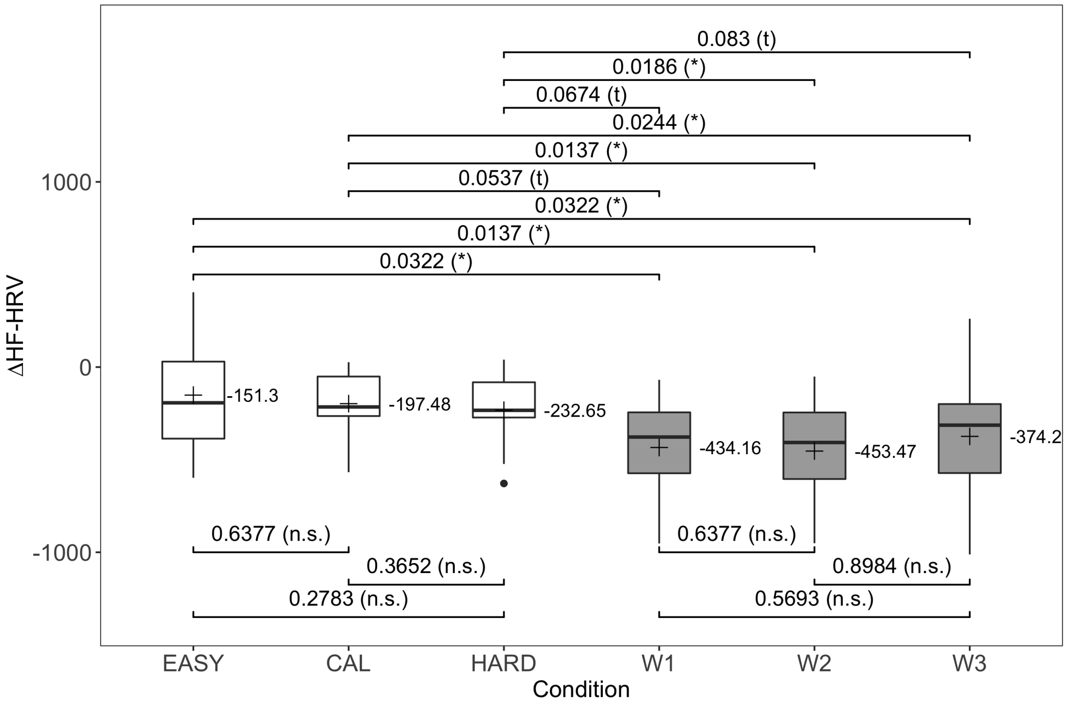

FIGURE 5.12: ΔHF-HRV Results in Experiment 1. Crosshairs and Numbers Next to Them Represent Means.

EEG Results

Variations in EEG frequency band powers were assessed to examine how frontal regions of the brain operate during flow, both within the traditional DM paradigm and across paradigms and tasks. Two analyses were conducted that assess the averaged power changes for each condition, and also the pattern of temporal variation over 30s windows within each condition.

For the between-condition analyses, Friedman tests were computed for each frequency band (over one Region of Interest (ROI) - the pooled frontal electrodes: AF3, F3, F7, FC5, FC6, F8, F4, AF4) to detect main effects across sampling points. In a preparatory step, the presence of hemispheric differences for the Alpha frequency ranges (due to possible effects in terms of FAA - see Labonté-Lemoyne et al., 2016; Harmon-Jones, Gable, and Peterson, 2010) was assessed for an Alpha range hemispheric difference score: Right Hemisphere (RH) - Left Hemisphere (LH). No main effect was found for these FAA features (see Table 5.6). The analysis of hemispheric differences was, therefore, not pursued further.

For the pooled frontal frequency band features, a main effect was found for the HiAlpha band. No effects were found for any other frequency range. Post-hoc pairwise Wilcoxon tests were conducted on the frontal HiAlpha feature (see Figure 5.13). Within the mental arithmetic task, the HiAlpha band shows significantly higher levels in the EASY condition than in the HARD condition, and the CAL condition indicated on a trend level. No difference was found between the CAL and HARD conditions. Within the writing task, no significant differences are found in the HiAlpha band. Across tasks, significantly lower HiAlpha is found in all writing rounds compared to the mental arithmetic EASY condition (trend level for writing round 2). Furthermore, trend level differences indicate lower HiAlpha in W1 and W3 than in the mental arithmetic CAL condition, and one trend level indication of lower HiAlpha in W1 compared to the mental arithmetic HARD condition.

ROI	Freq. Band	Test Statistic	P-Value	Effect Size
Frontal	ΔTheta	$\chi^2(5) = 5.0476$	0.4101	-
Frontal	ΔLoAlpha	$\chi^2(5) = 4.3810$	0.4960	-
Frontal	ΔHiAlpha	$\chi^2(5) = 12.1905$	0.0323	0.4063 (M)
Frontal	ΔLoBeta	$\chi^2(5) = 8.9524$	0.1110	-
Frontal	ΔMidBeta	$\chi^2(5) = 6.7619$	0.2390	-
Frontal	ΔHiBeta	$\chi^2(5) = 0.9524$	0.9663	-
FAA (RH-LH)	ΔLoAlpha	$\chi^2(5) = 3.8095$	0.5772	-
FAA (RH-LH)	ΔHiAlpha	$\chi^2(5) = 8.0952$	0.1511	-

Notes: Reported Effect Sizes are Kendall's W (Kendall and Smith, 1939).

TABLE 5.6: Friedman Tests for EEG Features in Experiment 1.

Together these results indicate primarily that, with increased task difficulty, frontal brain regions become active, and not generally de-activated as is suggested in the Transient Hypofrontality Theory (THT) (Dietrich, 2004). In contrast, very easy and monotone tasks show higher frontal Alpha power, expressed here in the higher Alpha frequency ranges. For flow, this indicates rather that at least some prefrontal areas are

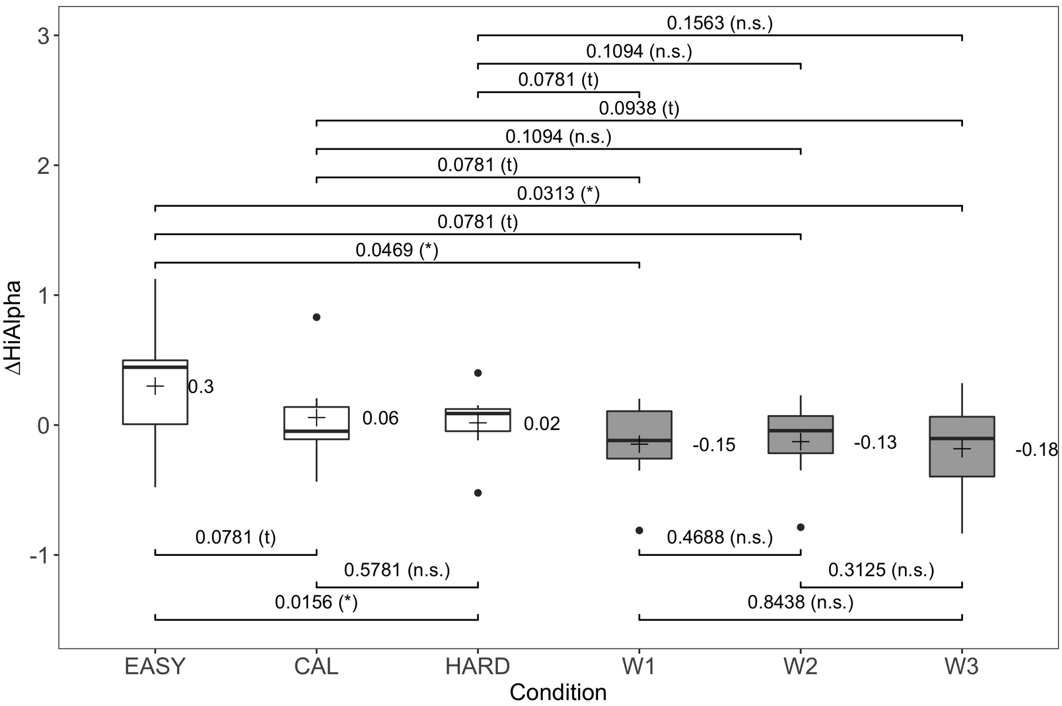

FIGURE 5.13: Frontally Pooled ΔHiAlpha Results in Experiment 1. Crosshairs and Numbers Next to Them Represent Means.

required to process the task. However, given a lack of specificity of the frontal Alpha pattern to changes in reported flow, this observation is instead considered to represent an engagement of top-down attention on the task at hand. Therefore, from these results across two tasks and paradigms, flow is considered to be related to increased attentional engagement (and not frontal downregulation). A surprising observation is the lack of frontal Theta power changes, as these have been repeatedly reported in related work (see Section 4.3). It is possible that this is caused by too coarse frequency band extraction methods (as related work uses more elaborate frontal Theta power isolation approaches - see Ewing, Fairclough, and Gilleade, 2016) or by a lack of frontal midline electrode positions in the used wearable EEG device (as frontal Theta changes are typically observed close to midline electrodes).

As a last series of analyses, an interesting proposition is followed up, that flow might be better identified through EEG measures by studying frequency power changes in time (see for example Soltész et al., 2014; Wolf et al., 2015). For this purpose, within-condition frequency power variations were investigated. Friedman tests on 30s-based segments of each round were computed (resulting in 10 and 14 segments for the mental arithmetic and the writing task respectively) (see Table 5.7). Results show main effects in the mental arithmetic EASY condition for the HiAlpha band, in the Writing Round 1 (W1) for the MidBeta band and the Writing Round 3 (W3) for the HiAlpha band. For the CAL and HARD mental arithmetic task conditions and Writing Round 2 (W2), no changes over time were observed.

For the follow-up assessment, a descriptive analysis was conducted. In a similar format, such analyses have previously been pursued for flow PNS measures (Harmat et al., 2011) for the sake of hypothesis development from small sample neurophysiological data. The descriptive analysis relies on rank transformed data (to further

Freq. Band	Test Statistic	P-Value	Effect Size
Mental Arithmetic - EASY Difficulty			
ΔTheta	$\chi^2(9) = 7.7091$	0.5637	-
ΔLoAlpha	$\chi^2(9) = 13.5636$	0.1387	-
ΔHiAlpha	$\chi^2(9) = 18.0727$	0.0343	0.3347 (M)
ΔLoBeta	$\chi^2(9) = 11.4182$	0.2481	-
ΔMidBeta	$\chi^2(9) = 8.0727$	0.5268	-
ΔHiBeta	$\chi^2(9) = 10.0000$	0.3505	-
Mental Arithmetic - CAL Difficulty			
ΔTheta	$\chi^2(9) = 7.8182$	0.5526	-
ΔLoAlpha	$\chi^2(9) = 10.9455$	0.2795	-
ΔHiAlpha	$\chi^2(9) = 5.6727$	0.7722	-
ΔLoBeta	$\chi^2(9) = 12.4727$	0.1880	-
ΔMidBeta	$\chi^2(9) = 9.9636$	0.3534	-
ΔHiBeta	$\chi^2(9) = 14.6182$	0.1020	-
Mental Arithmetic - HARD Difficulty			
ΔTheta	$\chi^2(9) = 1.5818$	0.9965	-
ΔLoAlpha	$\chi^2(9) = 10.3527$	0.3227	-
ΔHiAlpha	$\chi^2(9) = 6.7745$	0.6606	-
ΔLoBeta	$\chi^2(9) = 15.2400$	0.0846	-
ΔMidBeta	$\chi^2(9) = 8.7818$	0.4577	-
ΔHiBeta	$\chi^2(9) = 8.6509$	0.4701	-
Scientific Writing - Round W1			
ΔTheta	$\chi^2(13) = 9.3510$	0.7459	-
ΔLoAlpha	$\chi^2(13) = 8.7306$	0.7930	-
ΔHiAlpha	$\chi^2(13) = 9.4816$	0.7357	-
ΔLoBeta	$\chi^2(13) = 19.4735$	0.1091	-
ΔMidBeta	$\chi^2(13) = 32.6163$	0.0019	0.3584 (M)
ΔHiBeta	$\chi^2(13) = 22.1673$	0.0528	-
Scientific Writing - Round W2			
ΔTheta	$\chi^2(13) = 11.3905$	0.5781	-
ΔLoAlpha	$\chi^2(13) = 13.6571$	0.3984	-
ΔHiAlpha	$\chi^2(13) = 13.9238$	0.3792	-
ΔLoBeta	$\chi^2(13) = 15.1810$	0.2962	-
ΔMidBeta	$\chi^2(13) = 16.0000$	0.2491	-
ΔHiBeta	$\chi^2(13) = 12.1524$	0.5152	-
Scientific Writing - Round W3			
ΔTheta	$\chi^2(13) = 12.2286$	0.509	-
ΔLoAlpha	$\chi^2(13) = 14.3143$	0.3521	-
ΔHiAlpha	$\chi^2(13) = 24.8571$	0.0241	0.478 (M)
ΔLoBeta	$\chi^2(13) = 6.8286$	0.9108	-
ΔMidBeta	$\chi^2(13) = 12.2000$	0.5113	-
ΔHiBeta	$\chi^2(13) = 17.4286$	0.1804	-
Notes: Reported Effect Sizes are Kendall's W (Kendall and Smith, 1939).			

TABLE 5.7: Friedman Tests for EEG Within-Task Features in Experiment 1.

remove inter-individual differences - for a similar logic see Manzano et al., 2010) plotted against the progression in time. Furthermore, a Locally Estimated Scatterplot Smoothing (LOESS) smoothed fit line (Cleveland, 1979) is included in the plot as a conservative representation of a possible time-frequency power progression. The minimally possible span for the LOESS algorithm was selected to aid towards this goal of conservative representation. The progressions are shown in Figure 5.14. This progression indicates that HiAlpha peaks slightly in the first 1-3min. Toward the end of the phase (minutes 4-5), more volatility in the HiAlpha pattern is visible. For the writing task, some HiAlpha fluctuation is also indicated in round W3, in particular with early, mid and late-phase peaks (60s, 120s, 420s), and valleys in-between (150s, 330s). Furthermore, within the writing task, MidBeta progressions indicate a second and different pattern. Specifically, with the most pronounced differences for early vs late segments, pointing to a MidBeta activity increase in the first minutes of round W1. Overall, it should also be noted that no repeated start or end effects were found in all bands and experiment phases. Also, besides the MidBeta pattern in writing round W1, the phases showing higher flow reports are more strongly marked by consistency than volatility.

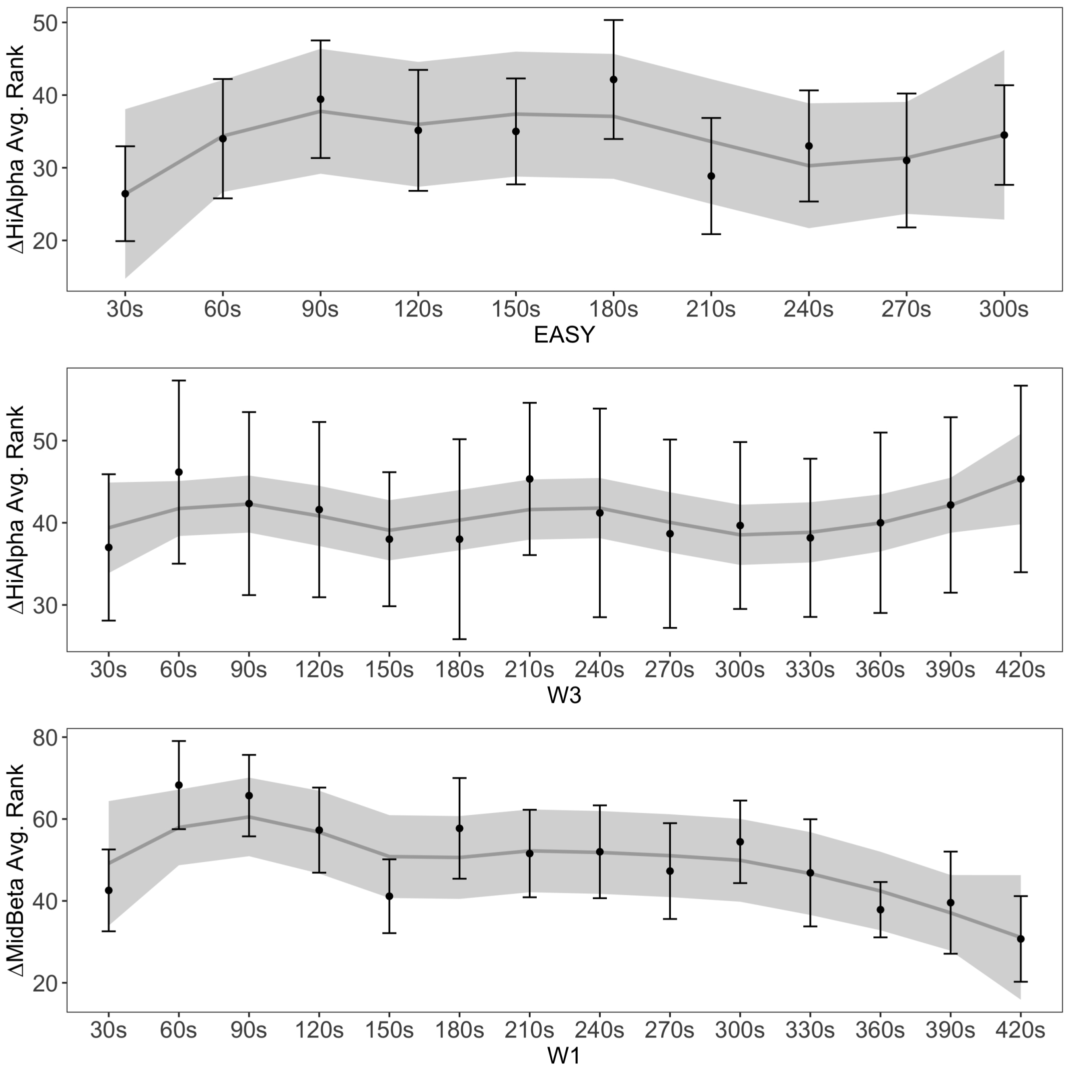

FIGURE 5.14: Frontal ΔHiAlpha and ΔMidBeta Within-Condition Changes in Experiment 1. Error Bars Are One Standard Error (SE).

Together these neurophysiological results from two tasks and paradigms indicate that flow is likely related to increased attentional engagement (as indicated by reduced and stable frontal HiAlpha power) and possibly to brief mental workload increases (as indicated by a short frontal MidBeta increase at the beginning of the writing task - the round that showed the highest reported flow levels amongst all conditions).

5.5 Discussion

5.5.1 Critical Findings

In these analyses, central gaps in the (neurophysiological) research on flow in more naturalistic settings were addressed that pertain to (1) the comparison of flow intensities across two experimental flow elicitation paradigms, of which one is newly developed for the Knowledge Work (KW) context, and to (2) the detection of flow experiences using wearable neurophysiological sensors that build on within-subject observation over two tasks and paradigms. Concerning the derived research questions, novel insights emerged that need to be critically appraised.

RQ3 - Mental Arithmetic DM Task Replication

Research Question 3 asked if the laboratory-based flow elicitation using a mental arithmetic Difficulty Manipulation (DM) task is reproducible. Within the mental arithmetic task, the results suggest a successful manipulation of difficulty (and subsequently flow) with comparable results to previous research. Specifically, the results show that flow experience (as indicated by self-reports) is experienced most strongly when task demands are optimally balanced with participants' skill levels. It needs to be appraised though, that this intensity of flow compared to low-difficulty task situations (the EASY condition), was not reaching statistically significant higher levels. This lack of separation between the EASY and CAL condition is potentially caused by high levels of perceived fluency during the very easy task. Similar complications have been reported in related work (Peifer et al., 2014; Peifer et al., 2015; Tozman, Zhang, and Vollmeyer, 2017), and some work has therefore also included participant action slowing-mechanisms during these easy task conditions (Keller et al., 2011; Harmat et al., 2015). Nevertheless, together with the results from the other variables below (perceived stress and neurophysiological data), it is considered that within the mental arithmetic task flow was most likely at its highest intensity in the CAL condition. This is why RQ3 is considered to be confirmed, as the elicitation of flow experience intensities was successfully reproduced, even in a different setting and using different self-report instruments than in (Ulrich et al., 2014).

RQ4 - cESM Utility

Research Question 4 asked if flow elicitation in the laboratory is intensified by a more naturalistic task setting (i.e. through the controlled Experience Sampling - cESM approach). First of all, within the writing task, all presented variables indicate a rather stable experience, despite repeated interruptions. This finding is important as interruptions are often considered a prime cause for a reduction or lack of flow experiences (Rissler et al., 2017). Therefore, initially, more experiential variance was anticipated due to repeated task interruption. It is possible that other factors in the writing task design (like the goal setting process) helped to mediate this interruption impact. Goal setting has been found to be an important step in the writing process that facilitates high-quality work (Flower and Hayes, 1981) and is theorised a prime conductor for flow experiences in the original theory (Csikszentmihalyi, 1975). The structured goal-setting approach that was included in the cESM task design was reported as being helpful to guide participants actions in the writing task. Also, participants did not report that substantial adjustment of these task goals was necessary as the task progressed. Therefore, it is herein considered that this structured form of goal setting for KW tasks might be a useful element for flow research with more naturalistic tasks to enable the elicitation of flow experiences. In comparison to the mental arithmetic DM task, flow in cESM writing was reported to be at least as high as in the CAL mental arithmetic condition, and with tentatively higher levels (in one instance bordering statistical significance - Writing Round 1). The results, therefore, show support that the cESM approach with a writing task can be used to elicit flow, at least at similar intensities that are elicited with a standard DM paradigm (the mental arithmetic task). Therefore, RQ4 is considered to be partially confirmed.

Furthermore, though, another clear difference is visible between the two tasks. Even though tasks are reported as similarly important, writing appeared to be experienced as less stressful and difficult and showed stronger Heart Rate Variability (HRV) reductions. A key reason for the stress difference could be the design-related, contrasted presence of multiple stress factors (difficulty overload, social-evaluative threat, lack of control) (Tozman, Zhang, and Vollmeyer, 2017). These factors have in the past purposefully been introduced to flow experiment designs to elicit motivated task performances (Tozman et al., 2015; Ulrich et al., 2014; Tozman, Zhang, and Vollmeyer, 2017) and have resulted in repeated sightings of increased stress or arousal in CAL and HARD conditions (compared to EASY conditions), even in contexts where threat experiences could be less likely (i.e. gaming tasks) (Harmat et al., 2015; Tozman et al., 2015; Klarkowski, 2016; Tozman, Zhang, and Vollmeyer, 2017). The results indicate that a task that is naturally important to the individual, yet lacks these stressors results in similarly reported flow intensities without perceptions of strain. The critique on the aptitude of the DM paradigm to elicit deep flow experience could, therefore, receive some support (Moller, Meier, and Wall, 2010), as could the proposition that naturalistic tasks are perceived as less attentionally effortful and therefore more in the realm of deeper flow experiences (Hommel, 2010). It should also be pointed out, that these results might indicate a central psychometrical limitation of flow self-report instruments, as there could be experiential components to flow (e.g. a perception of effortlessness) that are not captured by established scales. With a similar thought, some related work has started to include measures of effortless concentration (Harmat et al., 2015) and mental workload (Harris, Vine, and Wilson, 2017a).

RQ5 - Cross-Situation Flow Neurophysiology

Research Question 5 asked which neurophysiological correlates of flow can be observed across different cognitive task scenarios with wearable sensors. Two physiological feature spaces (HRV as PNS measures and frontal EEG frequency band powers as CNS measures) delivered interesting and consolidating insights.

Regarding the ECG features, within the mental arithmetic task, the observed HRV results are comparable with previous work showing increased PNS activation from CAL to HARD conditions (that would indicate a moderate level of physiological activation in flow) (Tian et al., 2017; Klarkowski, 2017; Tozman et al., 2015; Klarkowski, 2016). Within the writing task, the HRV similarity across the sampling points further supports the observation of a consistent experience. However, given this consistency, it is hard to tell if the reduced HRV (compared to the mental arithmetic task) is due to a qualitatively different flow experience, or due to other variables (e.g. task complexity or effort). Nevertheless, the finding that even though the writing is perceived as less stressful, the observation of further lowered HRV levels is interesting and could alternatively indicate that the proposition, that flow is a state of high physiological activation, is correct (Manzano et al., 2010; Keller et al., 2011; Ulrich, Keller, and Grön, 2016b). In any case, the comparison of the two tasks could explain these previously contradictory findings (i.e. why in some studies the physiological results point to moderate PNS activation and in some to high PNS activation) to be caused by task or paradigm based confounds. These results highlight a complication for future adaptive NeuroIS work that uses physiological thresholds to infer experiential states based on single-task calibration. Multi-task observation is likely going to be necessary to calibrate the detection of classification of flow intensities even on a within-subject level. What is interesting about the present results is also, that only indicators of parasympathetic cardiac modulation (RMSSD and HF-HRV) were found with main effects. Related work has also indicated an increasingly robust utility of these markers for the separation of optimal and non-optimal difficulties (see Section 4.2).

Regarding the EEG features, within the math task, observations of frontal regions integrate in multiple ways with previous work. First of all, it needs to be outlined that one of the more robust findings from related work (elevated frontal Theta levels with difficulty increases) is not supported by the findings in this work. The absence of the frontal Theta pattern could point to a need to further specify Theta band activity (like Ewing, Fairclough, and Gilleade, 2016 who select individualised Theta band activity in a 1 Hz range within the 4-7 Hz range, where the largest power modulation from the difficulty treatment is visible). On the other hand, it is possible that this frontal Theta effect is not visible as it typically emerges closer to midline electrodes, that are not available for the EEG headset that was used in this work.

Second, the finding of lower HiAlpha activity with increasing task difficulty is interesting in multiple ways. First, the separation of the Alpha band shows that HiAlpha is a more differentiating feature for the mental arithmetic task conditions. This finding has not been outlined as such in previous work, yet would explain why some of the work that includes separation does find frontal Alpha to contribute valuable diagnostic information between difficulty conditions (see, e.g. Ewing, Fairclough, and Gilleade, 2016; Katahira et al., 2018), while others that work with the broad Alpha band do not (see, e.g.Chanel et al., 2011; Klarkowski, 2017). Whether or not the HiAlpha band provides a diagnostic potential for flow observation beyond the indication of a difference to EASY conditions, remains however a subject of future work. Presently it appears that CAL and HARD conditions show a similar level of HiAlpha, that is lower than in the EASY condition (thus showing a potentially reduced activity in frontal brain regions in the EASY condition). The results of Alpha decreases with increased task difficulty are in line with previous EEG research on mental workload (Borghini et al., 2014) and are a possible indicator of top-down attentional task engagement. These results are also somewhat similar to a recent fNIRS-based study that finds frontal brain activity to be reduced in low difficulty conditions and to increase when task difficulty increases (Barros et al., 2018). These authors attribute this activity to attention on the task. The finding is plausibly transferrable given that the volatility of the HiAlpha power is only present in the mental arithmetic EASY condition. Besides, mind wandering during this condition was noted explicitly by one participant in the final experiment survey comment section. However, it needs to be appraised that frontal Alpha reduction is not a unanimous finding in the related work (see Section 4.3). For example, the results by Léger et al. (2014) and Labonté-Lemoyne et al. (2016) point in the opposite direction. These differences might be mainly caused by experimental approaches and analyses. Labonté-Lemoyne et al. (2016) observe two interacting participants and do not manipulate difficulty externally. Léger et al. (2014) derive their conclusion not through a DM paradigm but through regression analysis of reported flow and frontal Alpha power at the end of a prolonged learning session. The present data are, however, amongst the first to show this reduction of frontal Alpha across two tasks and paradigms, which represents a vital contribution to the literature.

Third, an absence of frontal Beta main effects was observed in the present results, yet a pattern was detected for the temporal dynamics in the writing task. In related work, lower beta activity is linked to higher flow experience self-reports (Léger et al., 2014), and has also been found to increase with task difficulty increases from EASY levels (Klarkowski, 2017). However, other studies have also found no beta difference at all on frontal sites (Soltész et al., 2014; Katahira et al., 2018). As Beta frequencies have generally been found to be related to changes in mental workload (Michels et al., 2010), the absence of Beta changes with task difficulties and tasks in this work is rather unexpected. It could be expected that at least some elevation in Beta levels would occur with flow, given that it unlikely occurs in very low difficulty tasks. The absence of such findings can perhaps be explained by methodological limitations, specifically the somewhat large area of pooled electrodes (as related work shows that workload-related Beta changes emerge less over medial central regions - see Michels et al., 2010), or the criteria for the sub-segmentation for the Beta bands that might have been insufficiently sensitive in capturing these expected patterns. Future work should apply a more refined perspective to further elaborate on potential Beta frequency relationships with flow. However, a possible relationship is identified in the temporal analysis, specifically for the frontal MidBeta range. Within the writing task, EEG results mostly support the view of a consistent experience across writing trials. The only effect that shows variation is the initial Beta increase within the first part of writing round W1 (temporal analysis). On the one hand, this round showed tentatively elevated flow reports when compared to the mental arithmetic CAL condition and the other writing task rounds. This elevation could mean, that at least some early elevation of frontal Beta could be related to intensified flow experiences. Given that this variation is not apparent in later phases, it is most likely attributable to a type of task initiation activity. It could perhaps be interpreted as an indicator of taking on a challenging task. It has been reported in flow and writing research (Flower and Hayes, 1981; Csikszentmihalyi, 1996) that initiation of a writing session takes additional effort to structure the task that may be required less at later stages. Given that Beta activity is in neuroscience literature often related to increased mental workload (Michels et al., 2010), this finding could signify an initial increase in cognitive effort and mental workload that dissipates after a while. Whether or not this observation is related to flow experience intensities will have to be studied further, but provides a novel proposition on how to possibly observe flow experiences through a time-dynamics perspective.

As an answer to RQ5, these neurophysiological results from two tasks and paradigms indicate that flow is related to increased physiological activation (moderate or low HRV as indicated by parasympathetic cardiac modulation features), to increased attentional engagement (and not frontal downregulation - as indicated by reduced and stable frontal HiAlpha power) and possibly to brief mental workload increases (as indicated by a short frontal MidBeta increase at the beginning of the writing task - the round that showed the highest reported flow levels amongst all conditions).

5.5.2 Limitations & Future Directions

Several limitations to the presented experiment need to be appraised. In general, the focus on a sample of German students and the small sample size is a critical limitation of this study, which is why the results can only be treated as preliminary. Through the integration with related work, this limitation was attempted to be overcome to some degree. Nevertheless, future work will have to repeat and extend these observations with larger samples. Due to this small sample, only indirect comparisons (non-parametric ANOVAs) were pursued between self-reports (flow and stress) and neurophysiological data. While this is common practice in the related work and was selected especially due to this relatability, future work should explore the use of more direct modelling of relationships. To that regard, especially rank-based Repeated Measures (RM) models (see Manzano et al., 2010) and non-linear regression analyses (e.g. Bian et al., 2016; Tozman et al., 2015; Chin and Kales, 2019) have helped to identify relationships between HRV metrics and reported flow. The application of these methods to the EEG feature space could provide similarly interesting additions to the flow neurophysiology literature and for the development of adaptive NeuroIS.

In more specific limitations, it was found that while the mental arithmetic DM task was replicable, it was also seen that flow contrasts could be improved. On the one hand, this might be achieved by further inducing boredom through waiting periods to prevent participants from rushing through very easy tasks and thus experiencing highly fluent actions. Similar recommendations have been made in related work (Harmat et al., 2015; Keller, 2016). On the other hand, flow intensities in optimal difficulty conditions could be further intensified by improving the calibration mechanism. In this instance, it is possible that difficulties in the CAL condition were slightly too high. Therefore, slower level increases in the calibration stage might help to set better difficulties. However, another direction seems particularly interesting, namely the integration of a more self-determined optimal difficulty calibration. Such an approach has been used in Barros et al. (2018) and was found to increase flow intensities compared to automatically calibrated difficulties. This approach is also interesting, as there was a tentative increase in flow experiences seen in the writing task. It was a primary argument that the writing task provides a more naturalistic task experience through the provision of more freedom and autonomy to work on a task that fits one's preferences and abilities. Therefore, the inclusion of a self-selected (optimal) difficulty could provide additional comparability between the two paradigms (DM and cESM) through added autonomy in the DM task. Furthermore, for the cESM approach, the consistency of elicited flow (and other) experiences was considered as a benefit. However, it can also be seen as a limitation in terms of low experiential variance that might be needed to identify neurophysiological contrasts. Therefore, future work should increase variance in the paradigm, for example by including a boredom phase (for example a text copying stage with slowed keyboard input), or by temporally varied sampling (with short and long task intervals, more or less flow might occur). Such an approach could also provide highly interesting insights into the time-dynamics of flow, one area that was considered to provide promising new grounds for the study of flow in this work and related work (Soltész et al., 2014; Wolf et al., 2015). In such future analyses, more sophisticated methods for feature extraction and time series modelling should be applied. For example, EEG band power extraction through Morlet Wavelet Transformation (MWT) is known to provide high temporal resolution up to the original sampling frequency (Cohen, 2014). It could, therefore, be used to derive much more accurate time series patterns than is possible with the herein used Fast Fourier Transformation (FFT) power extraction.

Lastly, the present work is limited in particular to frontal brain sites, yet with a fairly high aggregation. Future work should explore other topographical regions of interest that could be providing valuable information on what differentiates flow from other experiences. Some research points to the explicit role of temporal (Wolf et al., 2015), or parietal and occipital brain regions (Chanel et al., 2011). Also, more fundamental neuroscientific studies point to the relationship of higher frequency powers (Beta and Gamma) to mental workload over widespread regions of the scalp (Michels et al., 2010). As flow is related to task difficulties, such an observation could provide valuable information on flow boundary conditions. Lastly, since recent hemodynamic imaging work has found more nuanced frontal activation patterns (differences in activations of medial and dorsolateral Prefrontal Cortex - PFC areas during flow), a refined analysis of frontal regions could help to learn if the herein found frontal Alpha patterns are more topographically isolated.

5.6 Conclusion

In sum, an extensive analysis of the self-report and neurophysiological data was conducted in two flow elicitation paradigms. The results from this experiment, therefore, contribute in theoretical and practical ways to a foundation for the theory-driven development of adaptive NeuroIS. First, evidence is provided for the applicability and utility of the controlled Experience Sampling (cESM) approach to study flow in more naturalistic tasks in the context of Knowledge Work (KW). The writing task design was found to elicit a consistent flow experience that is at least as high in intensity as in an established Difficulty Manipulation (DM) paradigm. At the same time, this cESM approach elicited lower perceptions of stress, which makes the approach an interesting alternative for flow research. Specifically, the question arises, if it might be possible to study a difference between the experience of flow as a state of effortless (cESM) or effortful (DM) attention (Hommel, 2010) through comparison of these paradigms. Second, consolidating evidence is provided for neurophysiological configurations (of heart and brain) during flow experiences. Importantly, this evidence emerges across two experimental paradigms, which is a central contribution of this work. These results indicate that flow is likely related to increased physiological activation (moderate or low HRV as indicated by parasympathetic cardiac modulation features), to increased attentional engagement (and not frontal downregulation - as indicated by reduced and stable frontal HiAlpha power) and possibly to brief mental workload increases (as indicated by a short frontal MidBeta increase at the beginning of the writing task that showed the highest reported flow levels).

These results converge with recent understandings (Harris, Vine, and Wilson, 2017b) that the Transient Hypofrontality Theory (THT) is likely too simplistic since no general frontal downregulation is indicated. The results also provide more consolidation for related work through the application of frequency band segmentation. Previous results (of absent frontal Alpha relations to flow) may have been caused by the use of broad Alpha bands (e.g. 8-12 Hz). Such broad brands have been reported to possibly mask effects in narrower frequency ranges, especially such related to task-related and general attentional engagement (Klimesch, 1999). In the present work, the Alpha band separation provided useful diagnostic detail, isolating Alpha changes to the HiAlpha band. For the Beta band, however, this seemed less to be the case, although a few results point to potentially higher diagnostic properties of the MidBeta band. In particular, a direction that has been previously highlighted as promising for the identification of neural flow configurations - the study of time-dynamics of flow - received further support by this frequency separation approach.

For the development of adaptive NeuroIS, the neurophysiological results point to the opportunity to instantiate systems able to differentiate situations of non-optimal and optimal task difficulty by including HRV features (specifically RMSSD and HF-HRV) and frontal EEG features (specifically HiAlpha power). However, such a system would currently most likely only be used to identify situations of low task difficulty and mind wandering, as the separation of moderate from high task difficulties needs further insights. As related work that has combined DM paradigms with increased autonomy has found minima in HRV during these more autonomous conditions (Barros et al., 2018), it could be possible that HRV reductions beyond what is expected in higher task difficulties together with stable, frontal HiAlpha activity could be used to detect flow experience likeliness or at least its corollary of increased task attention. The convergence of these two features might thereby be explained by shared regulatory mechanisms (Peifer, 2012; Barros et al., 2018). Whether such diagnosticity is achievable and useful will be an interesting direction for future work. So too is the search for neurophysiological differences that could explain the stress perception difference and with it, the potential difference of flow experience as a state of effortless attention (Hommel, 2010).

Chapter 6

Experiment 2 - Difficulty, Autonomy, & Social Context

Contents of this section are in part adopted or taken from Knierim, Nadj, and Weinhardt (2019) and Knierim et al. (2019). See Section A.1 for further details.

6.1 Exploring Flow with Autonomy & Social Interaction

To provide the groundwork for the development of flow-facilitating, adaptive Neuro-Information Systems (NeuroIS), the work in this dissertation focuses on studying flow experiences in primarily cognitive tasks, to identify how flow could be intensified in experimental research (RG2). In this approach, flow neurophysiology research converges on more Knowledge Work (KW) related scenarios and refines how neurophysiological processes related to flow can be described across different situations using wearable sensor systems (RG3).

Experiment 1 confirmed that varying flow intensities could be elicited in laboratory research by using the mental arithmetic paradigm. Furthermore, it showed that alternative paradigms could elicit similarly strong, potentially stronger, and potentially qualitatively different flow experiences (high flow without stress perceptions). These differences were mainly considered to be driven by increased task naturalism and autonomy. In addition, physiological results showed that flow could be accompanied by high physiological activation (reduced HRV) and by reduced frontal Alpha power (EEG - considered to reflect increased attentional engagement). However, the EEG results also showed the absence of expected effects in the form of elevated mental workload levels (expected frontal Theta and Beta increases). Following this dissertation's research goals and building on the results from Experiment 1, Experiment 2 was set up, to further investigate some of the identified patterns, especially the patterns regarding autonomy and stress. Therefore, in Experiment 2, findings from Experiment 1 were integrated into alternative research designs, in particular in the form of providing increased autonomy by allowing participants to self-select an optimally challenging difficulty level for the mental arithmetic task. Also, Experiment 2 pursued an additional direction to intensify flow in the laboratory, namely completing a task during social interaction in a small group.

Two reasons drove the decision to investigate flow in small group interaction. On the one hand, related work has recently found repeatedly, that flow is more intense when experienced in a group (Magyaródi and Oláh, 2017; Tse et al., 2016; Walker, 2010). On the other hand, there is a high practical relevance that is currently placed on small group work. Groups are increasingly used to solve problems that surpass the capabilities of individuals, for example in academia where the dominant role of authoring groups is becoming apparent for high-quality research (Wuchty, Jones, and Uzzi, 2007). Also, in the industry, where Artificial Intelligence (AI) is leading to an increased demand for human knowledge problem solving, demand for group cooperation is increasing (Keith et al., 2016; Frey and Osborne, 2017). However, the Aristotelian promise ("the whole is more than the sum of its parts") is not automatically realised, and groups are often found to perform below their potential (Kerr and Tindale, 2004). For this reason, an emphasis has been put on the support of performance conducive states in small group research and thus on flow in small groups in particular (Keith et al., 2016).

The study of flow in social interaction is relatively young (Hout, Davis, and Weggeman, 2018). The flow experiences of small group members have been found to impact not only individual-level, but also group-level performance, satisfaction and growth outcomes (Heyne, Pavlas, and Salas, 2011; Walker, 2010). Furthermore, a central observation has so far been, that flow in social interaction might even be more intense than when experiencing flow alone (Walker, 2010). Nonetheless, while considerable correlational research has been conducted, the controlled study of flow in small groups has only sparsely attracted scholar's attention. Therefore, by extending established work on flow elicitation in the laboratory to the small group level (i.e. by adapting the mental arithmetic task), the intensification of flow in the laboratory can be further explored. Doing so, important contributions are made to understand flow experiences in a more encompassing nature. Importantly, a particular gap in flow in small groups can be bridged, that is the paucity of research addressing digitally-mediated interactions. As social interaction processes deviate strongly between face-to-face and digitally-mediated settings (Derks, Fischer, and Bos, 2008; Chanel and Mühl, 2015), the extension of previous work on flow in groups to the digital context represents an important research gap. Also, while there is an increasing interest to elucidate the underlying physiological processes of the flow experience, there has so far been almost no related research in small group settings. If scholars and practitioners want to better understand and facilitate flow in groups, laboratory experiments using physiological measures must be conducted. By analysing experiences across solitary and small group interactions during varied difficulties, the aim of Experiment 2 is to answer the main research questions of how well manipulations of autonomy and social context (= Autonomy Manipulation - AM and Social Context Manipulation - SCM) can serve to elicit more intense flow in the laboratory. More specifically, the research questions in Experiment 2 are:

- **RQ6:** Is flow elicitation in the laboratory intensified by:
  - **RQ6.a:** increased task autonomy?
  - **RQ6.b:** performing tasks in groups?
- **RQ7:** Is the flow elicitation using of a mental arithmetic Difficulty Manipulation (DM) task extensible to social interaction settings?
- **RQ8:** Which correlates to flow can be identified amongst different knowledge work scenarios using wearable sensors in:
  - **RQ8.a:** the PNS – in particular parasympathetic HRV indicators?
  - **RQ8.b:** the CNS – in particular mental workload and attentional engagement EEG power indicators?

To summarise, in Experiment 2, the experience of flow is observed in isolation and in digitally-mediated small group work on a cognitive task (SCM), with additional treatments that offer increased autonomy (AM). Altogether this second experiment contributes to the literature on flow experience, and the development of adaptive NeuroIS by (1) being the first to extend and compare a previously validated experimental task for solitary flow to the group level, and (2) extending the sparse knowledge on flow in digitally-mediated social interaction, and identifying two novel theoretical propositions as to why some flow experiences might be experienced as more intense than others. Furthermore, the refined analyses of neurophysiological data (specifically, parasympathetic HRV indicators, and EEG power indicators of mental workload, cortical idling, and attentional engagement) further consolidate the empirical knowledge on how flow can be described through changes in the heart and the brain. Importantly, these results are strengthened through the inclusion of various mechanisms for the elicitation of flow experiences in the laboratory (DM, AM, and SCM), which represents the major contribution of this work to the flow neurophysiology literature.

6.2 Experiment Design & Preparation

6.2.1 Materials

As in the first experiment, the second experiment was built on the pre-validated mental arithmetic DM paradigm (see Section 5.2). Replicating the design by Ulrich et al. (2014), participants sum multiple numbers, depending on the active and dynamically adjusted difficulty level. Adaptations and extensions were included to investigate additional hypotheses on the individual level and to enable comparability with small groups completing the task. To reduce the impact of potential confounding factors in small groups and to keep consistency with related work, constant groups sizes of three anonymous members were chosen.

Optimal Difficulty Mechanisms (AM)

First of all, given the results from Experiment 1 and emerging research, an additional mechanism for difficulty calibration was included, namely the self-selection of optimal difficulty (denoted AUTO). This alteration was used to test if increased autonomy would lead to intensified flow experiences - as this was found to be a possible conclusion from the cESM writing task in Experiment 1. Support for this proposition was also found in recent research that found more intense flow experiences in the laboratory when comparing self-selected and objectively calibrated optimal difficulties (Barros et al., 2018). This approach also factors in the discussion of the adequacy of objective mechanisms for optimal difficulty calibration. Specifically, it has been highlighted that such optimality might be a subject-dependent state with some individuals preferring underload and others preferring overload (Fong, Zaleski, and Leach, 2015; Tse et al., 2016). Thus, to achieve optimal difficulty balance, two different approaches were included. In contrast to the Self-Selected Optimal Difficulty (AUTO) condition, the established optimal difficulty calibration approach was used (Ulrich et al., 2014). In this calibrated difficulty condition (here denoted: CAL), the optimal level was computed from the mean level of the last 25% of trials in a calibration phase. For the optimal difficulty self-selection, at the start of the AUTO condition, participants were shown examples of how the levels would look in different difficulties, and were asked to select the level that would optimally challenge them. A screenshot of this mechanism is shown in Figure 6.1.

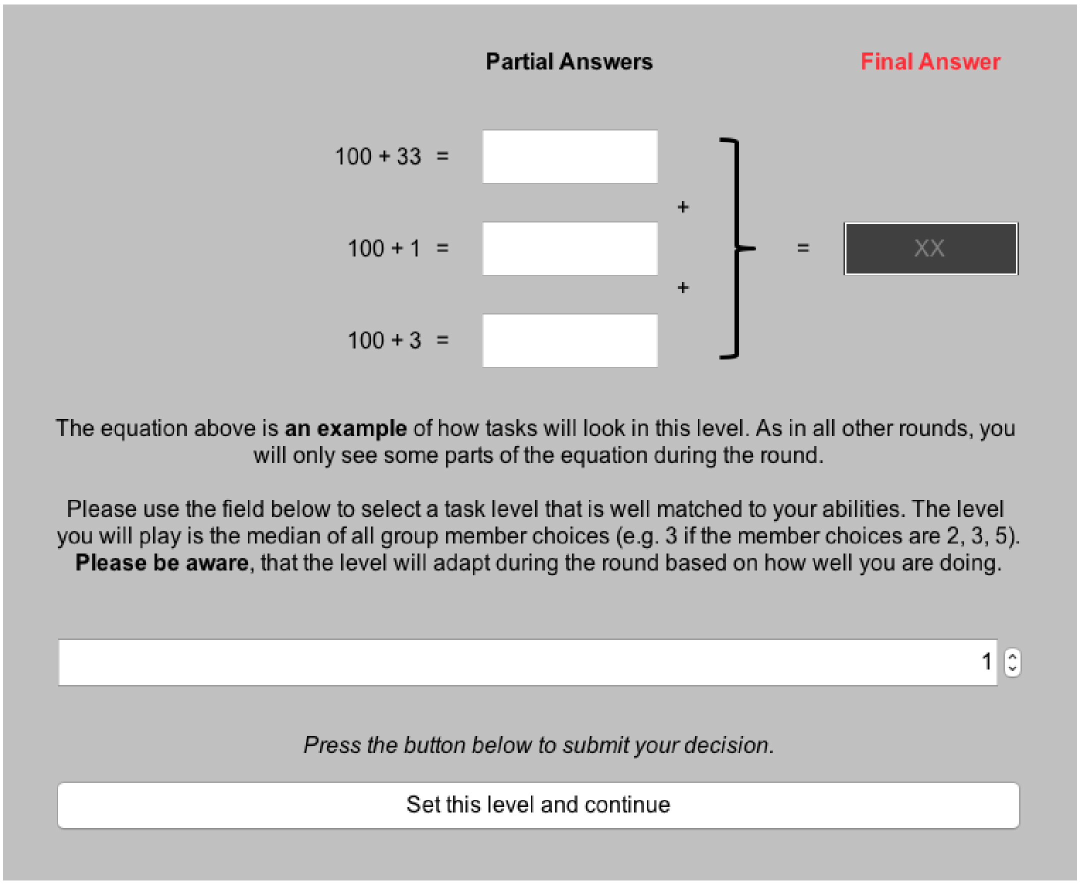

FIGURE 6.1: Selecting the Optimal Mental Arithmetic Difficulty Before the AUTO Condition in Experiment 2. The Displayed Equation was Updated When the Selected Level Was Changed.

Mental Arithmetic DM Task Adaptation

Secondly, given the emerging research on intensified flow experiences in groups, it was decided to alter the mental arithmetic task in ways to allow for its completion by multiple members of a small group (here: 3 people). As a starting point, it was decided to increase the number of equations to be solved from one to three. Still, a final solution had to be provided in the form of a three-digit number, yet partial answers could now be entered. This separation was used to allow participants working the task in a small group (Multi Person Condition (MP) condition) to focus on sub-tasks, a process typical for small group work. This change meant, that always, at minimum, six summands were presented (instead of a minimum of two), which increased the task difficulty for participants working in isolation (Single Person Condition (SP) condition) considerably. For this reason, the trial durations were extended to 28 seconds (from the previous 18 seconds). To further re-balance the difficulty in the SP condition, the digit ranges and the levelling logic were adjusted. In the EASY condition, only the results 303 or 304 were possible. This solution represented the minimal possible difficulty, that would require participants to keep their attention on the task. For the other conditions (CAL, AUTO, and HARD), difficulties were again adapted dynamically, and the mechanism of adding/removing a digit with level increases/decreases was also retained. However, the ranges of digits and the number of summands varied between the two latter conditions. Equations were changed to be of a similar form as in the EASY condition, always starting with the number 100 to be added to the remainder of each partial equation. This format allowed to further lower the difficulty in the CAL and AUTO conditions by decreasing the range of possible digits to [1-3] while still generating a resulting three-digit number. In the HARD condition, this digit range was increased to [1-6] to increase the likelihood of additional mental carries, that were found to moderate difficulty changes together with the overall number of digits per equation. As in the original design by Ulrich et al. (2014) and in Experiment 1, the difficulty levels could vary freely in the optimal difficulty conditions (both CAL and AUTO). However, the levelling logic was further altered to slow the increase in difficulty, again to counterbalance the already increased difficulty through the added number of summands, especially for the SP condition. Specifically, difficulties were increased when 3/3 consecutive answers were correct and decreased when 2/3 answers were incorrect. In the HARD condition, a balance had to be achieved between creating trials that are hard to solve for small groups (especially in MP), but that also do not cause prolonged frustration (especially in SP) (Keller, 2016). For this reason, the initial level in the HARD condition was set to be twelve levels higher than calibrated the starting level for the CAL condition. Yet, the difficulty could still fall three levels from this point on. Furthermore, to ensure difficulty was high enough but could fall quickly enough if necessary, levels increased when 2/3 consecutive answers were correct and decreased when 3/3 answers were incorrect. Additional adjustments were made to the task design to provide better contrasts in difficulty and thus, ideally, flow. A waiting time for each trial was introduced (i.e. participants had to wait for 7s before entering their final result). Similar approaches have been used by other scholars to prevent participants from rushing through very easy tasks (Keller, 2016). Lastly, the introduction phase was prolonged to six minutes to induce boredom during the EASY condition further, and to account for increased learning time in the small group condition.

Small Group Specific Task Extensions (SCM)

Furthermore, a set of extensions were specifically created for the MP condition. To integrate the requirements for cooperation and integration (see Section 3.3), two adjustments were made to the original task format. To create integration (in particular interdependence between group members), one part of the equation was hidden for each group member. This partial privatisation of information meant group members were never able to solve a trial on their own, as they only saw two of three equations (random allocation at each trial – no unsolvable trials). Similar approaches have been used in so-called Hidden Profile Paradigms in other small group research (Schulz-Hardt et al., 2006). To enable cooperation (in particular coordination), in the MP version, by clicking on a partial answer field, participants could signal to other participants on which part of the equation they were currently working. All participants shared the partial answer fields, that means they were able to correct the partial answers of each other. Each participant, however, was able to provide their final answer, which could not be altered by the other group members. This step was found to be a necessary to provide a simple feedback mechanism about the group's performance. It was also important, as participants were informed that a trial could only be solved correctly if at least two out of three group members provided the correct answer (to keep all group members engaged). Comparably, in the SP treatment, internal feedback about how well one is doing is already available. Through this sharing of final answers, at least a similar monitoring process was enabled. It was determined in pre-tests, that without such monitoring, participants found it impossible to assess how well their group was doing. Lastly, for the determination of the self-selected optimal difficulty for the group, the median of the selected individual optimal difficulties was chosen, and participants were informed (with an example) about this process while making their decision (see Figure 6.1). The median was chosen over the arithmetic mean as it was least likely to be manipulated by other group members. A screenshot of the math task during the MP condition is shown in Figure 6.2, and simplified examples of the task difficulties are shown in Table 6.1.

Condition	Level	Ulrich et al. 2014	Experiment 1	Experiment 2
EASY	0	103 + 6	101 + 2	100 + 1 (+) 100 + 1 (+) 100 + 2
CAL	1	65 + 73		100 + 22 (+) 100 + 3 (+) 100 + 1
CAL	2	58 + 91 + **4**		100 + 13 (+) 100 + **22** (+) 100 + 3
HARD	15	72 + 12 + 32 + 67 + 29 + 58 + 63 + 14 + 45		100 + 14 + 15 + 62 (+) 100 + 55 + 65 + 23 (+) 100 + 23 + 53 + 11
HARD	16	19 + 46 + 55 + 26 + 73 + 49 + 57 + 10 + 34 + **5**		100+35+22+16 + **2** (+) 100 + 64 + 45 + 26 (+) 100 + 25 + 51 + 31

Notes: Additional Digits with Level Increase are Highlighted in Bold.

TABLE 6.1: Mental Arithmetic Task Difficulties in Experiment 2.

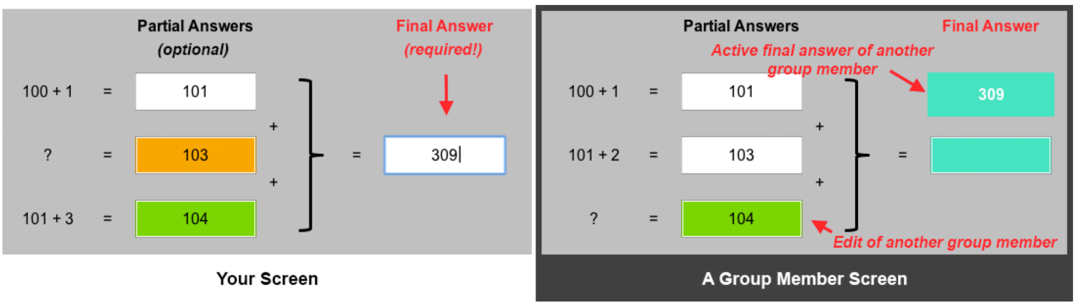

FIGURE 6.2: Mental Arithmetic Task in Experiment 2. Excerpt From the EASY MP Condition. This Picture was Shown to Participants in the MP Condition to Explain the Task.

6.2.2 Procedure & Sampling

Experiment 2 was conducted in a laboratory setting with sound-proofed and air-conditioned booths. Participants completed either the mental arithmetic DM task (difficulty as a within-subject factor) in isolation SP or as part of a three-person group MP (social context as a between-subject factor). Four participants were invited per session and allocated to SP or MP treatments randomly to increase anonymity and to contain potential confounding effects of group member familiarity. The order of the four mental arithmetic task conditions was randomised. To keep the number of condition order variations low, a uniform and balanced Latin Square Williams Design was chosen for the randomisation (four different orders) (Williams, 1949). At the start of the experiment, participants were welcomed, informed about the upcoming procedure and measurements. Afterwards, participants were asked to sign a consent form for their participation. Next, participants were guided to a booth in the laboratory and were fitted with the physiological sensors on the head (EEG) and chest (ECG), and the signal quality adequacy was checked. Following, participants completed a first survey collecting demographic information and some initial state variables. The setup in the laboratory booth is almost identical as for Experiment 1 - which was shown before in Figure 5.5. To complete this preparation stage, participants then completed eyes-open and eyes-closed baseline phases in which they were asked to "let their mind wander to wherever it takes them", and to avoid unnecessary movements. In the eyes-open phase, the participants were further asked to keep their eyes focused on a white fixation cross on a grey screen. The same message and fixation cross were shown for the washout screens before each task condition. Afterwards, an introduction to the math task was shown, and participants could familiarise themselves with the task in a practice round using the EASY condition design. Afterwards, the task was shown in the CAL condition design, starting at level 1 to calibrate the optimal difficulty. Next, all four math task conditions were presented in randomised order for five minutes. Participants responded to surveys after each condition (round surveys) and at the end of the experiment. After the last survey was completed, sensors were removed, and participants were debriefed. The complete experiment procedure is outlined in Figure 6.3.

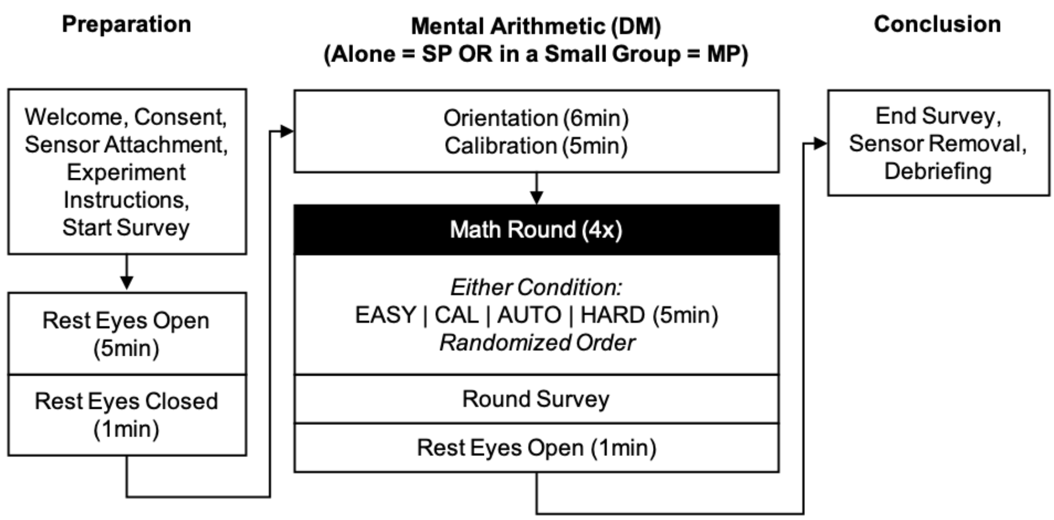

FIGURE 6.3: Procedure of Experiment 2 for Each Participant (Mixed Between/Within-Subject Design).

Data were collected for 41 participants in the SP treatment and 120 participants (i.e., 3 participants per group x 40 groups) in the MP treatment. Additional information on the sample is summarised in Table 6.2. Participants were sampled from a public student pool at the KIT and received a compensation of 21 Euros for the SP and

23 Euros for the MP condition. The compensation was based on the duration, and the MP version of the experiment lasted longer on average as participants had to wait on the slowest group member in the survey stages. Participants were screened for being generally healthy, not taking any mind-altering medication, having full eye-sight (with or without correction), and abstinence of the consumption of alcohol, marijuana or other recreational drugs in the past 24 hours before the experiment. Furthermore, participants were asked to arrive at the laboratory with washed hair and not using hair gel, hairspray, or similar products. SP and MP samples were assessed for comparability in terms of task preferences (3 items by Ulrich et al., 2014) and flow proneness (21 items by Ullén et al., 2012). Two-sided Welch's t-Test comparisons revealed no significant sample differences for flow proneness (p = .6278) and math task preferences (p = .3929). These results suggest an absence of a general bias for the likeliness of flow experience among the two participant samples.

Variable	Counts or Distributions	
	SP	MP
N (Indiv. / Groups)	41 / –	120 / 40
Age (Mean / Median)	24.15 / 23	23.34 / 22
Gender (Female / Male)	17 / 24	54 / 66
Handedness (R / L / Ambi)	39 / 1 / 1	108 / 11 / 1
English Abilities $\geq$ B1	97.44%	100%

TABLE 6.2: Sample Description for Experiment 2.

6.2.3 Measures

Demographic information about age, gender, handedness, study majors, and language abilities were collected at the start of the experiment. Afterwards, self-reported perceptions of experiences were collected at two levels: (1) after mental arithmetic task conditions (herein termed "rounds"), and (2) at the end of the experiment (herein termed "end"). Round questionnaires contained scales for individual difficulty (and group difficulty in MP) (one item by Engeser and Rheinberg, 2008, and one item by Ulrich et al., 2014), flow (ten-item FKS scale by Engeser and Rheinberg, 2008, individual performance (and group performance in MP) (one item adapted from the NASA TLX by Hart and Staveland, 1988, and affective experience (two single-item SAMs for affective valence and arousal by Bradley and Lang, 1994, amongst others. In the MP condition, additionally, perceptions of information sharing (two items by Aubé, Brunelle, and Rousseau, 2014) were collected after each condition. End surveys included scales to explore some of the relationships of flow experience with group experiences. These group experiences span multi-item measures for individual satisfaction and growth opportunities in the social unit (Wageman, Hackman, and Lehman, 2005), and perceptions of collective efficacy (Zumeta et al., 2016). These constructs are collected once at the end of the experiment. Almost all questions used 7-point scales (SAM arousal and valence used 9-point scales) that indicate the level of agreement with a presented statement. As additional control variables, trait variables named to influence flow experience like flow proneness (Ullén et al., 2012) were also collected. An overview of all measured variables, including the item operationalisations is summarised in Appendix A.3.

Neurophysiological data were collected using wearable sensors. The selected sensors were chosen, balancing the trade-off of acceptable signal quality and wearability. The same two sensor systems, as in Experiment 1, were used (see Subsection 5.3.3).

In terms of ECG features, this experiment focuses on further elucidating flow-HRV patterns, in particular with regards to the role of the parasympathetic activity during flow (see Section 4.2). A central challenge in this approach is that almost all of the related research has been conducted with DM paradigms. This focus brings with it a critical confound that parasympathetic activity patterns might relate more to changes in difficulty, than to changes in flow. In the first experiment in this dissertation, changes in HRV (specifically parasympathetic activity indicators RMSSD and HF-HRV) highlighted that flow experiences could be accompanied by both moderately and strongly reduced HRV levels. These levels are either determined by the task (academic writing might require more physiological effort than mental arithmetic), or possibly by the increased autonomy that individuals had in the writing situation. This autonomy might have caused the elicitation of a more intense flow experience that is insufficiently captured by the flow self-reports alone. The latter is based on the observation that stress perceptions were also much lower in the writing task, despite similarly high flow levels and despite a substantial increase in physiological activation in this writing scenario. To follow up on these central HRV lines of research, in this second experiment, the patterns of parasympathetic activation related HRV features (RMSSD and HF-HRV) are analysed in two additional manipulations (AM and SCM).

Given the high dimensionality of the EEG data, it was again decided to focus on select Regions of Interest (ROI) and frequency band ranges to study flow-related neural activity patterns which have been proposed in related work (see Chapter 4.3). These variables of interest focus again on primarily frontal features, namely frontal midline Theta power (a known correlate of mental workload - Borghini et al., 2014; Ewing, Fairclough, and Gilleade, 2016) and frontal Alpha power (for the assessment of frontal downregulation as proposed in the Transient Hypofrontality Theory (THT)). However, as in Experiment 1 frontal Theta and Beta observations did not differentiate between difficulty conditions (possibly due to the electrode distributions of the EEG headset), Beta power over the whole scalp was additionally analysed given its well-documented relationship to changes in mental workload (Michels et al., 2010).

6.2.4 Data Processing

For all types of data, the data processing followed a five-step strategy. First, problematic data instances were identified and removed (e.g. failed control questions, implausible or erroneous physiological values in resting phases). Second, outliers were removed from the data using a conservative distance metric of $\geq$2 SD from the variable mean. Third, variable distributions were checked for normality using skewness and kurtosis inspections with cut-offs of $\leq \pm 1.5$. Fourth, main effect analyses were conducted using Mixed ANOVAs or Linear Mixed Model (LMM)s (whenever individual and/or group level random effect structures were indicated). The appropriateness of employed models was assessed in a stepwise manner with increasing model complexity. This process means that initially, null models were created to assess the support of random effect structures for individuals (level-1) and groups (level-2) subsequently. If random effects were not supported, Mixed ANOVAs were used afterwards. Otherwise, LMMs were used. Next, the main effects were included in the model and model quality criteria (Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and χ^2) assessed regarding fit improvements. Then, an interaction term between the two treatment variables was added, followed by the addition of various potential covariates. Models shown below always represent the best-fitting model as identified by this process. Fifth, whenever of interest, adequate regression analyses (e.g. non-linear regressions for report-physiology relationships) were utilised to further identify direct links between variables (as previous main effect tests only allow to test variable changes with the experiment conditions). Finally, it should be noted that in each analysis, p-value Benjamini-Hochberg (BH) correction is performed to reduce the inflation of false-positive error rates. The details per data domain (report, behaviour, or neurophysiology) are outlined in the following.

Data from participants who repeatedly failed control questions and who showed signs of uncooperative responding (close to zero variance in the survey responses) were removed. Overall, 39 usable SP data sets and 116 usable MP data sets (40 groups) were retained. Self-report constructs with multiple items were further analysed for their internal consistency (see Table 6.3). Having removed items that strongly reduced internal consistency, all variables exceed the recommended thresholds (e.g. 0.6 for Cronbach's Alpha - see Griethuijsen et al., 2015). In particular, one item each was removed from the flow scale, the general group satisfaction scale, the personal growth scale, the group diversity scale and the interaction quality scale. In total, adequate internal consistency was indicated for all constructs in both SP and MP conditions, with Alpha and Omega values passing general threshold recommendations. For two control variables (interdependence and adequate group size), the internal consistencies were poor, which is why only single items were used for the related analyses. The data were normally distributed as assessed by skewness and kurtosis values of $\leq \pm 1.5$. The only variable exceeding these thresholds is the individual difficulty variable in the SP EASY condition, likely caused by floor effects through the difficulty manipulation. As random effects (neither level-1 or level-2) were not supported for this variable, a parametric, mixed ANOVA test was employed that is robust against slight deviations from normality.

Objective data from the task conditions were extracted and processed analogously to the self-report data. The variables extracted for analyses include the average task (difficulty) level per condition and the count of correct trials per condition for DM checks. Similar metrics have been employed in related work with comparable experiments (Katahira et al., 2018; Ewing, Fairclough, and Gilleade, 2016). Furthermore, the elicited preferences for optimal difficulty were collected as a metric per individual and per group (range of preferences). After removal of outliers ($\leq$2 SD from the variable mean), normal distribution was confirmed (with the exception of the task level variable that has no variance in the EASY condition, and the SP HARD condition due to floor effects (often no correct trials in this condition). As random effects (neither level-1 or level-2) were not supported for these two variables, mixed ANOVAs were employed. These tests have been shown to be robust against slight deviations from normality and were therefore deemed appropriate in these cases. Overall, 41 usable SP data sets and 40 MP (group) data sets were retained.

ECG data were processed following the guidelines of Malik et al. (1996), using the Python toolbox BioSppy (Carreiras et al., 2015). The complete ECG processing

Variable (Items Retained)	Cronbach's Alpha	Avg. Item-Total Cor.	McDonald's Omega
Flow (9/10)	0.80 / 0.77	0.62 / 0.60	0.80 / 0.76
Stress (5/5)	0.83 / 0.82	0.77 / 0.76	0.83 / 0.82
Autonomy (3/3)	0.81 / 0.82	0.85 / 0.86	0.81 / 0.82
Soc. Presence (5/5)	– / 0.88	– / 0.82	– / 0.89
Int. Qual (2/3)	– / 0.78	– / 0.90	– / 0.79
Grp. Skills (3/3)	– / 0.70	– / 0.79	– / 0.71
Grp. Effort (3/3)	– / 0.73	– / 0.80	– / 0.75
Grp. Divers (2/3)	– / 0.77	– / 0.90	– / 0.77
Grp. Engmt. (3/3)	– / 0.70	– / 0.79	– / 0.70
Pers. Growth (2/2)	– / 0.80	– / 0.91	– / 0.80
Grp. Sat. (2/3)	– / 0.66	– / 0.86	– / 0.66
Coll. Effic. (4/4)	– / 0.86	– / 0.84	– / 0.85
Inf. Sharing (2/2)	– / 0.81	– / 0.92	– / 0.82

Notes: First Value = SP, Second Value = MP; Numbers in Parentheses = Retained/Measured Items; Int. = Interaction; Qual. = Quality; Grp. = Group; Divers. = Diversity; Engmt. = Engagment; Pers. = Personal; Sat. = Satisfaction; Coll. = Collective; Effic. = Efficacy; Inf. = Information.

TABLE 6.3: Latent Variable Internal Consistencies in Experiment 2.

pipeline is summarised in Appendix A.4 Table A.10. Notably, participants with implausible baseline values (e.g. >200 beats per minute in the eyes open resting phase) were removed from the data as they indicate data collection errors. Overall, 35 usable SP data sets and 104 usable MP data sets (38 groups) were retained. After outlier removal and natural log transformation of frequency domain features (a typical transformation for power-law distributed variables - see Harmat et al., 2015; Berntson, Quigley, and Lozano, 2007; Cohen, 2014) normal distribution of variables were confirmed. In this analysis, particular emphasis was placed on HRV features that reflect parasympathetic heart rate modulation (RMSSD, HF-HRV) due to their previous prevalence in Experiment 1, and their salient role in related work.

EEG data were processed primarily following the guidelines of Cohen (2014). The automated EEG data preparation process is outlined in detail in Appendix A.4 Table A.12. Data sets with recording errors or insufficient data quality were excluded before the execution of the pipeline through visual inspection. Signal data were additionally screened after signal processing to ensure no critical errors occurred in the pipeline. Overall, 34 usable SP data sets and 103 usable MP data sets (39 groups) were retained. Parameters for the processing steps were tuned for the Epoc+ EEG headset. For the feature aggregations, median averaging was used to reduce the impact of outliers in the data (Cohen, 2014). Frequency bands were extracted following Klimesch (1999). To account for inter-individual differences, Individualized Alpha Frequency (IAF) peaks were identified. As Alpha is also known to vary regionally (being slower at anterior sites), yet as not all participants showed such clear peaks for all sites, a global IAF maximum was determined as lying 0.5 Hz below the occipital Alpha maximum during an eyes-closed resting phase (see Figure 6.4). Based on this IAF, 2 Hz Theta and Alpha sub-bands were extracted (see Figure 6.5). To extend the personalised and band-refined approach, the Beta band was similarly decomposed. In line with previous research that has extracted low, mid, and upper Beta bands with 3 Hz, 5 Hz, and 10 Hz ranges respectively (Berta et al., 2013), the previous IAF-based decomposition was continued using these ranges. For the following analyses, wherever spatial aggregation (into ROI) was performed, only those observations were retained where >50% of electrodes had available data. After such aggregations, outliers (≤2 SD from the mean) were removed. Frequency band power features showed normal distributions. For the analyses of frontal Alpha activity, before analyses on homologous electrode pairs were conducted, difference scores were created (RH-LH), to assess the potential presence of FAA effects. FAA is related to approach-avoidance motivation (Harmon-Jones, Gable, and Peterson, 2010; Smith et al., 2017). This step is conducted not only to assert the adequacy of electrode pair pooling but also because FAA has been suggested to be present during flow (a state of high intrinsic motivation that elicits a desire to repeat the task) (Labonté-Lemoyne et al., 2016).

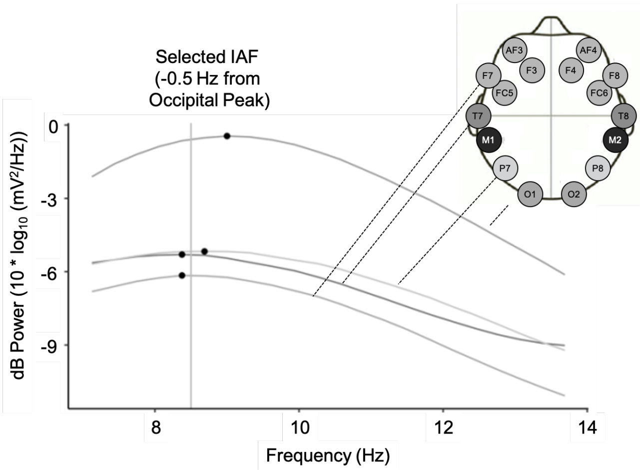

FIGURE 6.4: PSD for One Participant During Eyes-Closed Resting With Pooled Frontal (AF3, AF4, F3, F4, F7, F8, FC5, FC6), Temporal (T7, T8), Parietal (P7, P8), and Occipital (O1, O2) Sites. Dots Are Regional Maxima.

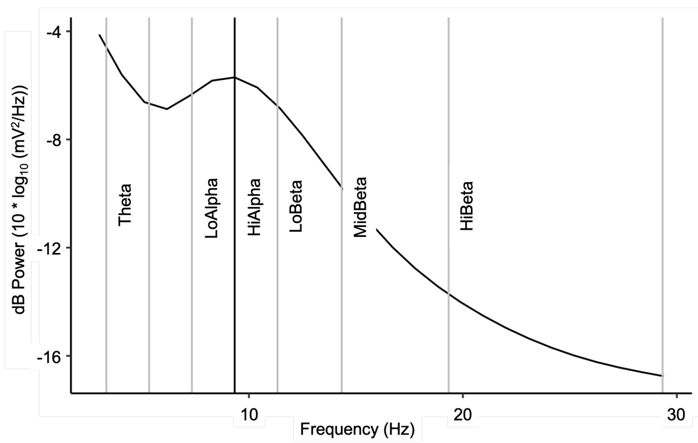

FIGURE 6.5: Grand Average PSD for All Participants During an Eyes-Closed Resting Phase With All 14 Electrodes Pooled, Demonstrating the Decomposition Into Narrow Theta, Alpha and Beta Frequencies.

6.3 Results

6.3.1 Manipulation Checks

Subjective & Objective Difficulty Variation

To check the success of the manipulation of difficulties, subjective perceptions and objective metrics of task difficulty were evaluated. Specifically, questions regarding perceived difficulty and perceived optimal difficulty were assessed (subjective difficulty) and task level and counts of correct trials per condition were analysed (objective difficulty). For all four variables, no model supported the inclusion of individual (level-1) or group (level-2) random effects. Therefore, two-way mixed ANOVAs (with Greenhouse-Geisser (GG) correction for violations of sphericity) were used to analyse the variable changes across conditions. Model statistics are shown in Table 6.4. Variable Distributions, including follow-up tests (one-way or RM ANOVAs and post-hoc Welch's t-tests), are shown in Figures 6.6 to 6.8.

Factor(s)	Test Statistic	P-Value	Effect Size
Perceived Difficulty			
Social Context	F(1, 133) = 7.3351	0.0153	0.0225
Difficulty	F(2.88, 383.45) = 344.6897	<0.001	0.6018
Social Context * Difficulty	F(2.88, 383.45) = 3.6663	0.0153	0.0158
Perceived Optimal Difficulty			
Social Context	F(1, 133) = 0.7801	0.3787	0.0018
Difficulty	F(2.59, 343.96) = 79.8077	<0.001	0.2940
Social Context * Difficulty	F(2.59, 343.96) = 9.3493	<0.001	0.0465
Task Level			
Social Context	F(1, 69) = 2.4139	0.2497	0.0156
Difficulty	F(1.42, 97.96) = 2172.9226	<0.001	0.9450
Social Context * Difficulty	F(1.42, 97.96) = 1.6441	0.2497	0.0128
Nr. of Correct Trials			
Social Context	F(1, 68) = 3.6103	0.0617	0.0183
Difficulty	F(1.84, 125.19) = 1583.6239	<0.001	0.9380
Social Context * Difficulty	F(1.84, 125.19) = 4.8990	0.0214	0.0447
Notes: P-Values are BH Corrected; Reported Effect Sizes are η^2_G.			

TABLE 6.4: Two-Way Mixed ANOVA for Subjective and Objective Task Difficulty in Experiment 2.

For individual task difficulty, a significant interaction between social context (SP or MP) and difficulty (EASY, CAL, AUTO, or HARD) was found. Subsequent tests for simple main effects for social context show higher perceived difficulty in the SP CAL and AUTO conditions. Simple main effects for difficulty show significant effects in both SP and MP conditions. Follow-up pairwise Welch's t-Tests show significant, stepwise increases in perceived task difficulty. Flow theory states not just moderate task difficulties (as would be supported so far), but optimally adapted difficulties as a requirement for deeper flow intensities (Nakamura and Csikszentmihalyi, 2009). Therefore, one item used by Ulrich et al. (2014) ("the task demands were well matched to my ability") was further analysed. For perceived optimal difficulty, a significant interaction was present between social context and difficulty. Subsequent tests for simple main effects for social context reveal better balance in the SP CAL condition (trend level) and the AUTO condition, and lower levels in the SP HARD condition. Simple main effects for difficulty show significant differences for both the SP and MP conditions. Follow-up Welch's t-tests show a significant, stepwise increase in optimal difficulty from EASY to the CAL and AUTO conditions and a decrease from these two conditions to HARD in both social contexts. In the MP condition, there are no significant differences between EASY and HARD and no significant differences between CAL and AUTO. In the SP condition, the perceived optimal difficulty is increased in the AUTO condition when compared to the CAL condition at trend level, and the HARD condition is significantly less optimal than the EASY condition.

To further evaluate the manipulation success beyond perceived measures, two objective metrics of task difficulty were evaluated. First, the average task level per difficulty condition was used to assess changes in difficulty. For this variable, no interaction effects were present for social context and difficulty. Main effects for difficulty were present, and no main effect was indicated for social context. Thus, subsequent pairwise tests were conducted that show a significant stepwise increase in average task level from EASY to CAL to AUTO to HARD. As a second metric of objective difficulty, the count of correctly solved trials per difficulty condition was evaluated. A significant interaction is found for this variable across social context and difficulty conditions. Subsequent tests for simple main effects for social context reveal a significantly lower number of correct trials in MP EASY and a higher number of correct trials in the MP HARD condition. Simple main effects for difficulty show significant changes in both the SP and MP conditions. Follow-up pairwise comparisons show significant, stepwise decreases in the number of correct trials with the difficulty conditions in both social context conditions.

In summary, these results illustrate a successful manipulation of (subjective and objective) difficulties with large effect sizes and a maximum of perceived optimal difficulty in the balanced difficulty conditions (CAL and AUTO). The optimality of difficulty varies with medium to large effect sizes across the difficulty conditions. It is maximised in the CAL and AUTO conditions, to a larger extent so in the SP condition. These results suggest that the manipulation meets not only necessary but sufficient criteria to elicit flow in different intensities. The comparison of objective difficulties indicates that the mental arithmetic tasks showed a broader range in difficulty for those completing them in isolation (SP).

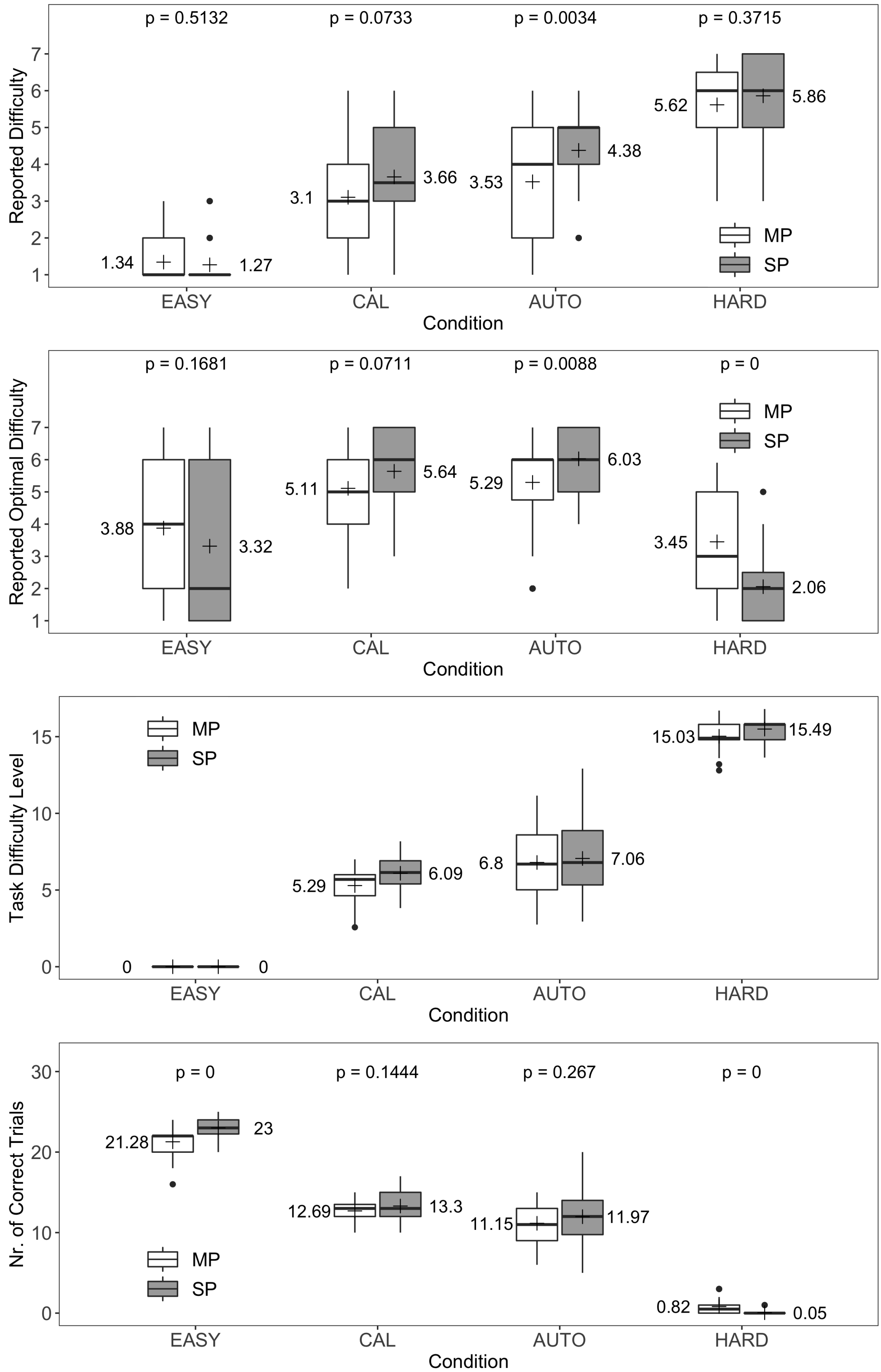

FIGURE 6.6: Difficulty Variable Distributions in Experiment 2. Follow-Up Tests are One-Way ANOVAs (MP vs SP). P-Values are BH-Corrected. Crosshairs and Numbers Next to Them Represent Means.

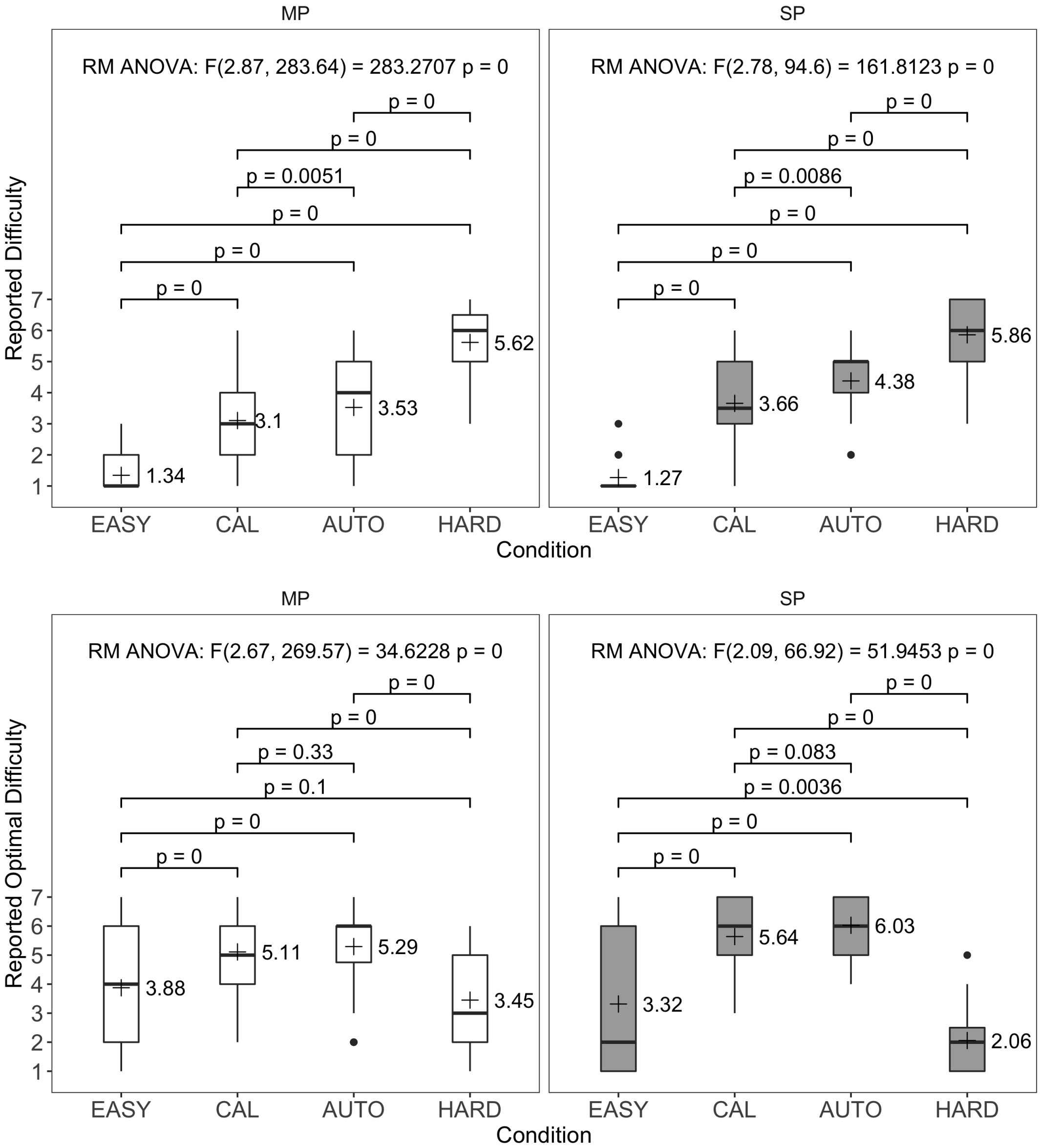

FIGURE 6.7: Subjective Difficulty Variables Pairwise Comparisons. P-Values are BH-Corrected. Crosshairs and Numbers Next to Them Represent Means.

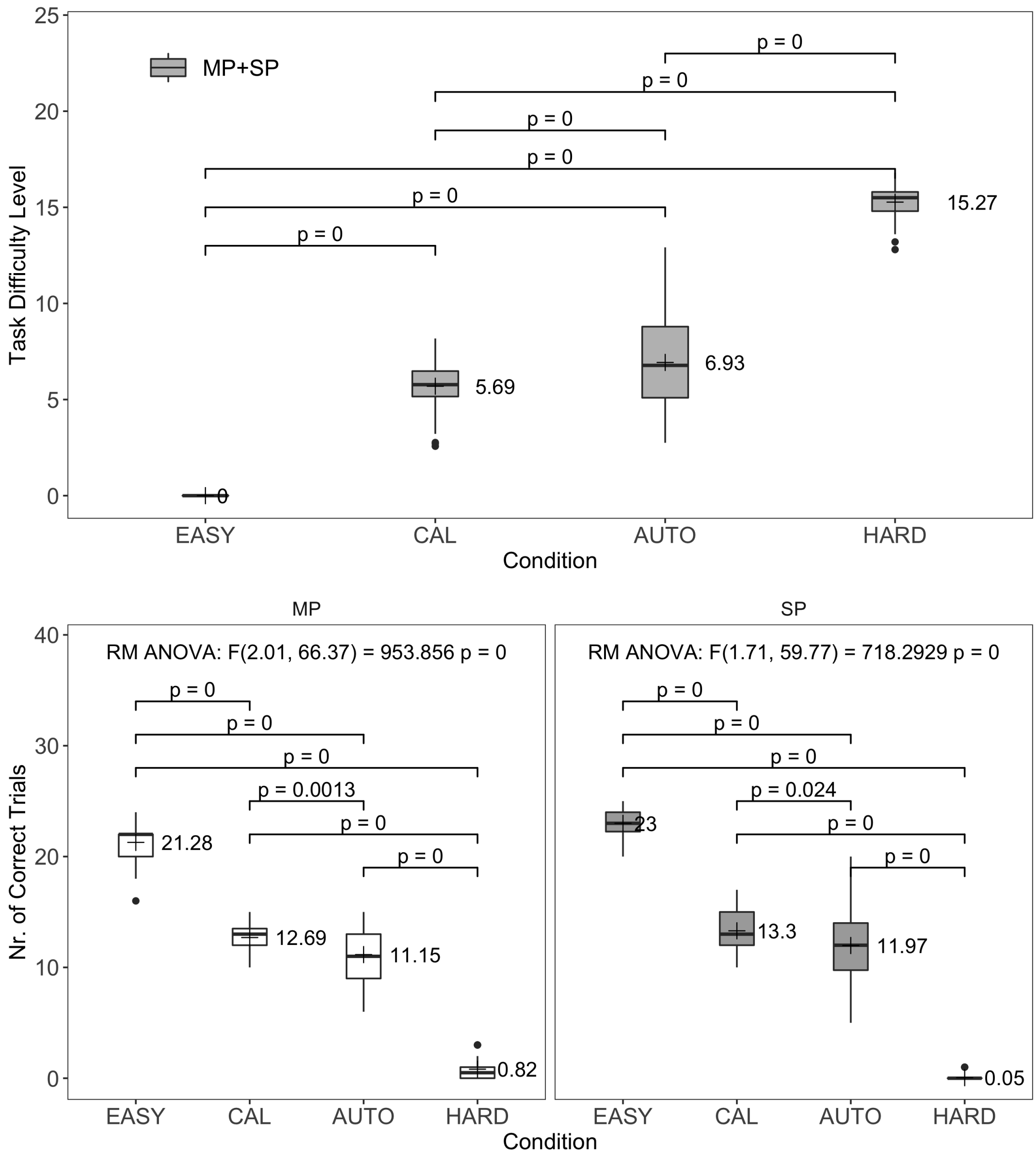

FIGURE 6.8: Objective Difficulty Variables Pairwise Comparisons. P-Values are BH-Corrected. Crosshairs and Numbers Next to Them Represent Means.

Autonomy in Optimal Difficulty Conditions

In a second step, the characteristics and comparability of the CAL and AUTO conditions were assessed. To report and observe these differences is important because previous research has found differences in flow intensity from different optimal difficulty calibration mechanisms (i.e. automatic calibration vs self-selection), yet reported mainly similar average difficulty levels in both treatments (Barros et al., 2018). As was previously seen, in both SP and MP conditions, the difficulty (subjective and objective) was higher in AUTO than in CAL. Thus, on average, the self-selected optimal difficulties were higher, which might indicate that the automatic calibration might have lead to sub-optimal, too low difficulties. However, the distributions of difficulty deltas in AUTO (i.e. the difference to the calibrated optimal difficulty in the CAL condition) show that not merely a too low difficulty was calibrated (see Figure 6.9). Instead, it can be seen that the majority of participants selected only slightly different difficulty levels (-3 to +3), and some preferred lower difficulties as well. An F-test does not indicate that the range in selected difficulties is different for the SP and MP conditions (p = 0.151). Therefore, the efficacy of the AUTO condition (and therefore the AM) receives support to calibrate optimal difficulties better.

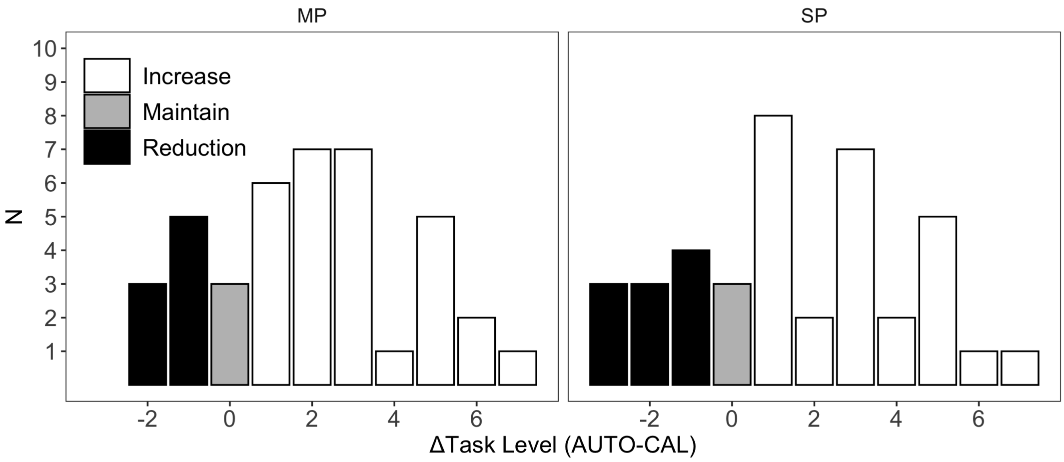

FIGURE 6.9: Selected Optimal Difficulties in Experiment 2.

In addition, the structure of the selected optimal difficulties in the MP AUTO condition was assessed. This step was taken to understand better if there are systematic variations in how groups might differ in terms of preferences for difficulty. To provide an overview, Figure 6.10 shows that it can be considered, that substantial range is present in the subjective difficulty preferences per group (median range = 6 levels of difficulty). As the shading in this figure shows, this range does not appear to be connected to the direction of how the difficulty level was realised (relative to the calibrated optimal difficulty in the CAL condition). This pattern means that some groups chose harder or lower difficulties unanimously or with more heterogeneity. Altogether, these results demonstrate, however, that substantial range is present in the difficulty preferences in the MP AUTO condition. As will be discussed below (specifically regarding the perceptions of autonomy) and as is apparent from the optimal difficulty variable, the MP AUTO condition might not have been similarly able to elicit more intense flow experiences due to this insensitivity to all group members' difficulty preferences. Therefore, the AM is only considered to have been partially successful (specifically more clearly for the SP condition).

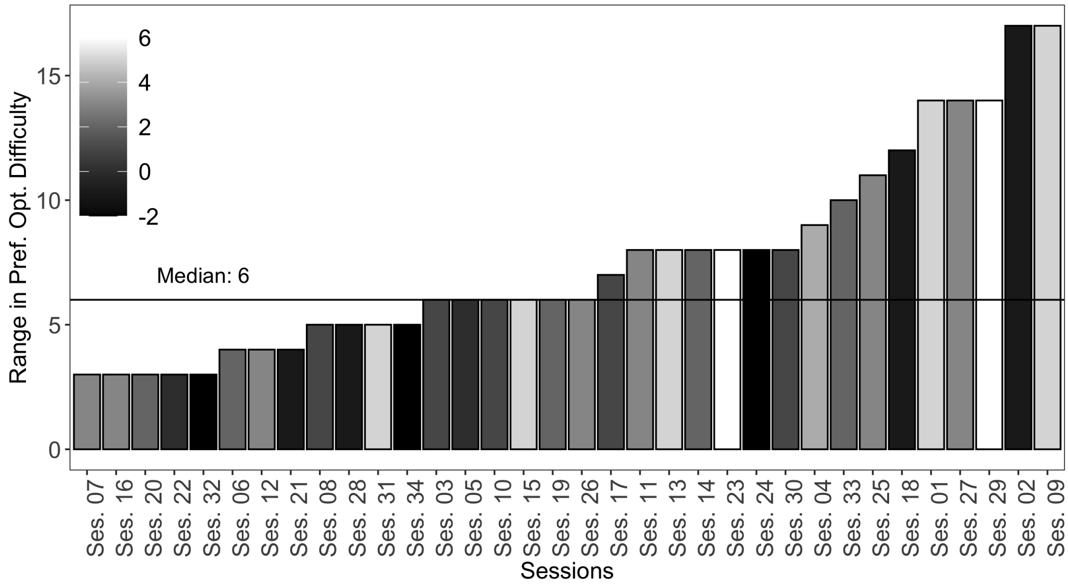

FIGURE 6.10: Selected Optimal Difficulties By Group in Experiment 2. Shadings Represent the Change in Optimal Task Level (AUTO - CAL).

Assessment of the Social Interaction Format

To further assess the quality of the social interaction, and especially to validate the successful integration of the key requirements for cooperation and integration, distributions of respective variables were assessed descriptively (see Figure 6.11). The boxplots show support for the successful integration of the requirements. The coordination was found to be sufficiently possible (single self-developed item), group members found to be mutually dependent, and the group size to be just right (both single items from Wageman, Hackman, and Lehman, 2005).

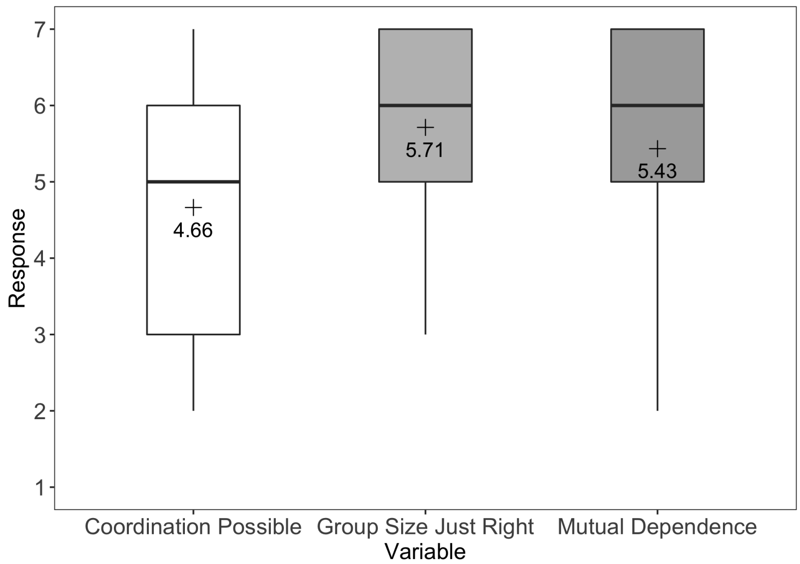

FIGURE 6.11: Perceptions of the Group Interaction Format in the MP Condition in Experiment 2. Crosshairs and Numbers Next to Them Represent Means.

6.3.2 Flow & Related Experiences

Flow

Flow reports were analysed using a multilevel LMM as the data had a hierarchical structure with individual-level variables (level-1) nested within small groups (level-2). In the stepwise model development process (see Table 6.5), the best model fit was indicated for model 3 that includes the main experiment effects (difficulty and social interaction) with their interaction, and additionally condition sequence (the order in time) and flow proneness as covariates. Results indicate a significant interaction between the manipulated variables social context and difficulty. Corrected post-hoc contrasts using the Estimated Conditional Means (ECM)s reveal a significant maximum of flow in the SP AUTO condition (compared to other SP and MP conditions). No other differences emerged between the SP and MP conditions. In both social contexts, flow was minimal in HARD compared to all other conditions. All pairwise comparisons are shown in Figure 6.12 and Figure 6.13.

Metrics	Flow Model 1	Model 2	Model 3
Fixed Effects			
Social Context^F	1.390	1.356	2.095
Difficulty^F	49.876***	52.635***	52.384***
Social Context * Difficulty^F	–	5.871***	5.916***
Condition Sequence^C	–	–	-0.0499 (0.0276)^t
Flow Proneness^C	–	–	0.6438 (0.1421)***
Random Effects			
Level 1 Resid. (Indiv.)	0.1614***	0.1712***	0.1392***
Level 2 Resid. (Grp.)	0.0368***	0.0361***	0.0259***
Goodness-of-Fit			
χ^2 (df)	130.611 (4)***	17.452 (3)***	22.917 (2)***
AIC	1423.7609	1412.3092	1393.3918
BIC	1458.5260	1460.1112	1449.8851
Marginal R^2	0.165	0.183	0.230
Conditional R^2	0.388	0.413	0.418

Notes: ^{t}p <.1; *p <.05; **p <.01; ***p <.001; ^{F} = Factor Variable (Showing F-Test Results); ^{C} = Continuous Variable (Showing Beta-Coefficients with SE); Random Effects Shown as Variance; χ^2 Goodness-of-Fit Statistics Obtained by Comparing Target Models with Their Previous Models (e.g. Model 1 to Model 0); P-Values are BH-Corrected.

TABLE 6.5: LMM for Perceived Flow in Experiment 2.

Altogether, it can be stated, that the typical, inverted U-shaped pattern of flow with increasing difficulty is fully visible for the SP condition (where the AUTO condition showed highest flow) and partially for the MP condition (where flow is only reduced in the HARD condition). Therefore variation of flow intensities by manipulation of difficulty is confirmed. However, with the SP AUTO condition showing the highest intensity, the intensification of flow by social context is considered to be rejected.

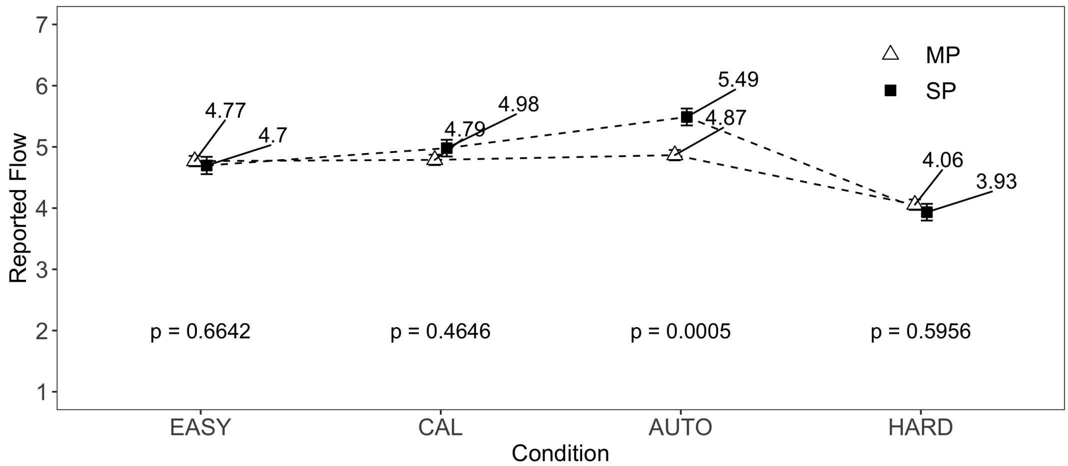

FIGURE 6.12: Flow Reports Per Condition in Experiment 2. P-Values are BH-corrected. Dots Are ECMs, Error Bars One SE.

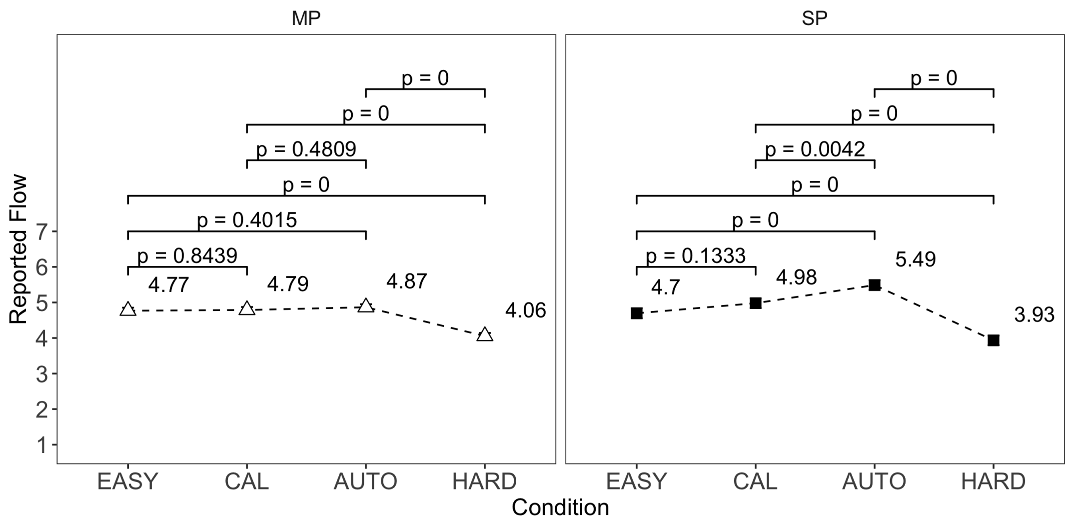

FIGURE 6.13: Flow Reports Post-Hoc Contrasts in Experiment 2. P-Values are BH-Corrected. Crosshairs and Numbers Next to Them Represent Means.

In much of the previous work on flow experiences in social interaction, the empiric tenor has been that social flow experience is intensified or facilitated (see, e.g. Magyaródi and Oláh, 2017; Tse et al., 2016; Walker, 2010). However, in the present results, not only, does flow in groups appear to be at most on par with the solitary intensities, but also in some instances is lower than in individuals working alone, which is the first finding of this type. This difference may be caused by less optimal difficulties in the MP than in the SP condition. Previous work might have found more intense flow in social interaction due to inverse difficulty patterns (i.e., a task for individuals might have been slightly too hard and for groups have had just the right difficulty). However, two additional explanations are possible given the two major novel elements in this experiment design (the optimal difficulty self-selection condition and the digitally mediated environment).

Autonomy Follow-Up

The first explanation for the unexpectedly similar or lower flow in the small groups could be caused by an inversion of opportunities for optimal difficulty selection (i.e. contrary to previous work, in isolation/groups, participants had more/less flexibility in difficulty selection). Considering, that the task was designed to be very simple (three equations for three participants – all with the same subtask difficulty), a major reduction of freedom to self-select a sub-task that is optimally difficult might have reduced flow in the MP condition. Previous studies might not have encountered this by using more complex task designs. Compare the present design, for example, to a puzzle task in related work (Tse et al., 2016). In their study, dyads cooperatively finished easy (3x3 tiles), moderate (4x4 or 5x5 tiles) or hard (6x6 tiles) puzzles. Such a task provides more opportunities and autonomy for individuals to select subtasks (i.e. puzzle piece combinations) that match their current preference for difficulty, than does the experiment design in this work. If this difference is a vital factor, the previous finding of intensified flow in groups could be mainly driven (i.e. moderated) by degrees of freedom and autonomy – not just the social interaction itself. In line with this reasoning, self-selected difficulty approaches have been found to elicit more intense flow in SP settings (Barros et al., 2018). Such an effect might be caused by better matching of task difficulty preferences (the one-size fits all for optimal difficulty setting has been generally critiqued for the DM paradigm - see Fong, Zaleski, and Leach, 2015; Løvoll and Vittersø, 2014). From the present data, it also appears that the SP AUTO condition has elicited the most intense flow experience. The same effect might have further reduced flow potentials in the MP condition.

To assess the possibility that autonomy differences might be related to more intense flow, follow-up analyses were conducted. First, autonomy perceptions (collected after each difficulty condition) were compared for the optimal difficulty conditions and the social context conditions (using an LMM with individual-level random effects and an interaction term for the variables difficulty and social context - see Table 6.6). Post-hoc

Metrics	Autonomy
Fixed Effects	
Social Context	1.7902
Difficulty	2.2211
Social Context * Difficulty	4.9390*
Random Effects	
Level 1 Resid. (Indiv.)	1.7584***
Level 2 Resid. (Grp.)	–
Goodness-of-Fit	
AIC	971.8618
BIC	994.0644
Marginal R^2	0.014
Conditional R^2	0.777
Notes: ^{t}p <.1; *p <.05; **p <.01; ***p <.001; Fixed Effects Shown as F-Test Results; Random Effects Shown as Variance; P-Values are BH-Corrected.	

TABLE 6.6: LMM for Perceived Autonomy in Experiment 2.

contrats (see Figure 6.14) show (with trend level significances), that the perceptions of autonomy were not generally higher in the SP condition, but only in the SP AUTO condition. Thus, a weak indirect link appears to be present, between increases in autonomy and flow. Note also the low values for marginal R^2 but high values for conditional R^2 that indicate that perceptions of autonomy are more strongly varied between persons than by the experimental manipulations. For the MP conditions, it was further assessed if perceived autonomy could be related to the range in group members' preferences. For this purpose, an artificial grouping variable was created (median split on the preference ranges) and entered into the previously described LMM (excl. social context due to focus on MP condition only). Model fit indices did not support the inclusion of this split or interaction with the condition variable. The distribution of the data after this split (see Figure 6.15) also does not indicate the presence of such an effect.

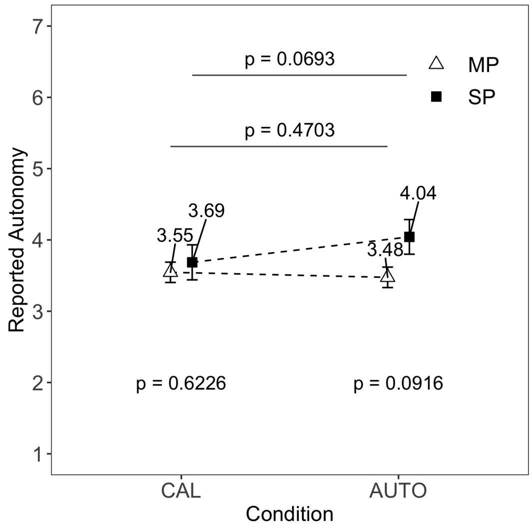

FIGURE 6.14: Perceived Autonomy in the Optimal Difficulty Conditions in Experiment 2. P-Values are BH-corrected. Dots Are ECMs, Error Bars One SE.

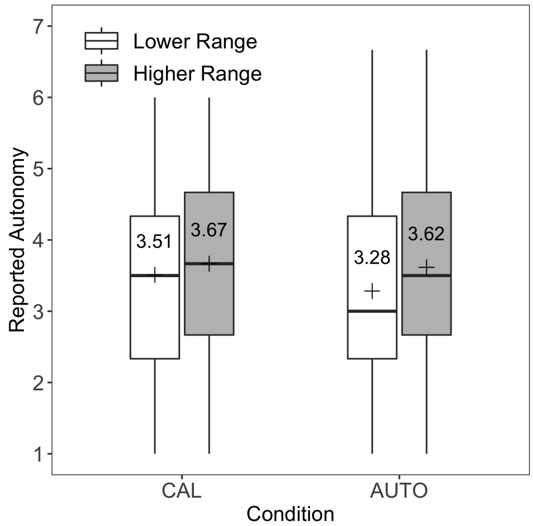

FIGURE 6.15: Perceived Autonomy in the MP Optimal Difficulty Conditions After a Median Split on Difficulty Preference Ranges. Crosshairs Are Means.

Second, an LMM was created to assess a possible direct link between flow and perceived autonomy. The results (see Table 6.7 and Figure 6.16) show significant positive and moderately sized relationships between the two variables, with a steeper slope in the SP condition. The interaction effect suggests that the social context moderates the relationship between autonomy and flow. Altogether, the proposition that increased autonomy is linked to an increase in flow receives further support.

Third, to assess whether or not increased autonomy is related to an improvement in difficulty calibration, a Linear Regression Model (LM) was created for perceptions of autonomy and optimal difficulty. The results (see Table 6.8 and Figure 6.17) show a significant and positive relationship between the two variables, that is again more pronounced for the SP condition, suggesting that changes in autonomy translate more directly into difficulty optimisation when no social interaction is present.

In summary, the findings for perceived autonomy, flow, and optimal difficulty, indicate that with increased autonomy (that was specifically higher in SP AUTO - and

Metrics	Flow
Fixed Effects	
Intercept	3.8195 (0.1230)***
Autonomy	0.2269 (0.0319)***
Social Context (SP)	-0.4065 (0.2326)
Social Context * Autonomy	0.1546 (0.0601)*
Random Effects	
Level 1 Resid. (Indiv.)	0.1148***
Level 2 Resid. (Grp.)	0.0266***
Goodness-of-Fit	
AIC	1400.9069
BIC	1431.1017
Marginal R^2	0.188
Conditional R^2	0.341
Notes: ^{t}p <.1; *p <.05; **p <.01; ***p <.001; Fixed Effects Shown as Beta-Coefficients with SE; Random Effects Shown as Variance; P-Values are BH-Corrected.	

TABLE 6.7: Autonomy-Flow LMM in Experiment 2.

Metrics	Optimal Difficulty
Fixed Effects	
Intercept	3.7512 (0.2323)***
Autonomy	0.1997 (0.0624)**
Social Context (SP)	-1.1245 (0.4443)
Social Context * Autonomy	0.2726 (0.1177)*
Goodness-of-Fit	
AIC	2247.4812
BIC	2269.0489
R^2	0.0578
Adj. R^2	0.0526
Notes: ^{t}p <.1; *p <.05; **p <.01; ***p <.001; Fixed Effects Shown as Beta-Coefficients with SE; P-Values are BH-Corrected.	

TABLE 6.8: Autonomy-Optimal Difficulty LM in Experiment 2.

at least not higher in MP in this experiment) better calibration of difficulty occurs. This calibration is again linked to higher flow experience intensities. Therefore, there is additional evidence present, that the lower flow intensities in the MP condition might be caused by an inverted pattern of freedom of choice for task difficulty. Previous findings of intensified flow in social units might, therefore, be confounded with the nature of how more freedom is provided in more complex tasks that are typically present in small group interaction settings.

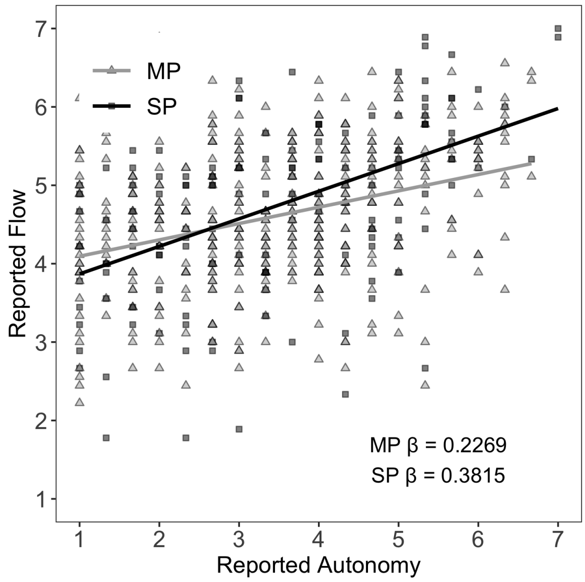

FIGURE 6.16: LMM Slopes of Autonomy and Flow in Experiment 2.

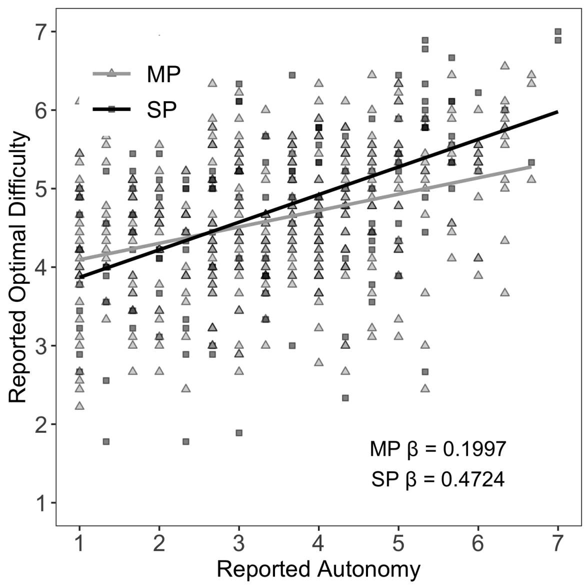

FIGURE 6.17: LMM Slopes of Autonomy and Optimal Difficulty in Experiment 2.

Social Context Follow-Up

Another central difference in the experiment design might have resulted in an inhibition of social context manipulation-based flow intensification, namely the social interaction format. Having opted for a digitally-mediated cooperation scenario, the interaction possibilities in the present design have been reduced to task-related action signalling only (i.e. participants could only show on which part of the problem they were currently working). This restriction is a central difference to previous work that has almost exclusively opted for face-to-face interaction settings (see Table 3.1). It could be possible, that not just any type of social interaction will do similarly well to facilitate flow and could even be the limiting factor to it. This possibility falls in line with the general understanding that digital media can be altering and limiting the "normal" (i.e. known from face-to-face environments) exchange of socio-affective communication (Derks, Fischer, and Bos, 2008; Chanel and Mühl, 2015).

To further explore the possibility of this effect, a relationship between perceptions of social presence (collected in the MP condition at the end of the experiment) and the average level of flow experiences across all difficulty conditions was compared. First of all, the distribution of the social presence rating indicates a rather low level of social presence (mean = 3.28, SD = 1.34, median = 3.2 - on a 7p Likert scale), which indicates the possibility, that a lack of social signals has been perceived in this experiment. Furthermore, a LM analysis between average flow and social presence indicates a (trend level) significant, weak positive relationship between the two variables (see Table 6.9 and Figure 6.18). As this effect is weak, the possibility of a link between social information and flow in social interaction does not receive much support. The social presence variable can however only be seen as a proxy for a variety of social factors that might influence the intensity of flow (e.g. emotional contagion - see Labonté-Lemoyne et al., 2016, or stress-buffering - see Tse et al., 2016). Therefore, while not receiving strong support in this instance, a relationship between social signals and flow intensity should be further investigated. However, as the present study did not directly compare a face-to-face to a digitally-mediated interaction scenario, or did not manipulate social signals in any other form, conclusions on these relationships require additional experiments.

Metrics	Flow
Fixed Effects	
Intercept	4.3153 (0.1548)***
Social Presence	0.0783 (0.0437)^t
Goodness-of-Fit	
AIC	201.5334
BIC	209.5798
R^2	0.0294
Adj. R^2	0.0202
Notes: ^{t}p <.1; *p <.05; **p <.01; ***p <.001; Fixed Effects Shown as Beta-Coefficients with SE; P-Values are BH-Corrected.	

TABLE 6.9: Social Presence - Avg. Flow LM in Experiment 2.

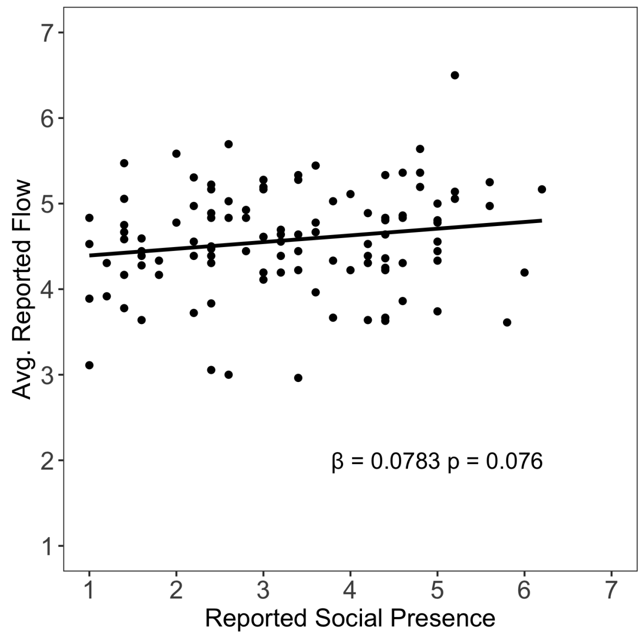

FIGURE 6.18: LM of Avg. Flow and Social Presence in Experiment 2.

Stress Follow-Up

To follow up on the relationships of flow with stress that was identified in Experiment 1, main effect and direct comparison (LMM) analyses were conducted. To reiterate, in Experiment 1, flow was found to be both accompanied by moderate and low stress levels. This observation prompted the question as to whether the experienced flow in the writing task might be qualitatively different. It could represent a flow experience without stress perceptions that would appear to be more in line with theoretical predictions (e.g. the concept of flow as a unique experience of effortlessness - see, e.g. Bruya, 2010; Ullén et al., 2010). To investigate this possibility further, first, a main effect analysis (i.e. for the manipulated variables) was conducted similar to the flow reports in the previous section. This indirect comparison was chosen specifically to identify whether increases in flow were accompanied with reductions of stress - that would be expected given the results in Experiment 1. Most notably, such a flow increase appeared here in the SP AUTO condition. The model with best indicated fit includes the manipulated variables as fixed effects (without interaction) and individual (level-1) random effects (see Table 6.10). The absence of an interaction effect is a first indicator that an increase in flow is not accompanied by a reduction in stress, as this would be expected given the disordinal interaction effect for flow (maximum in SP AUTO). The main effect for difficulty and the pairwise post-hoc contrasts indicate a slight increase of stress in the AUTO condition compared to the EASY condition, and a reduction of stress (significant at trend level) from the CAL condition to the AUTO condition (see Figure 6.19). First of all, these results show that stress perceptions increase disordinally with difficulty (that increases stepwise - see Figure 6.7). Furthermore, in the AUTO condition, while difficulties are elevated, and flow stays constant compared to lower difficulties (MP) or increases compared to lower difficulties (SP), stress perceptions show a tentative decrease. This observation adds weak, but additional support to the findings from Experiment 1 that higher flow levels might be associated with lower perceptions of stress. It is possible, however, that the results alternatively indicate that flow and stress are partially unrelated.

Metrics	Stress
Fixed Effects	
Social Context	2.0279
Difficulty	41.7950***
Random Effects	
Level 1 Resid. (Indiv.)	0.5392***
Level 2 Resid. (Grp.)	–
Goodness-of-Fit	
AIC	1774.4230
BIC	1805.1429
Marginal R^2	0.119
Conditional R^2	0.468
Notes: ^{t}p <.1; *p <.05; **p <.01; ***p <.001; Fixed Effects Shown as F-Test Results; Random Effects Shown as Variance; P-Values are BH-Corrected.	

TABLE 6.10: LMM for Perceived Stress in Experiment 2.

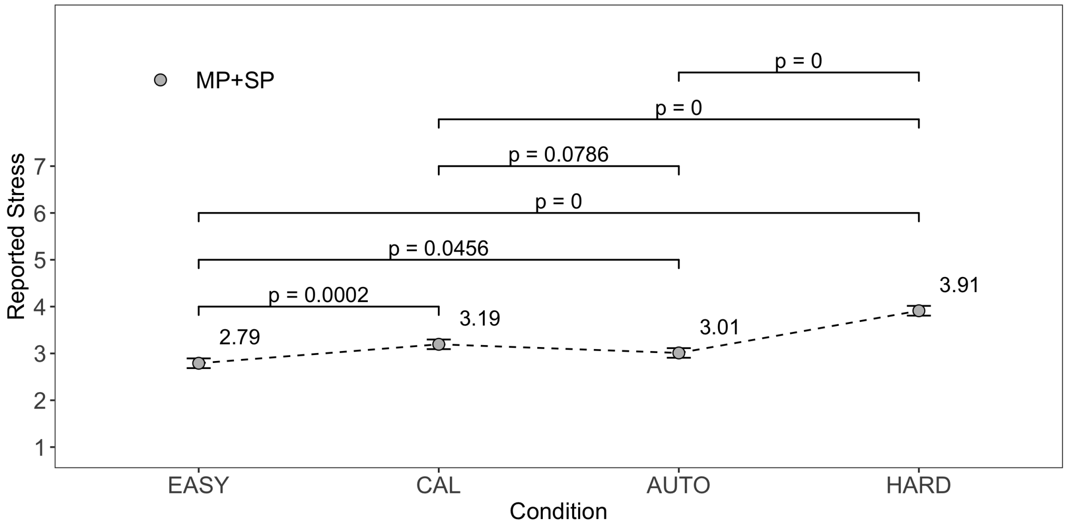

FIGURE 6.19: Stress Report Post-Hoc Contrasts in Experiment 2. P-Values are BH-corrected. Dots Are ECMs, Error Bars One SE.

To further assess the relationship between stress and flow, an LMM analysis was pursued. Given the previous results and after inspection of a scatterplot, polynomial terms were included in the model development. The best-fitting model included stress (as a quadratic term) and social context as fixed effects and individuals and small groups as random effects (see Table 6.11 for the model development). The results (see Figure 6.20) show a significant quadratic relationship of stress and flow that indicates an increasingly declining flow experience with higher stress perceptions.

Metrics	Flow
Fixed Effects	
Intercept	4.6798 (0.0516)***
Stress	-8.3939 (0.9002)***
Stress^2	-2.5019 (0.8648)**
Random Effects	
Level 1 Resid. (Indiv.)	0.1633***
Level 2 Resid. (Grp.)	0.0359***
Goodness-of-Fit	
AIC	1458.9573
BIC	1485.0731
Marginal R^2	0.150
Conditional R^2	0.359

Notes: ^{t}p <.1; *p <.05; **p <.01; ***p <.001; Fixed Effects Shown as Beta-Coefficients with SE; Random Effects Shown as Variance; P-Values are BH-Corrected.

TABLE 6.11: Stress-Flow LMM in Experiment 2.

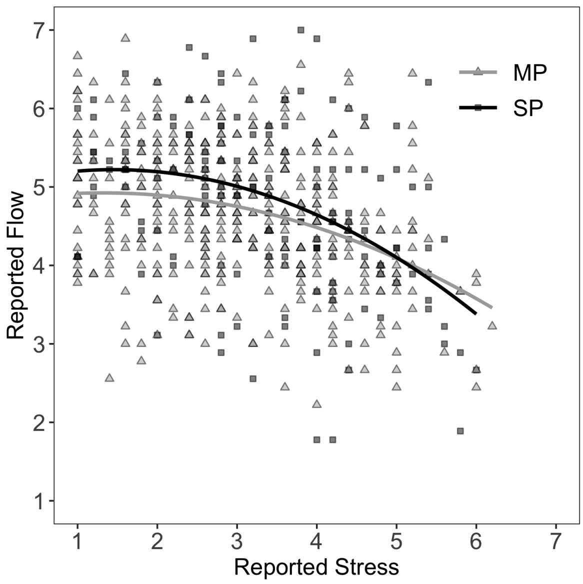

FIGURE 6.20: LMM of Flow and Stress Perceptions in Experiment 2.

Therefore, it can be stated that higher intensities of flow occur with lower perceptions of stress. However, the incompatibility of stress and flow increases mostly with higher levels of stress. On the one hand, these results suggest that with higher flow, stress is rather reduced, which together with the findings from Experiment 1, could indicate a shortcoming of present self-report instruments to capture the stress-absence (effortlessness) facet of the flow experience. Yet, this finding could also be seen as further support for the alternative proposition that flow and stress are partially unrelated (at least at the lower end of stress perceptions). Theoretical integrations proposed by (Tozman and Peifer, 2016) would account for this effect by explaining, that flow and stress become only incompatible when a task becomes too difficult, such that self-evaluative threats emerge. To further elaborate on what the actual pattern is, more data from naturalistic settings and laboratory experiments will have to be acquired. Nevertheless, the results highlight an important aspect for further development of the situational assessment of flow. Inclusion of stress or effortlessness dimensions into future flow self-report constructs could provide a valuable addition to assess the internal validity of flow measurements (the quality of experienced flow). An additional aspect to the questions about flow quality from Experiment 1 was the observation of high flow with lower stress in writing and a concomitant increase in physiological activation. This pattern suggested, that while flow was present and perceived as stress-/effortless, individuals still exerted a higher amount of energy, which further added to the possibility that high flow is perceived as effortless yet shows increased physiological effort. This pattern is similarly observed by (Harris, Vine, and Wilson, 2017a) and investigated in the next section.

6.3.3 Neurophysiology

Electrocardiography (ECG)

As previously discussed (Subsection 6.2.4) and to follow up on the results from Experiment 1, HRV-indicators of parasympathetic cardiac modulation were analysed to observe how the body balances physiological activation during flow experiences. Specifically, time-domain (RMSSD) and frequency-domain (HF-HRV) features were analysed. First, main effects from experimentally manipulated variables (social context and difficulty) were assessed, followed-up by direct (i.e. regression) analyses of flow-HRV relationships. For HRV-features, fixed effects were stepwise included for (1) experimentally manipulated variables, (2) their interactions, (3) time (stage one to four in the difficulty condition sequence), and (4) demographic covariates (age and gender - see, e.g. Valentini and Parati, 2009). To account for time covariates is important in HRV analyses with prolonged task exposures (see, e.g. Barros et al., 2018). For readability, the results from the best fitting models are reported. For the main effect analyses for both HRV features, LMMs include difficulty, social interaction, and time as fixed effects and individual (level 1) and small group (level 2) random effects (see Table 6.12). Time was found to be a significant and strong positive covariate. This finding indicates that with time, HRV levels increase. Demographic covariates (age and gender) were eliminated as they did not show improvements in model fits. The results (Table 6.12) show significant main effects for social context (trend level for RMSSD) and a main effect for difficulty for RMSSD at trend level. Therefore, first of all, a contrast is visible in how physiologically demanding the two social contexts were. Figure 6.21 shows that in the MP condition, parasympathetic activation was higher, than in the SP condition, which indicates a less demanding experience in MP.

It was further assessed if this finding represents a difference amongst the samples by comparing HF-HRV and RMSSD feature levels during the eyes open resting phase. One-way ANOVAs with social context as factor indicate no significant difference for HF-HRV ($F(1) = 1.1506$, $p = 0.2855$), and also not for RMSSD ($F(1) = 0.0769$, $p = 0.7820$). Therefore, the lower physiological demand in the MP condition is considered a consequence of the experiment design.

For the effect of task difficulty (Figure 6.21), it can be seen that HRV levels show a known pattern of reduction with difficulty from related work (see Section 4.2,

Metrics	HF-HRV	RMSSD
Fixed Effects		
Social Context^F	7.965*	3.892^t
Difficulty^F	1.661	2.426^t
Condition Sequence^C	0.0827 (0.0145)***	1.9243 (0.2514)***
Random Effects		
Level 1 Resid. (Indiv.)	0.2937***	87.4028***
Level 2 Resid. (Grp.)	0.0051*	0.1878***
Goodness-of-Fit		
AIC	727.5956	351.5257
BIC	765.5097	355.3116
Marginal R^2	0.066	0.059
Conditional R^2	0.725	0.719
Notes: ^{t}p <.1; *p <.05; **p <.01; ***p <.001; ^{F} = Factor Variable (Showing F-Test Results); ^{C} = Continuous Variable (Showing Beta-Coefficients with SE); Random Effects Shown as Variance; P-Values are BH-Corrected.		

TABLE 6.12: LMM for Parasymp.-HRV Features in Experiment 2.

Table 4.1). However, these effects are too small in the present data to pass statistical significance thresholds. Only the RMSSD feature shows a trend level effect for difficulty, with post-hoc contrasts (see Figure 6.22) indicating a (trend level) HRV minimum in HARD. Together, these analyses show poor sensitivity to difficulty changes for the HRV features in this experiment. Yet, the descriptive patterns are convergent with one particular view in related work (represented, e.g. by Peifer et al., 2014; Tozman et al., 2015; Bian et al., 2016). The present patterns suggest a moderate reduction of HRV during flow experience, specifically in the form of moderate activation or withdrawal from parasympathetic cardiac modulation.

In the direct relationship follow-up analysis, flow and HRV show a similar picture. Given the previous patterns and in line with that body of related work, non-linear relationships of cardiac activity and flow were analysed through LMMs with quadratic terms. The best-fitting models included orthogonal linear and quadratic terms and eliminated social context as a covariate. As with previous flow models, individual (level 1) and group (level 2) random effects were included. Initially, the results (Table 6.13) support the presence of quadratic relationships of flow with both HF-HRV and RMSSD (at trend level). Follow-up sensitivity analyses (including quadratic predictors only) supported the pattern for RMSSD ($t = -1.6771$, $p = 0.0944$), but not for HF-HRV ($t = -1.6318$, $p = 0.1036$). Therefore, the regression results confirm that only RMSSD shows a weak, inverted-U-shaped pattern with flow (Figure 6.23). Again, while weakly sensitive, these results are in line with most of the latest work and are therefore taken as further confirmation of findings, that flow is most likely represented by moderate physiological activation level - here specifically moderate parasympathetic activation.

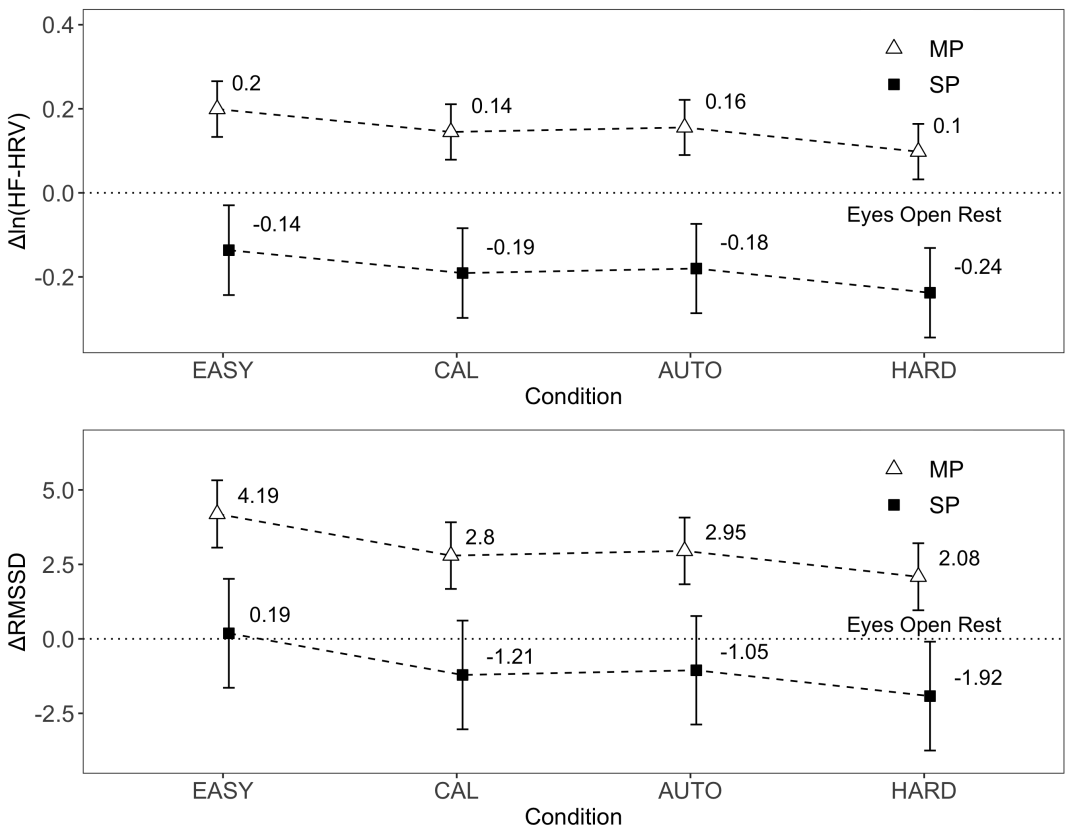

FIGURE 6.21: HRV in Experiment 2. P-Values are BH-corrected. Dots Are ECMs, Error Bars One SE.

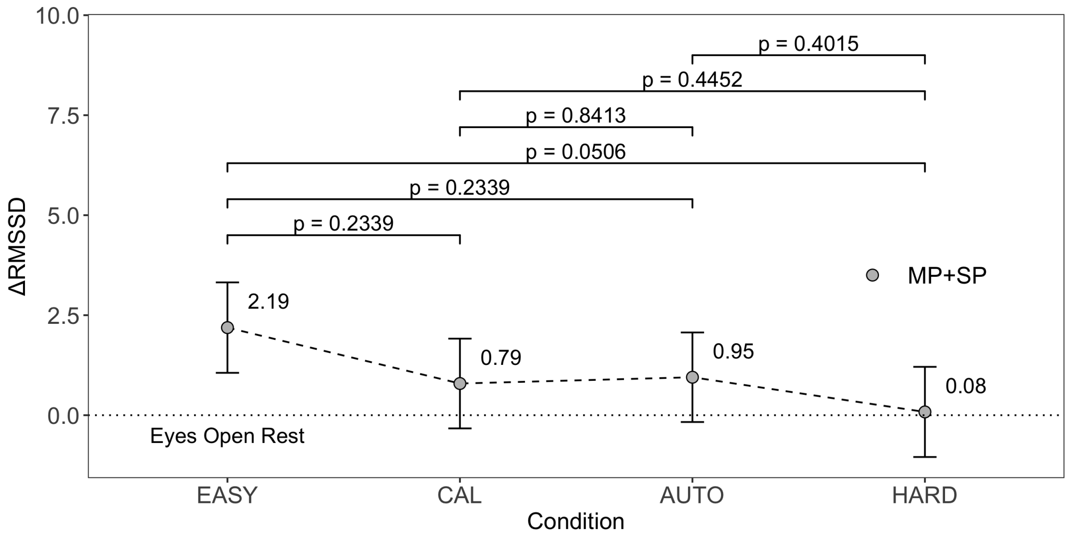

FIGURE 6.22: RMSSD Post-Hoc Contasts in Experiment 2. P-Values are BH-corrected. Dots Are ECMs, Error Bars One SE.

Metrics	Flow ~ RMSSD	HF-HRV
Fixed Effects		
Intercept	4.6768 (0.0534)***	4.6727 (0.0534)***
HRV	0.1689 (-1.0502)	0.5486 (-1.0489)
HRV^2	-1.8890 (-1.0062)^t	-1.8443 (0.9941)^t
Random Effects		
Level 1 Resid. (Indiv.)	0.1468***	0.1432***
Level 2 Resid. (Grp.)	0.0043**	0.0092**
Goodness-of-Fit		
AIC	1305.8518	1299.7415
BIC	1330.8569	1324.7717
Marginal R^2	0.008	0.029
Conditional R^2	0.172	0.188

Notes: ^{t}p <.1; *p <.05; **p <.01; ***p <.001; Fixed Effects Shown as Beta-Coefficients with SE; Random Effects Shown as Variance; P-Values are BH-Corrected.

TABLE 6.13: HRV-Flow LMMs in Experiment 2.

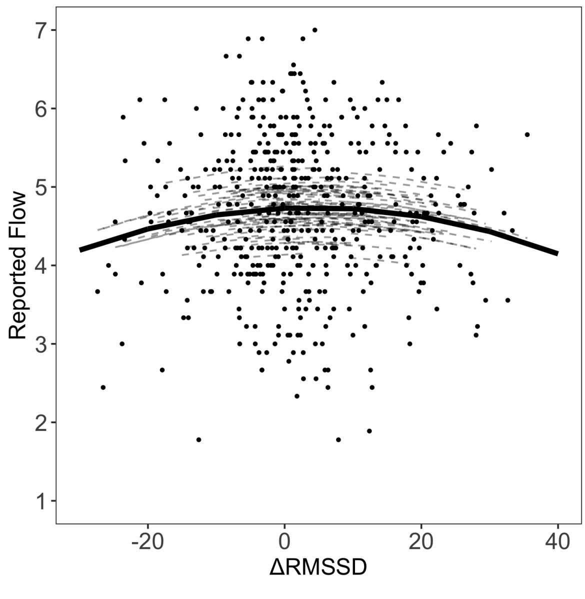

FIGURE 6.23: RMSSD-Flow LMM in Experiment 2. The Black Line Represents the LMM Quadratic Regression Prediction. Dashed Lines Represent Predicted Values per Subject with Random Intercept.

Electroencephalography (EEG)

To follow up on related work and the results from Experiment 1, EEG-indicators of mental workload (Theta and Beta frequency band powers over frontal and whole scalp ROI), and indicators of frontal downregulation (Alpha frequency band power) were assessed. In a central refinement to Experiment 1 and as a contribution to the body of related work, IAF-based frequency band personalisation, including band sub-segmentation, was employed to conduct more refined analyses of the patterns of interest. For EEG-features, fixed effects were stepwise included for (1) experimentally manipulated variables, (2) their interactions, (3) time (stage one to four in the difficulty condition sequence), and (4) demographic covariates (age, gender and handedness - see, e.g. Picton et al., 2000). To account for time covariates is important in EEG analyses with prolonged task exposures (see, e.g. Borghini et al., 2014; Wascher et al., 2014) For readability, the results from the best fitting models are reported.

Theta

Frontal Theta was assessed, that is changes in power in the -6 to -4 Hz range from the IAF at homologous frontal electrode pairs close to the midline - AF3 & AF4 (AF) and F3 & F4 (F-M). For AF, models did not support the inclusion of social context or difficulty main effects and were discarded from further analysis. At F-M, models supported interaction effects for these two experimentally manipulated variables at F-M. The latter model also includes time as a significant covariate and eliminated demographic covariates. In addition, individual (level-1) random effects are included. The results from the final model for F-M are shown in Table 6.14 and Figure 6.24. Given the corrected (trend level) interaction effect, pairwise post-hoc contrasts were assessed (see Figure 6.25). No differences between the SP and MP conditions were found. In the SP condition, the ECMs primarily show a pattern of increased frontal Theta from EASY, yet without significant differences. While this means that there is no statistical support for a relationship between frontal Theta and flow, the ECM distribution pattern is similar to related work. In such related work elevated frontal Theta levels during flow are found (see, e.g. Fairclough et al., 2013; Ewing, Fairclough, and Gilleade, 2016; Soltész et al., 2014). The MP condition contrasts the patterns in SP, as HARD shows significantly reduced Theta power (when compared to EASY and AUTO), and a tentatively elevated Theta level in EASY.

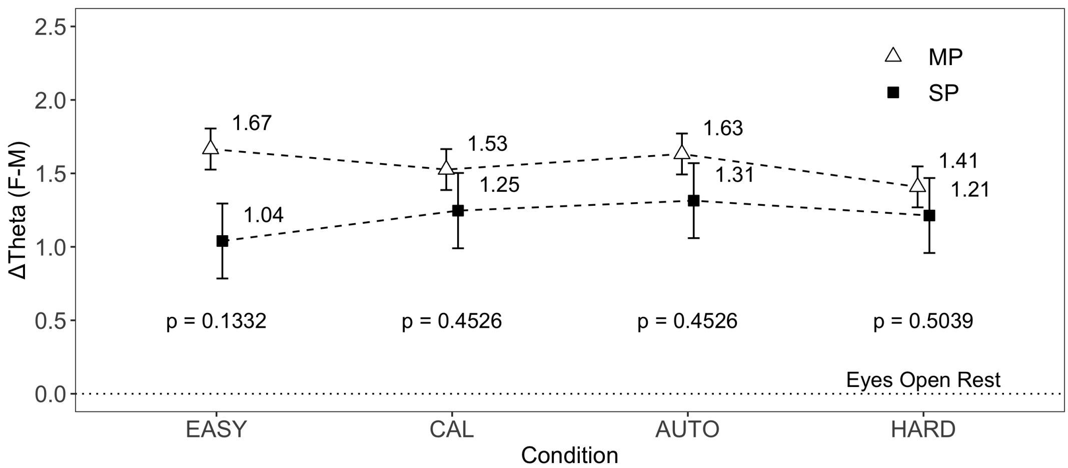

FIGURE 6.24: F-M Theta ECMs in Experiment 2. P-Values are BH-corrected. Dots Are ECMs, Error Bars One SE.

In subsequent regression analyses, the lack of sensitivity of frontal Theta in this data is indicated further. No differences or interactions with social interaction context (SP vs MP) were found. Yet, initially, a quadratic relationship between frontal Theta and flow is observed (see Table 6.15). However, follow-up sensitivity analysis that include only the quadratic term in the regression no longer support the presence of a quadratic

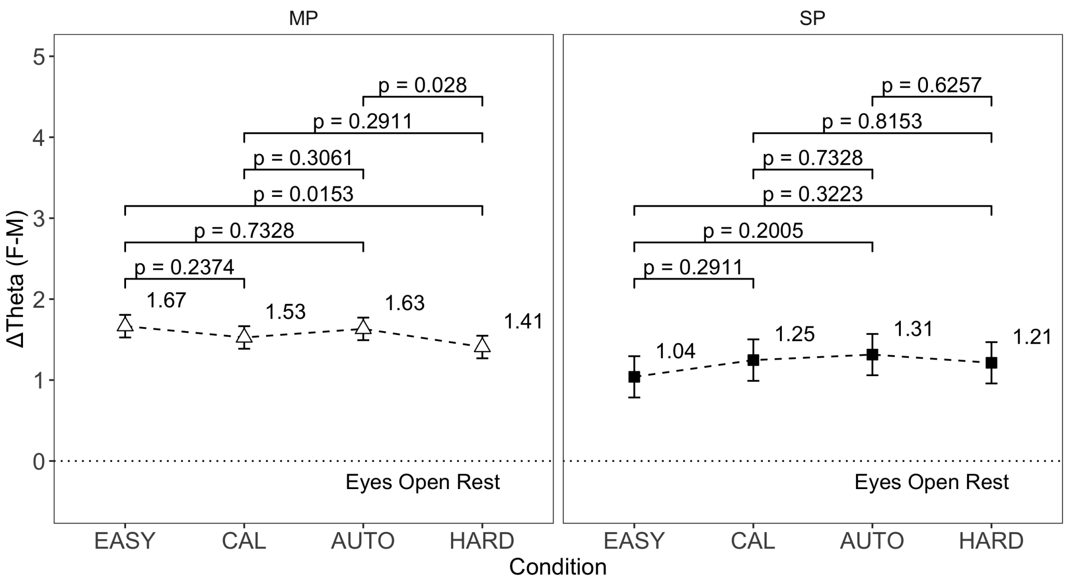

FIGURE 6.25: F-M Theta Post-Hoc Contrasts in Experiment 2. P-Values are BH-corrected. Dots Are ECMs, Error Bars One SE.

relationship (t = -0.3560, p = 0.7223). This finding further confirms the likeliness that a lack of sensitivity for frontal Theta activation is present. Such a finding might be best explained by the absence of midline electrodes, over which frontal Theta activation is typically found concerning changes in difficulty (see Borghini et al., 2014; Ewing, Fairclough, and Gilleade, 2016). To conclude the frontal Theta results presentation, it should be noted that the observed patterns showed no considerable difference when extracting Theta frequency band power using generalised thresholds (i.e. 4-7.5 Hz

Metrics	Theta F-M
Fixed Effects	
Social Context^F	1.6779
Difficulty^F	1.4832
Social Context * Difficulty^F	2.7406^t
Condition Sequence^C	0.0574(0.0219)*
Random Effects	
Level 1 Resid. (Indiv.)	1.324***
Level 2 Resid. (Grp.)	-
Goodness-of-Fit	
AIC	839.8971
BIC	883.3832
Marginal R^2	0.020
Conditional R^2	0.876
Notes: ^{t}p <.1; *p <.05; **p <.01; ***p <.001. ^{F} = Factor Variable (Showing F-Test Results); ^{C} = Continuous Variable (Showing Beta-Coefficients with SE); Random Effects Shown as Variance; P-Values are BH-Corrected.	

TABLE 6.14: LMM for F-M Theta in Experiment 2.

band power). Overall, the results can only indicate on a descriptive level that during flow, a moderate increase of frontal Theta activation might be present, that would represent moderate demands. At least, such a finding integrates with related work and with flow theory (flow as the result of a situation with demand-skill balance).

Metrics	Flow ~Theta F-M
Fixed Effects	
Intercept	4.7038 (0.0662)***
Power	0.7737 (1.0490)
Power^2	-2.8543 (1.0077)**
Random Effects	
Level 1 Resid. (Indiv.)	0.0722***
Level 2 Resid. (Grp.)	0.0873***
Goodness-of-Fit	
AIC	974.7435
BIC	998.0934
Marginal R^2	0.027
Conditional R^2	0.204

Notes: ^{t}p <.1; *p <.05; **p <.01; ***p <.001; Fixed Effects Shown as Beta-Coefficients with SE; Random Effects Shown as Variance; P-Values are BH-Corrected.

TABLE 6.15: F-M Theta-Flow LMM in Experiment 2.

Alpha

Frontal Alpha, that is power in the 2 Hz wide sub-bands ranging from -2 to +2 Hz range from the IAF, was inspected at homologous frontal electrode pairs (AF3+AF4 = AF, F3+F4 = F-M, F7+F8 = F-L, FC5+FC6 = FC). No lateral asymmetries were indicated for these paired sites (see Table 6.16 and Table 6.17), which indicated suitability of pooling for these frequencies and ROIs. For AF, F7 & F8 (F-L), and FC5 & FC6 (FC), models did not support the inclusion of social context or difficulty main effects and were discarded from further analysis. Thus, results for homologous electrode pairs show that only at F-M sites, the experimentally manipulated variables indicate main effects for both the LoAlpha and the HiAlpha band. The final models have the same form as those for frontal Theta. Specifically, models include interaction effects for social interaction and difficulty and condition sequence as a significant covariate. In addition, individual (level-1) random effects are included. The results from the final models are shown in Table 6.18 and Figure 6.26. Only difficulty was found to have a main effect on frontal Alpha. Pairwise post-hoc contrasts show a reduction of frontal Alpha from EASY for both LoAlpha and HiAlpha to all other difficulty conditions (see Figure 6.26). These findings are in line with studies that have found a frontal Alpha reduction during flow (Ewing, Fairclough, and Gilleade, 2016), and other phenomena such as Working Memory (WM) load, top-down attention, or mental fatigue (Borghini et al., 2014; Deiber et al., 2007; Wascher et al., 2014).

Metrics	LoAlpha AF	LoAlpha F-M	LoAlpha F-L	LoAlpha FC
Fixed Effects				
Social Context	1.2443	2.3373	3.5309	0.2646
Difficulty	1.3550	0.3503	1.4646	3.1415
Random Effects				
Level 1 Resid. (Indiv.)	0.7354***	1.8724***	0.5160***	1.0135***
Level 2 Resid. (Grp.)	0.1839***	-	0.1338***	-
Goodness-of-Fit				
AIC	796.4535	1097.0984	720.5125	911.8465
BIC	827.5201	1125.6419	751.7117	939.8217
Marginal R^2	0.014	0.019	0.034	0.007
Conditional R^2	0.789	0.867	0.775	0.800

Notes: ^{t}p <.1; *p <.05; **p <.01; ***p <.001; Fixed Effects Shown as F-Test Results; Random Effects Shown as Variance; P-Values are BH-Corrected; Model Fit Indices Also Showed No Improvement of the Shown Model Compared to Their Null Model.

TABLE 6.16: LoAlpha FAA Evaluation in Experiment 2.

Metrics	HiAlpha AF	HiAlpha F-M	HiAlpha F-L	HiAlpha FC
Fixed Effects				
Social Context	1.2888	2.6467	3.7732	0.0566
Difficulty	3.1348	0.3536	0.4033	2.8254
Random Effects				
Level 1 Resid. (Indiv.)	0.9824***	1.7282***	0.5110***	0.8908***
Level 2 Resid. (Grp.)	-	-	0.0921***	-
Goodness-of-Fit				
AIC	788.0449	1090.2653	699.7471	881.6616
BIC	815.2670	1118.8249	730.9682	909.6194
Marginal R^2	0.017	0.021	0.033	0.005
Conditional R^2	0.811	0.858	0.773	0.787

Notes: ^{t}p <.1; *p <.05; **p <.01; ***p <.001; Fixed Effects Shown as F-Test Results; Random Effects Shown as Variance; P-Values are BH-Corrected; Model Fit Indices Also Showed No Improvement of the Shown Model Compared to Their Null Model.

TABLE 6.17: HiAlpha FAA Evaluation in Experiment 2.

Metrics	LoAlpha F-M	HiAlpha F-M
Fixed Effects		
Social Context^F	0.0777	0.0053
Difficulty^F	5.4098**	13.3641***
Social Context * Difficulty^F	1.1937	11.197
Condition Sequence^C	0.0783 (0.0190)***	0.0706 (0.0182)***
Random Effects		
Level 1 Resid. (Indiv.)	1.3660***	1.3944***
Level 2 Resid. (Grp.)	-	-
Goodness-of-Fit		
AIC	954.8888	915.2611
BIC	999.8683	960.1404
Marginal R^2	0.014	0.017
Conditional R^2	0.879	0.892
Notes: ^{t}p <.1; *p <.05; **p <.01; ***p <.001; ^{F} = Factor Variable (Showing F-Test Results); ^{C} = Continuous Variable (Showing Beta-Coefficients with SE); Random Effects Shown as Variance; P-Values are BH-Corrected.		

TABLE 6.18: LMM for F-M Alpha in Experiment 2.

In subsequent regression analyses, no direct relationship of flow with frontal Alpha is found (see Table 6.19). No differences or interactions with social interaction context (SP or MP) were supported for either model. Initially, a quadratic relationship between frontal LoAlpha and flow is observed. However, a follow-up sensitivity analysis that includes only the quadratic term in the regression no longer support the effect (t = -1.4905, p = 0.1377). Also, for both LoAlpha and HiAlpha, no linear relationships are detected.

Therefore, the results suggest together, that during flow, no general frontal down-regulation is present. Instead, neural activity in frontal regions close to the midline (yet not spread out further) is increased. This increased activation is likely related to attentional engagement towards the task (see Klimesch, 1999; Deiber et al., 2007). Yet, while this is a known observation from other paradigms in neuroscience, the results further indicate that in situations with balanced (optimal) difficulty, no Alpha peculiarities emerge that might be related to the flow experience. The identification of the topographical restriction of Alpha activity is, however, important to note, as the observation of larger areas might have led scholars to different conclusions. The segmentation of the Alpha band was in this instance not found to be of relevance as the observed effects are similarly found in both Lo- and Hi-Alpha sub-bands. This means that the previous findings from Experiment 1, (and some related work) might also be artefacts, for example resulting from the utilisation of non-personalised frequency ranges.

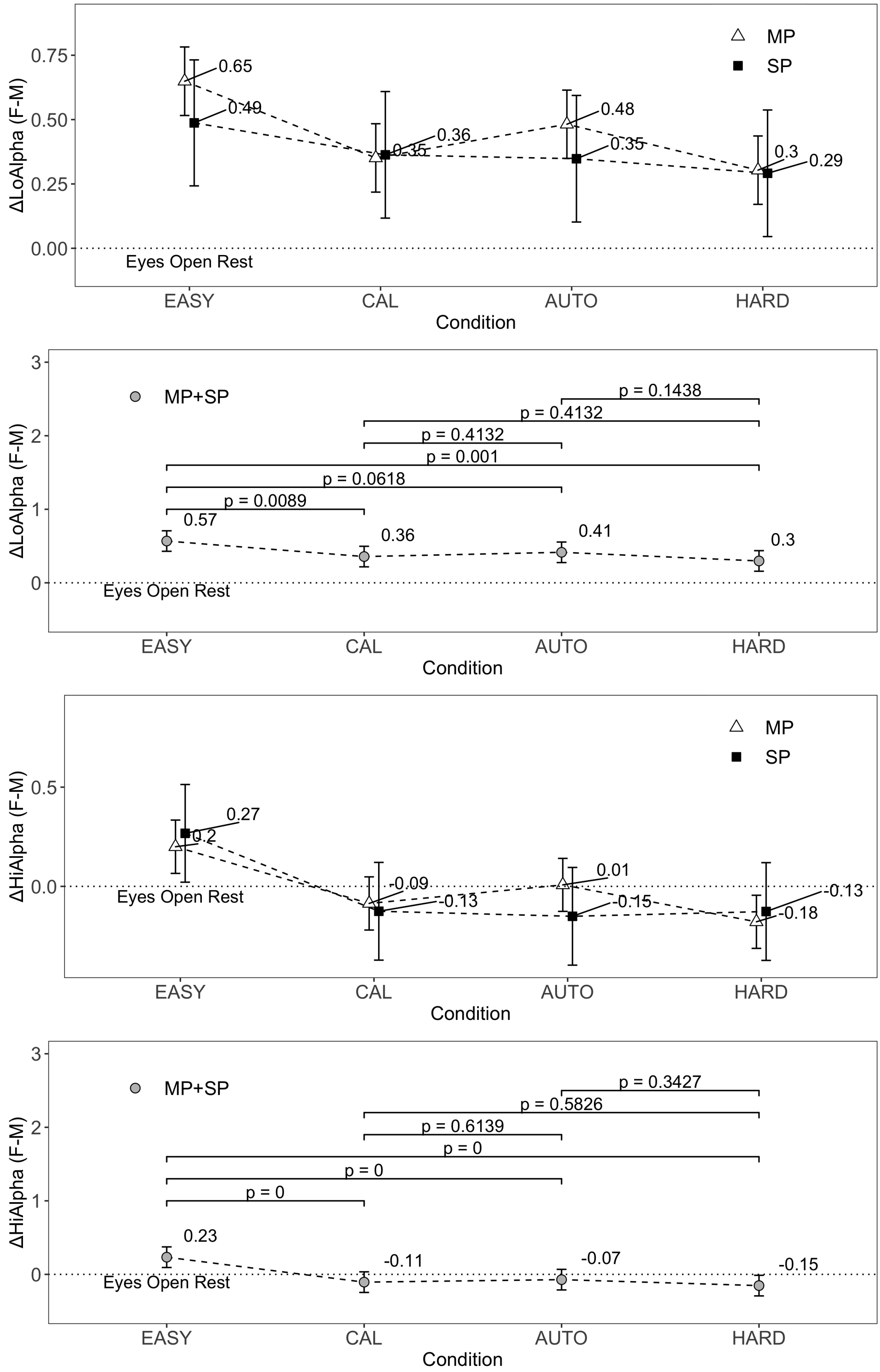

FIGURE 6.26: F-M Alpha ECMs & Pairwise Post-Hoc Contrasts in Experiment 2. P-Values are BH-corrected. Dots Are ECMs, Error Bars One SE.

Metrics	Flow ~ LoAlpha F-M	HiAlpha F-M
Fixed Effects		
Intercept	4.6920 (0.0585)***	4.7051 (0.0599)***
Power	0.6930 (-1.0282)	-0.0021 (0.0380)
Power^2	-2.0225 (-1.0087)*	-
Random Effects		
Level 1 Resid. (Indiv.)	0.0587***	0.0725***
Level 2 Resid. (Grp.)	0.0932***	0.0975***
Goodness-of-Fit		
AIC	1320.1541	1315.6773
BIC	1345.3572	1336.6391
Marginal R^2	0.010	0.000
Conditional R^2	0.182	0.193
Notes: ^{t}p <.1; *p <.05; **p <.01; ***p <.001; Fixed Effects Shown as Beta-Coefficients with SE; Random Effects Shown as Variance; P-Values are BH-Corrected.		

TABLE 6.19: F-M Alpha-Flow LMM in Experiment 2.

Beta

Beta activity, that is power in 3, 5 and 10 Hz wide sub-bands ranging from +2 to +20 Hz range from the IAF was analysed at four ROI over the whole scalp. This set of four ROI (Frontal = AF3 + F3 + F7 + F8 + F4 + AF4; Central = FC5 + T7 + T8 + FC6; Posterior = P7 + O1 + O2 + P8; Whole Scalp = All 14 electrodes) was chosen to involve fewer tests, and the selection was based on the reported reactivity of these locations to variations in mental workload (Michels et al., 2010). All models were developed in the same way as for previous frequency bands. For LoBeta, models did not support the inclusion of social context or difficulty main effects and were discarded from further analysis. For MidBeta, models supported the inclusion of these main effect variables for all sites. Posterior sites, however, showed no significant main effects, but frontal, central and whole scalp ROI indicated main effects, without interactions. All three sites showed elevated MidBeta in the MP condition (see Table 6.20 and Figure 6.27). Difficulty effects were indicated for frontal (at trend level) and central MidBeta. Post-hoc contrasts (see Figure 6.28) did not further confirm this effect for frontal sites, but show a stepwise increase with difficulty in central MidBeta from CAL to AUTO and HARD, with no difference between EASY and CAL.

Metrics	MidBeta Frontal	MidBeta Central	MidBeta Posterior	MidBeta Whole Scalp
Fixed Effects				
Social Context^F	6.3659*	6.4950*	1.2344	6.4671*
Difficulty^F	2.3200^t	7.6698***	0.1375	0.9989
Condition Sequence^C	-	-	0.0458^t (0.0231)	0.0498* (0.0178)
Diff. * Cond. Seq.^{F}	-	-	0.5114	0.2478
Soc. Co. * Diff.^{F}	-	-	0.4941	-
Soc. Co. * Cond. Seq.^{F}	-	-	3.9025*	-
Soc. Co. * Diff. * Cond. Seq^F.	-	-	0.3854	-
Age^C	-	-	-	0.0124 (0.0124)
Gender^F	-	-	-	2.2998
Handedness^F	-	-	-	0.7423
Random Effects				
Level 1 Resid. (Indiv.)	0.4978***	0.8216***	0.8134***	0.5198***
Level 2 Resid. (Grp.)	0.0036*	-	-	-
Goodness-of-Fit				
AIC	1009.2431	1136.8294	1195.0056	1012.1950
BIC	1043.2583	1166.3874	1271.0120	1075.8283
Marginal R^2	0.039	0.049	0.024	0.082
Conditional R^2	0.715	0.763	0.811	0.747

Notes: ^{t}p <.1; *p <.05; **p <.01; ***p <.001; ^{F} = Factor Variable (Showing F-Test Results); ^{C} = Continuous Variable (Showing Beta-Coefficients with SE); Random Effects Shown as Variance; Models Per ROI Chosen Due to Best Fit; P-Values are BH-Corrected.

TABLE 6.20: LMMs for MidBeta in Experiment 2.

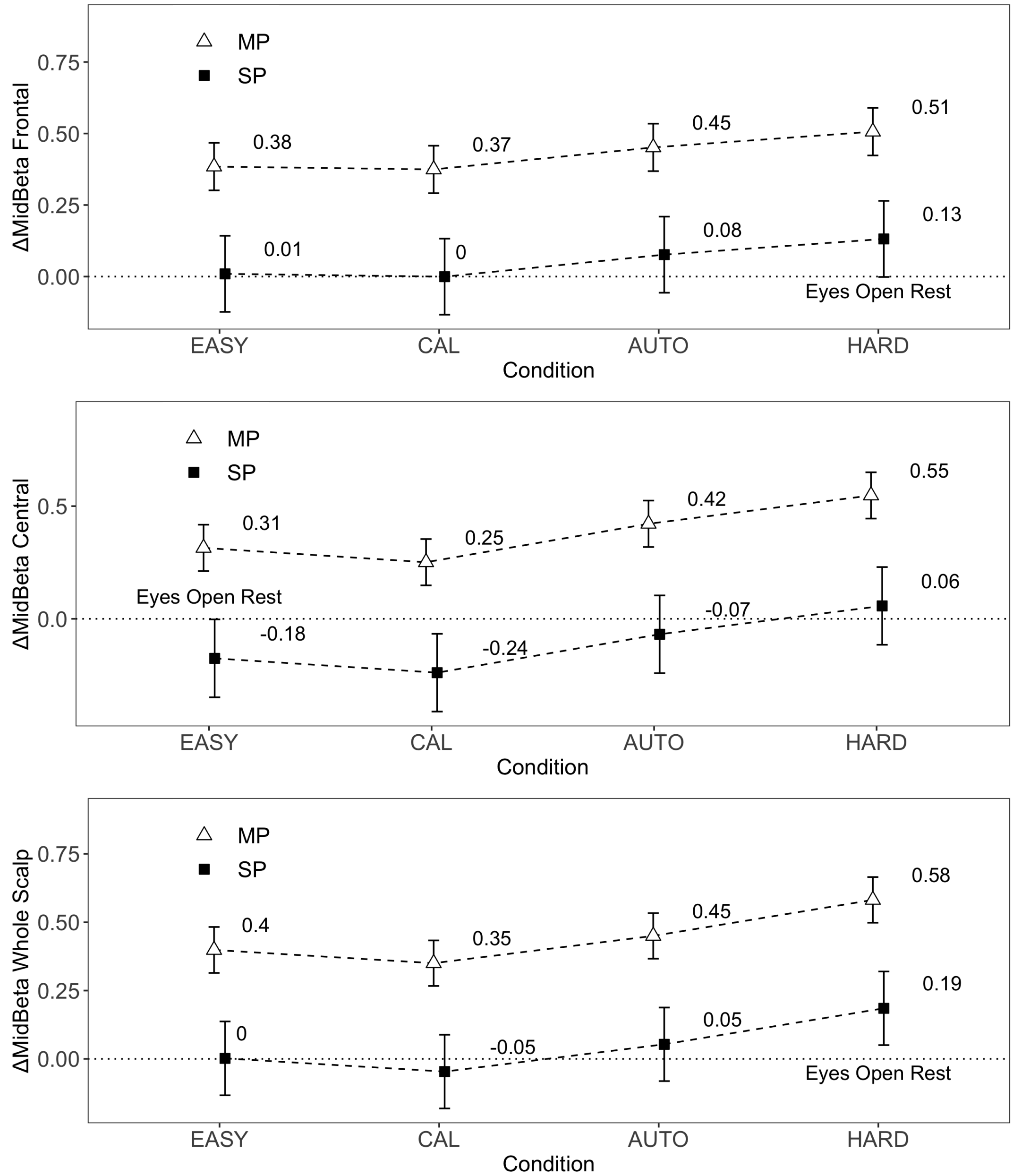

FIGURE 6.27: MidBeta ECMs in Experiment 2. P-Values are BH-corrected. Dots Are ECMs, Error Bars One SE.

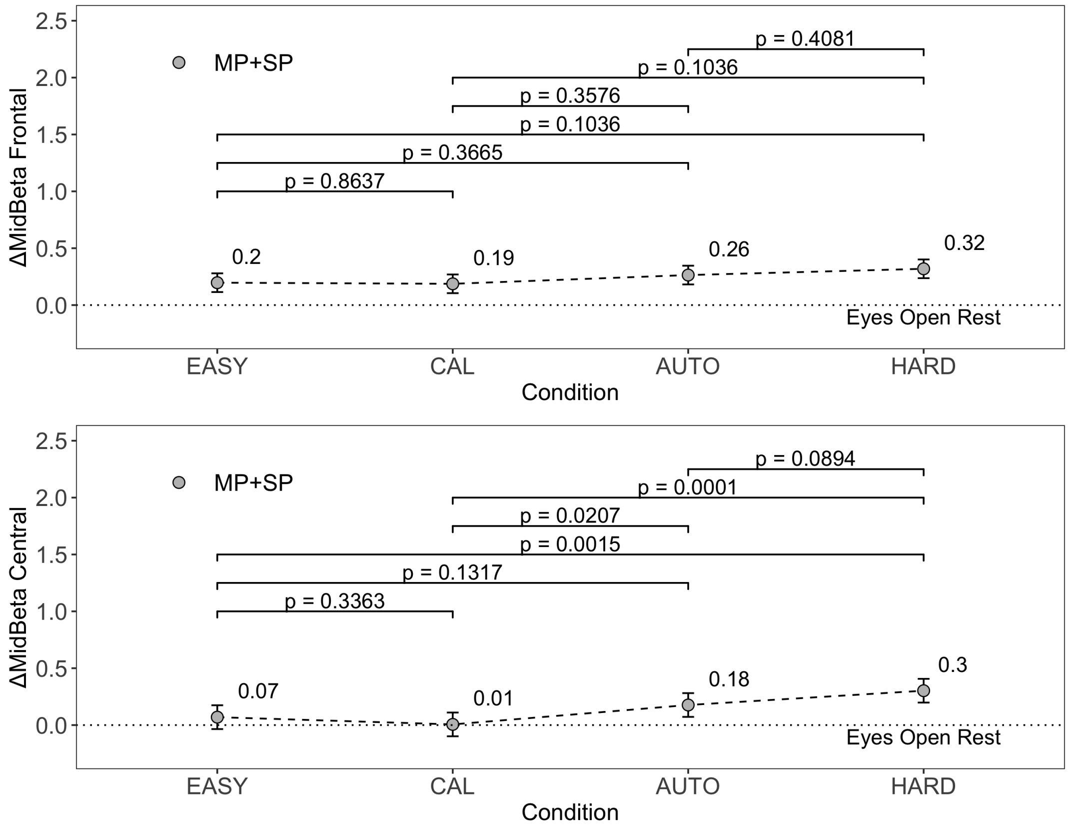

FIGURE 6.28: MidBeta Pairwise Post-Hoc Contrasts in Experiment 2. P-Values are BH-corrected. Dots Are ECMs, Error Bars One SE.

For HiBeta, all four ROI show significant main effects for social context and difficulty. Furthermore, these models did not support interactions between these experimentally manipulated variables and eliminated time and other covariates (age, gender, handedness). Also, the HiBeta models include group (level-2) random effects for all but the posterior ROI, which suggests a reciprocal influence from other group members for this particular frequency range. The results for the HiBeta LMMs are shown in Table 6.21 and Figure 6.29. For HiBeta at all ROIs, a main effect for social context is found with higher HiBeta in the MP condition. Also, for all ROIs, a main effect for difficulty is confirmed. Post-hoc contrasts (see Figure 6.30) show a stepwise increase of HiBeta with difficulty, with weaker sensitivity for the lower difficulties at central and whole scalp sites.

Metrics	HiBeta Frontal	HiBeta Central	HiBeta Posterior	HiBeta Whole Scalp
Fixed Effects				
Social Context	6.8131*	3.3005^t	4.6173*	8.5640**
Difficulty	14.6334***	24.3435***	38.4679***	25.4697***
Random Effects				
Level 1 Resid. (Indiv.)	0.7202***	0.9745***	0.8366***	0.5925***
Level 2 Resid. (Grp.)	0.0274***	0.0433***	-	0.0366***
Goodness-of-Fit				
AIC	1283.6867	1447.2835	1367.2711	1248.5461
BIC	1317.6089	1481.0482	1401.3017	1282.5150
Marginal R^2	0.064	0.067	0.094	0.094
Conditional R^2	0.680	0.662	0.686	0.666

Notes: ^{t}p <.1; *p <.05; **p <.01; ***p <.001; Fixed Effects Shown as F-Test Results; Random Effects Shown as Variance; Models Per ROI Chosen Due to Best Fit; P-Values are BH-Corrected.

TABLE 6.21: LMMs for HiBeta in Experiment 2.

It was further assessed if the difference by social context (SP vs MP) finding represents a difference amongst the samples by comparing HiBeta levels during the eyes open resting phase. One-way ANOVAs with social context as factor indicate no significant difference for Frontal (F(1) = 0.0001, p = 0.9752), Central (F(1) = 0.0181, p = 0.8932), Posterior (F(1) = 0.1085, p = 0.7424), or Whole Scalp (F(1) = 0.0045, p = 0.9463) ROIs. Therefore, the increased HiBeta levels in the MP condition can be considered a consequence of the experiment design. Together these findings indicate, that HiBeta is visible as a specific and sensitive feature, most likely related to changes in mental workload. In contrast, lower Beta ranges show no (LoBeta) or only sparse (Central MidBeta) reactivity to changes in difficulty. The robustness of the findings across ROIs that is only identified in the higher Beta ranges represents an important contribution of this work for flow EEG research, as it might allow to more robustly identify instances of (sub-)optimal workload.

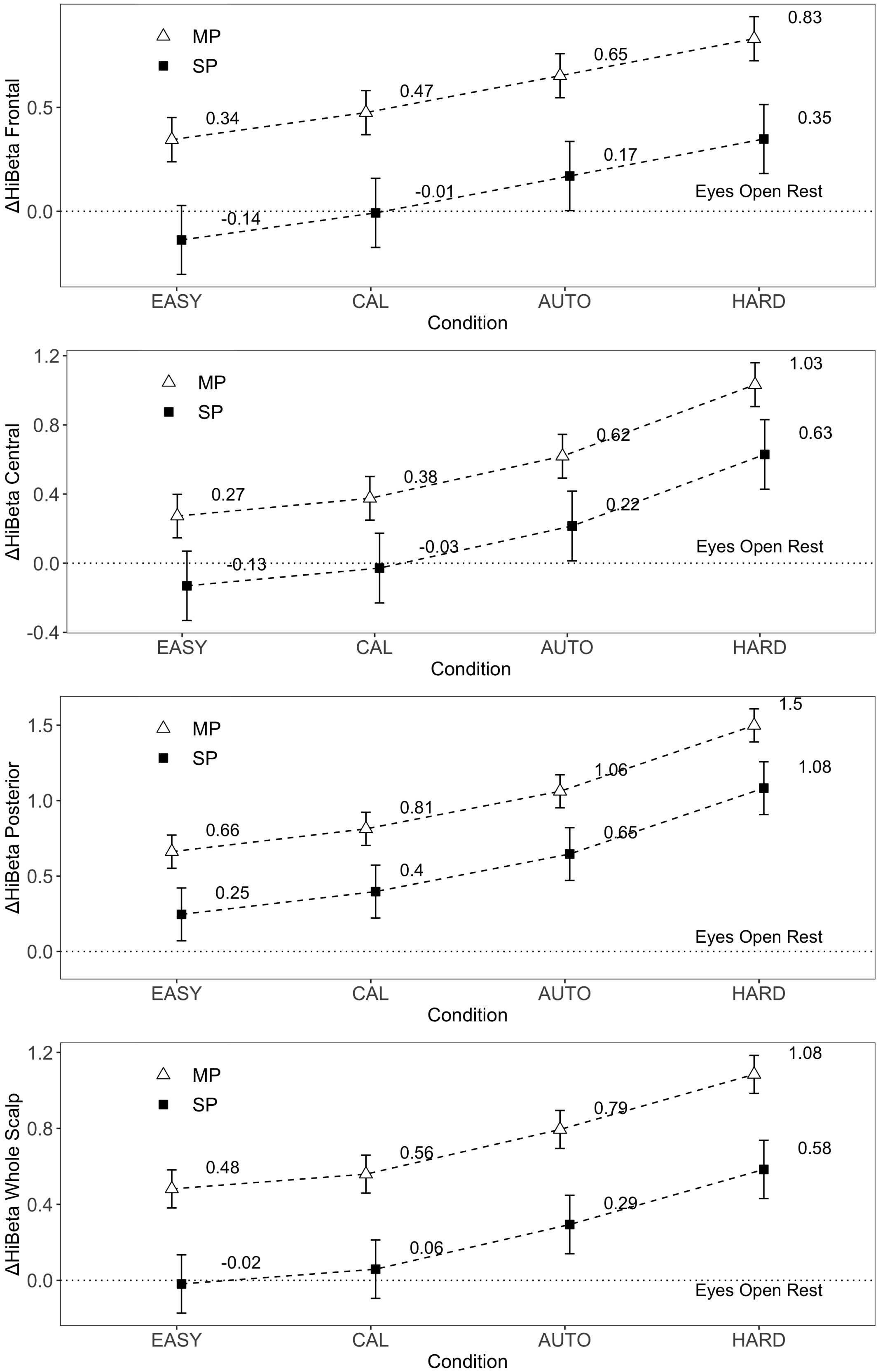

FIGURE 6.29: HiBeta ECMs in Experiment 2. P-Values are BH-corrected. Dots Are ECMs, Error Bars One SE.

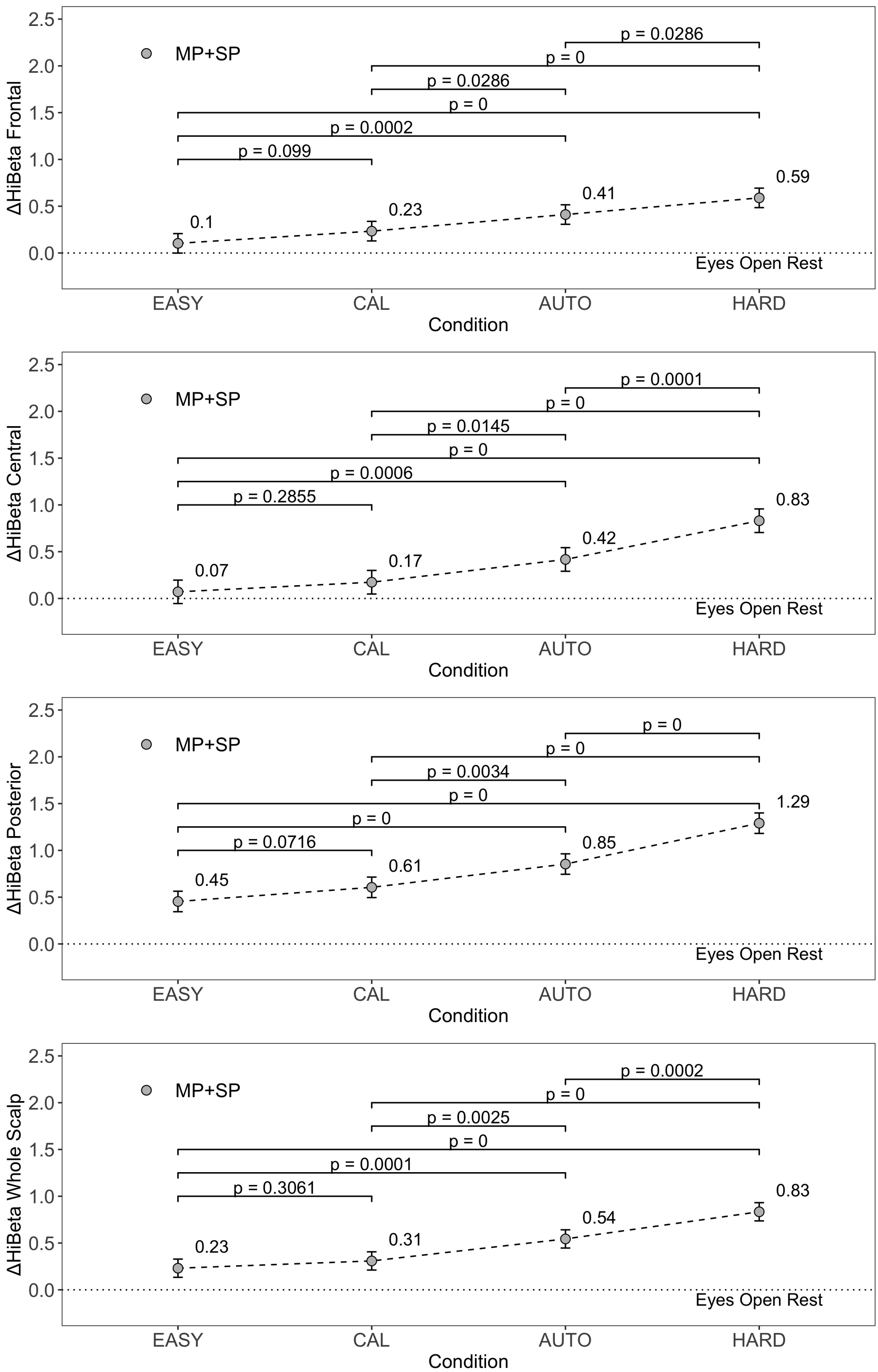

FIGURE 6.30: HiBeta Pairwise Post-Hoc Contrasts in Experiment 2. P-Values are BH-corrected. Dots Are ECMs, Error Bars One SE.

To follow up on the utility of HiBeta observations, subsequent regression analyses were conducted for the HiBeta feature (see Table 6.22). For all HiBeta ROIs, quadratic models were tested and supported the presence of non-linear relationships with perceived flow over all regions (except frontal sites - for distributions see Figure 6.31). No differences or interactions with social interaction context (SP or MP) were supported for any HiBeta-flow model.

These findings demonstrate the link between flow and HiBeta levels in an inverted-U shaped form. As HiBeta is herein understood as a sensitive indicator of mental workload, this finding is considered to integrate well with previous flow-related findings like moderately elevated frontal Theta band power (see, e.g. Fairclough et al., 2013; Ewing, Fairclough, and Gilleade, 2016; Soltész et al., 2014) or inverted-U-shaped flow relationships with physiological activation (see, e.g. Tozman et al., 2015; Peifer et al., 2014; Bian et al., 2016). Lastly, the observation of group (level-2) influences on HiBeta features is an interesting observation. No other physiological feature showed similarly robust influences from group membership. It is, therefore, possible that HiBeta features share a relationship with group influences on flow experiences.

Metrics	Flow ~ HiBeta Frontal	HiBeta Central	HiBeta Posterior	HiBeta Whole Scalp
Fixed Effects				
Intercept	4.7393***	4.7025***	4.7081***	4.7043***
	(0.0598)	(0.0526)	(0.0568)	(0.0540)
Power	-0.0693	-1.3006	-2.3347*	-2.7775**
	(0.0430)	(0.9976)	(-1.0012)	(-1.0017)
Power^2	-	-3.0002**	-2.0021*	-1.8309*
		(0.9378)	(0.9333)	(0.9318)
Random Effects				
Level 1 Resid. (Indiv.)	0.1403***	0.1373***	0.1029***	0.1187***
Level 2 Resid. (Grp.)	0.0449***	0.0198**	0.0675***	0.0377***
Goodness-of-Fit				
AIC	1322.9734	1314.8561	1321.4671	1315.6277
BIC	1343.9861	1340.0956	1346.7548	1340.8792
Marginal R^2	0.006	0.025	0.022	0.026
Conditional R^2	0.218	0.206	0.217	0.206

Notes: ^{t}p <.1; *p <.05; **p <.01; ***p <.001; Fixed Effects Shown as Beta-Coefficients with SE; Random Effects Shown as Variance; P-Values are BH-Corrected.

TABLE 6.22: HiBeta-Flow LMMs in Experiment 2.

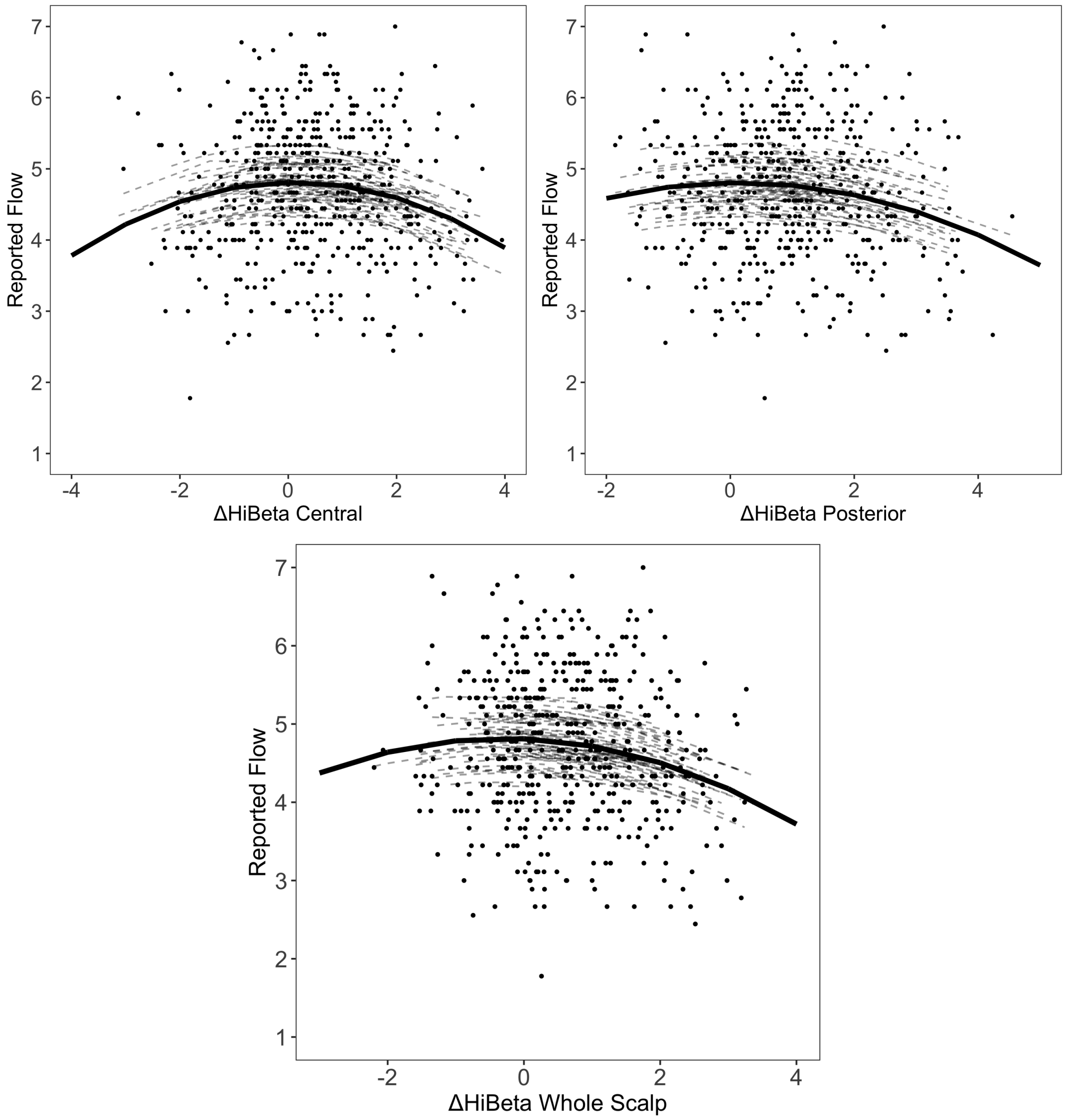

FIGURE 6.31: HiBeta-Flow Quadratic Fits in Experiment 2. The Black Line Represents the LMM Quadratic Regression Prediction. Dashed Lines Represent Predicted Values per Subject with Random Intercept.

6.4 Discussion

6.4.1 Central Findings

RQ6 & 7 - Flow Intensification

In this experiment, different flow experience intensities were elicited through a controlled, Difficulty Manipulation (DM) experiment design. Furthermore, additional manipulations of autonomy (AM) and social context (SCM) were included in the design. In doing so, as one of the first studies, a fully digitally-mediated cooperation scenario was used for the MP condition. These additions that are integrated with a previously validated task and paradigm represent a particular contribution of this work. In relation, Research Question (RQ) 6 asked if flow elicitation in the laboratory is intensified by increased task autonomy (RQ6.a) and by performing a task in small groups (RQ6.b), and RQ7 asked if the flow elicitation using a mental arithmetic DM task is extensible to social interaction settings. To this regard, the presented results reveal intriguing findings that require critical review and discussion.

Throughout the variables for subjective or objective difficulty, the designated manipulation of difficulties into low, balanced, and high difficulties was supported in both the SP and the MP conditions. Beyond this pattern, slight interactions are found for each variable, like the lower perceived (and less optimal) difficulty in the balanced difficulty MP conditions, and the opposite pattern in the objective difficulty variable of nr. of correct trials (fewer correct in EASY and more correct in HARD in the MP condition). These findings highlight the challenge of creating fully comparable conditions in terms of DM for SCM paradigms. The nature of the small group setting creates interaction effects, that must be considered in the interpretations of results. In the present form, they appear to most readily aid in explaining the weaker contrasts for flow experiences in the MP condition. Therefore, a takeaway from the presented approach is, that contrasts in creating easy, optimal, and hard task can be improved for MP settings. Especially towards the intensification of flow, it would appear that alternate mechanisms to create optimally balanced group difficulties ought to be employed. This shortcoming means that RQ7 is answered by partial affirmation, that the herein presented mental arithmetic task appears to be suitable for the extension for social flow research, yet that some refinement is further required. For the SP condition, one approach that appeared to provide desirable results in this direction was the inclusion of the AM manipulation, that was indicated as successful by the increase in perceived autonomy from CAL to AUTO. However, no such effect was found for the MP condition. Furthermore, such an effect did also not become apparent, even in situations when the range of preferred difficulties within groups was rather small. This finding indicates that the manipulation of autonomy in small group settings requires alternate mechanisms. Presently, it would appear that the approach of median preferred difficulty selection does not sufficiently integrate individual preferences. Instead, an approach could be to aggregate individual preferences more in the form of a summation instead of an averaging. This idea means that, for example, by providing sub-tasks in the preferred difficulty of each individual, every member's preference should be included more clearly. The third manipulation, namely SCM is considered to have been successful by task design and by descriptive reports. Participants could not complete the tasks in the MP condition by themselves and indicated an elevated level of perceived interdependence on average. Yet, besides the manipulation having occurred, its consequences for flow experience intensities were in contrast to expectation.

First of all, the pattern of flow experience, in particular in the MP condition, was not expected in this form. While the success of DM was confirmed, the SCM did not lead to intensified flow. In much of the previous work on flow experiences in social interaction, the empiric tenor has been that social flow experience is intensified or facilitated (see, e.g. Magyaródi and Oláh, 2017; Tse et al., 2016; Walker, 2010). However, in the present results, not only, does flow in groups appear to be at least on par with the solitary intensities, but also in some instances is lower than in individuals working alone, which is the first finding of this type. This effect may be to some degree caused by less optimal difficulties in the MP compared to the SP condition. Previous work which has found inverted patterns might have found more intense flow in social interaction due to inverse difficulty patterns (i.e., a task for individuals might have been slightly too hard and for groups have had just the right difficulty). Yet, it can be argued that also in the MP condition, two moderate difficulty levels were instantiated (difficulty increased from CAL to AUTO), which makes it unlikely that the lower flow intensities in MP are but the result of inadequate difficulty calibration. Given the two major novel elements in the present experiment design (the optimal difficulty self-selection condition and the digitally mediated environment), two other central explanations for these results were followed-up on in exploratory analyses.

The first explanation for the unexpectedly similar or lower flow experiences in groups could be rooted in an inversion of opportunities for optimal difficulty selection (i.e. contrary to previous work, in isolation/groups, participants had more/less flexibility in difficulty selection). It has been stated that intense flow might be difficult to elicit in laboratory setups, particularly due to a variance in preferences for optimal difficulty (i.e. some people prefer slight underload, some prefer slight overload - see Fong, Zaleski, and Leach, 2015; Tse et al., 2016). This variance is often not addressed through preference insensitive difficulty calibration algorithms (like the one used in this and most related work - see Keller, 2016; Ulrich et al., 2014). This problem means that through a one-size-fits-all difficulty calibration (that personalises difficulty based on performance only), the potentiality for flow intensification is systematically limited. In contrast, self-selected difficulty approaches have been found to elicit more intense flow in solitary tasks (Barros et al., 2018). From the present data, it similarly appears that the AUTO SP condition has elicited the most intense flow experience. The same effect might have further reduced flow potentials in the MP condition. The follow-up analyses on flow, autonomy, and optimal difficulty support this consideration, as significant positive relationships were found between perceived flow with autonomy and autonomy with optimal difficulty. This results suggests that increased levels of autonomy could be related to better calibration of difficulties which are theoretically grounded as preconditions for flow intensity (Nakamura and Csikszentmihalyi, 2009). Notably, the slopes for these relationships varied with social context and showed a stronger relationship amongst the variables in the SP. A possible explanation for this finding is that in the small group interaction, higher levels of individual autonomy must not directly lead to higher levels of flow, most likely because other members actions are still likely to interfere with flow emergence. These results support the idea, that previous findings of intensified flow in groups could be largely driven (i.e. moderated) by degrees of freedom and autonomy - not just social interaction. This finding provides a new rationale and starting point to differentiate further what might drive flow experience intensities in the laboratory and social interaction.

Another central difference in the experiment design might have resulted in an inhibition of social context manipulation-based flow intensification, namely the social interaction format. Having opted for a digitally-mediated cooperation scenario, the interaction possibilities in the present design have been reduced to task-related action signalling only (i.e. participants could only show on which part of the problem they were currently working). This design element is a significant difference to previous work that has almost exclusively opted for face-to-face interaction settings (see Table 3.1 in Section 3.3). It could be possible, that not just any type of social interaction will do similarly well to facilitate flow and could even be the limiting factor to it. This possibility falls in line with the general understanding that digital media can be altering and limiting the “normal” (i.e. known from face-to-face environments) exchange of socio-affective communication (Chanel and Mühl, 2015; Derks, Fischer, and Bos, 2008). To further explore the possibility of this effect, a relationship between perceptions of social presence and the average level of flow experiences was compared. While the overall level of perceived social presence is rather low, only weak support was found for a relationship between perceived social presence and flow in this experiment. Therefore, the possibility that social signals influence the intensity of flow is not receiving proper support. However, these results are limited to the extent as to how they can inform on such a hypothesis. First of all, the level of perceived social presence can only represent a proxy for the wide array of influential social signals and processes. Second of all, as the present study did not directly compare a Face-to-face (F2F) to a Digitally-Mediated Communication (DMC) scenario, nor manipulated social signals in any form, conclusions on these relationships require separate experiments.

To follow up on the relationships of flow with stress that was identified in Experiment 1, dedicated follow-up analyses were conducted. To reiterate, in Experiment 1, flow was found to be both accompanied by moderate and low stress levels. This observation prompted the question as to whether the experienced flow in the writing task is qualitatively different, namely a flow experience without stress perceptions that would appear to be more in line with theoretical predictions (e.g. the concept of flow as a unique experience of effortlessness - see, e.g. Bruya, 2010; Ullén et al., 2010). The results showed a significant quadratic relationship between stress and flow, that indicates an increasingly declining flow experience with higher stress perceptions. Therefore, it can be stated that higher intensities of flow occur with lower perceptions of stress. However, the incompatibility of stress and flow increases mostly with higher levels of stress. On the one hand, these results suggest that with higher flow, stress is rather reduced, which together with the findings from Experiment 1, could indicate a shortcoming of present self-report instruments to capture the stress-absence (effortlessness) facet of the flow experience. Yet, this finding could also be seen as further support for the alternative proposition that flow and stress are partially unrelated (at least at the lower end of stress perceptions). Theoretic integrations proposed by Tozman and Peifer (2016) would account for this effect by explaining, that flow and stress become only incompatible when a task becomes too difficult, such that self-evaluative threats emerge. To further elaborate on what the actual pattern is, more data from naturalistic settings and laboratory experiments will have to be acquired. Nevertheless, the results highlight an important aspect for further development of the situational assessment of flow. Inclusion of stress or effortlessness dimensions into future flow self-report constructs could provide a valuable addition to assess the internal validity of flow measurements (the quality of experienced flow). An additional aspect to the questions about flow quality from Experiment 1 was the observation of high flow with lower stress in writing and a concomitant increase in physiological activation. This pattern suggested, that while flow was present and perceived as stress-/effortless, individuals still exerted a higher amount of energy, which further added to the possibility that high flow is perceived as effortless yet shows increased physiological effort. These patterns have been similarly observed by (Harris, Vine, and Wilson, 2017a). The observations on neurophysiological patterns of flow are discussed in the next section.

Together the results from manipulation checks and reported flow intensities provide a two-sided answer for RQ6. RQ6.a can be affirmed as flow was found to be intensified through the inclusion of higher levels of autonomy. RQ6.b can be negated as flow was not found to be intensified through social interaction in digitally-mediated environments.

RQ8 - Cross-Situation Flow Neurophysiology

In the analyses of neurophysiological features, interesting findings emerged in relation to flow experiences that ought to be discussed. It should first of all be noted that the results are still strongly bound to the DM paradigm and a fairly restrictive laboratory setup. Nevertheless, the additional inclusions of AM and SCM manipulations in this experiment provide additional evidence and increased external validity when compared to most previous work that focused on flow experience in isolation in DM paradigms. Especially as the AM manipulation realised at least some level of intensification of in-laboratory induced flow, the data come with the potential to further deepen and consolidate previous findings from more naturalistic settings (e.g. Experiment 1). Thus, the data provide the potential to follow up on some of the salient hypotheses on flow neurophysiology (e.g. the level of calming physiological influences, or the degrees of frontal downregulation). The related RQ8, therefore, asked which correlates of flow can be identified amongst different knowledge work and is reviewed in the following paragraphs for Peripheral Nervous System (PNS) and Central Nervous System (CNS) feature spaces.

Throughout ECG and EEG features, even with refined approaches, no features were found that show variation with reported flow in the form of a highly diagnostic, one-to-one relationship. Given that related work has also not uncovered such marker-features, the present results suggest, that either such markers are not identifiable with the present feature space, or that flow elicitation effects are still too weak to allow for neurophysiological contrasts in this feature space to emerge. Instead, however, the amalgamation of findings on neurophysiological parameters (e.g. for the PNS and CNS), does serve to describe a coarser configuration picture of the neurophysiological state during flow. From the present data, it can be summarised that flow appears to be represented by moderate physiological activation (moderate HRV) and mental workload (moderate HiBeta power - and tentatively elevated frontal Theta power), and by increased attentional engagement (reduced frontal Alpha). In addition, flow appears to be represented by an absence of variation in approach-avoidance motivation or affective valence (as indicated by the absence of frontal alpha asymmetry changes - see Harmon-Jones, Gable, and Peterson, 2010 for the foundation of this interpretation). Altogether, these findings provide detailed answers to RQ8 and are in its subtleties discussed in the next paragraphs.

The HRV related findings suggest, that flow is not represented by a level of high physiological activation similar to stress, as has been previously documented in some research (Experiment 1 and Keller et al., 2011). Also, the findings further raise questions as to whether or not a strong calming influence is present and whether a configuration like non-reciprocal co-activation is present in the body. However, in the absence of sympathetic activity markers, the change in the interaction dynamics of the two Autonomous Nervous System (ANS) branches cannot be sufficiently assessed and still requires further research. For the HRV features in this experiment, two peculiar observations need to be discussed further. First, a lack of sensitivity to the difficulty manipulation was present, that is irregular in comparison to related work. Related work has found a high level of variance in the sensitivity of parasympathetic

HRV features to difficulty manipulations across two game tasks (Barros et al., 2018). Therefore, the lack of feature sensitivity in this experiment might be a result of the chosen mental arithmetic task. Further cross-task experiments are required to clarify this issue. Second, a significant difference in HRV levels across the SP and MP condition was visible, that was not expected in this form. Two explanations are considered for the lower physiological demand (higher HRV) in the MP condition. On the one hand, the task could have been less hard in MP in general. Yet, no other metrics would indicate such an effect (subjective and objective task difficulty measures). Furthermore, the herein found more highly diagnostic markers for mental workload (specifically HiBeta EEG power) point in an opposing direction. Therefore, a more likely explanation for the HRV differences across the social contexts (despite the absence of general differences in flow, stress, or task difficulty) could be found in the duration of breaks during the experiment. In the MP condition, participants often had to wait on other group members to complete surveys to move on to the next difficulty condition as a group. This waiting time might have lead to an overall higher level of physiological relaxation as compared to the SP condition. This interpretation is supported by the observation that time has a significant and strong positive relationship with the HRV features.

The frontal Theta and higher Beta findings over the whole scalp suggest that flow is related to moderate levels of mental workload. In these features, several important findings and contributions to flow research ought to be discussed. First of all, an initially surprising finding was the absence or weak sensitivity of one the more robust workload indicators that has previously emerged in flow EEG research, namely frontal Theta. In the SP condition, a pattern of increased frontal Theta from EASY is visible, yet without significant differences. While this means that there is no statistical support for a relationship between frontal Theta and flow, the pattern is similar to related work that has found elevated frontal Theta levels during flow (see, e.g. Fairclough et al., 2013; Ewing, Fairclough, and Gilleade, 2016; Soltész et al., 2014). It is only for this reason that the elevated frontal Theta effect is considered to be tentatively visible, with increased task difficulty. However, the effect is too small to reach significance, which is likely due to the absence of electrodes over midline positions (e.g. Fz) over which the frontal Theta effect is typically observed (see, e.g. Borghini et al., 2014; Ewing, Fairclough, and Gilleade, 2016). Conversely, frontal Theta effects have also been found to be absent in Difficulty Manipulation (DM) experiments that have also used the Emotiv Epoc+ EEG headset (Klarkowski, 2017). Beyond this initial limitation, a second peculiarity for the frontal Theta band variation was found, namely the divergence of Theta levels with social context (elevated from EASY in SP vs high and reduced in HARD in MP). Given the absence of small group EEG studies in related work, the explanation of this pattern is a bit more difficult to explain. Besides, another concomitant of mental workload (see Beta power below) does not show a similar reduction towards HARD (yet a generally elevated level in MP, that might reflect a generally higher workload in MP). As a possible explanation for these patterns, the specificity of Theta power is considered. In more fundamental neuroscientific work, both Theta and Beta power have shown reactivity to increases in task difficulty induced mental workload (Kahana et al., 1999; Deiber et al., 2007; Michels et al., 2010). However, frontal Theta has additionally been found to be influenced by task complexity (Kahana et al., 1999) and by (working memory-independent) attentional engagement to visual stimuli (Deiber et al., 2007). In this experiment, the MP condition does not only provide a higher level of visual stimulation in general (signals by other group members - see Figure 6.2) but also the requirement to attend to other group member's actions (due to incomplete task information). This added attentional demand in the MP condition is considered to have caused the (tentatively) elevated frontal Theta levels in the MP condition. The reduction of frontal Theta in MP HARD might then furthermore be explained by a reduction of attentional complexity towards the high task difficulty when it becomes harder to monitor the actions of others. Stated differently, in the high task difficulty, group monitoring might be given up to some degree in favour of focusing on one's task that is already very difficult. In support of the latter argument it can be said that even in this MP HARD condition, the frontal Theta levels are still reasonably elevated when compared to all SP Theta levels, which indicates an elevated level of mental workload.

Beyond the (weakly diagnostic) frontal Theta findings, another feature set indicated much clearer reactivity to the manipulations of difficulty and social context, namely Beta power over the whole scalp. Isolated by the frequency band sub-segmentation, in particular, the higher Beta ranges showed an apparent reactivity to the manipulation of difficulty in both SP and MP scenarios. As the separation of Beta bands has in neuroscientific work been found to emphasise mental workload changes in the higher frequency ranges (Michels et al., 2010), and as positive Beta band relationships with higher workloads have been reported often (Michels et al., 2010), the present results indicate a high sensitivity of HiBeta powers to changes in difficulty. Given the elevated HiBeta levels in the MP condition, the HiBeta range is considered to show high specificity to mental workload changes. The latter is based on the findings that workload is considered to increase with higher task complexity (Kahana et al., 1999; Deiber et al., 2007) (which is here considered to be the case). The additional group interaction is, therefore, likely driving the HiBeta increases throughout the MP condition. These considerations are in line with the previously discussed divergence of frontal Theta levels between SP and MP conditions. Given that HiBeta was found to be unconfounded by time and visible over the whole scalp, it currently represents one of the most useful EEG features for the observation of flow in this experiment. The observation that it appears to more clearly appear after segmentation of the broad Beta band (LoBeta and MidBeta showed almost no reaction to the experimental manipulation) makes the methodological approach a valuable contribution to flow research. The identification of a direct non-linear relationship in this region has not been reported before. However, some research has already found a relationship of higher frequency ranges with flow (Soltész et al., 2014 finds Gamma band relationships). Therefore, future work might want to more closely explore how to leverage the higher frequency ranges for flow experience observation. Presently, they would seem to appear as suitable indicators of (non-)optimal workload situations that can be used across situations and repeated measurements. Moreover, it should be noted that the HiBeta models include group (level-2) random effects for all but the posterior Region of Interest (ROI), which suggests a reciprocal influence on mental workload from other group members. This observation is interesting in itself, as such an effect has not been documented in previous flow research. Given the indicated sensitivity and specificity of the HiBeta band to changes in mental workload, it is possible, that HiBeta features might be related to group influences on flow experiences. Potentially, they could be indicating a situation of shared mental workload during shared flow. This possibility is further explored in the following Chapter 7.

The last set of findings relates to attentional engagement and frontal downregulation, that is, the frontal Alpha power patterns. The Alpha results suggest together that during flow, no general frontal downregulation is present. Instead, neural activity in frontal regions close to the midline (and not spread out further) is increased. This increased activation is likely related to attentional engagement towards the task (see Klimesch, 1999; Deiber et al., 2007; Borghini et al., 2014). While this is a known observation from other neuroscience experiments, the results further indicate that in situations with balanced (optimal) difficulty, no peculiar Alpha changes emerge that might be related to the flow experience. The identification of the topographical restriction of Alpha activity is useful to note here, as the observation of larger areas might have lead researchers to different conclusions. The segmentation of the Alpha band was in this instance not found to be of relevance as the observed effects are similarly found in both Lo- and Hi-Alpha sub-bands. Yet, the confidence levels show that for HiAlpha, the effect appears more pronounced. This result means that the previous findings from Experiment 1 (i.e. the absence of main effects from difficulty on the LoAlpha band), (and some related work - see, e.g. Ewing, Fairclough, and Gilleade, 2016) might also be artefacts, for example resulting from the utilisation of non-personalised frequency ranges. Overall, the present observations integrate with the findings from Experiment 1 that tasks with increased difficulty show a reduction (that was also more stable) of frontal Alpha. The requirement for a refined downregulation hypothesis in the Transient Hypofrontality Theory (THT), therefore, gains further support based on EEG-methodology as well. However, at least with the present EEG system, a much too low spatial resolution is available that would be needed to further integrate findings from hemodynamic studies on medial and dorsal Prefrontal Cortex (PFC) activation with EEG Alpha observation. Using co-registered EEG and fMRI, for example with this mental arithmetic DM paradigm (previously used in an fMRI study by Ulrich et al., 2014; Ulrich, Keller, and Grön, 2016b) could allow providing more insight here. Also, it would be interesting to investigate how source-localisation methods and spatial filters from high-density EEG recordings might lead to a better understanding of frontal activation patterns during flow. Presently, frontal Alpha can only be used in the form of identifying flow boundary conditions, specifically as it might lend itself to identify situations of low top-down attentional engagement towards a task (instances of mind-wandering). In a final note on frontal Alpha patterns, it should be noted, that no changes in Frontal Alpha Asymmetry (FAA) were detected with flow in the present experiment. This finding is intriguing since flow has been considered to be related to FAA because it is often considered to be related to approach-avoidance motivation (Smith et al., 2017; Harmon-Jones, Gable, and Peterson, 2010). The present findings (including the absence of such effects in Experiment 1) indicate that flow is more likely a state of affective neutrality, likely explained by the absence of self-monitoring and self--evaluative processes (see, e.g. Sadlo, 2016; Harris, Vine, and Wilson, 2017b). Such interpretations of affective neutrality have been put forward by psychologists (see, e.g. Engeser and Schiepe-Tiska, 2012), yet are contested by those who argue that flow is an experience of strong positive valence (see, e.g. Mauri et al., 2011; Manzano et al., 2010). Convergent with the latter view, some scholars have already suggested that flow could be related to FAA (Labonté-Lemoyne et al., 2016). Yet, the present results are the first reported evidence that refutes this suggestion. It is possible, however, that the findings indicate merely, that during the task, flow is not affectively valued, but that such connotations arise after the task is completed and self-evaluative processes emerge again. To study these temporal dynamics of FAA (see Allen and Cohen, 2010) could, therefore, be an interesting opportunity for flow EEG research.

6.4.2 Limitations & Future Directions

A few general and specific limitations to the presented experiment need to be appraised. General limitations of this work are the use of a German student sample (with a majority of business administration and engineering study majors - >45% of participants), the limited number of observations (no more than 40 observations on the individual and group level), the limited signal availability and quality from the wearable EEG devices, and the large number of statistical tests that were performed. To counter these limitations, it can be argued, that flow is found to be a universal experience (Nakamura and Csikszentmihalyi, 2009) that spans across cultures, ages, and professions. Therefore, the student sample might be quite representative of this particular kind of work. Sample size recommendations would typically lead to higher numbers of desired participants, it can also be argued that the presented efforts represents an unusually large account in relation to how many neuroscientific studies are conducted (see Turner et al., 2018; Szucs and Ioannidis, 2017), and in relation to other related work that has (except for a few studies) used smaller samples (see Table 4.2). As for the EEG sensor quality, it can further be argued that while there are certainly improvements possible in terms of signal quality, the apparatus was selected as it was found to deliver acceptable quality for power frequency analyses (Barham et al., 2017). Regarding the number of statistical tests, the issue was recognised by p-value corrections throughout the analyses to reduce false-positive errors.

Beyond these rather typical experiment limitations, some that are specific to the experiment ought to also outlined. Given the artificial laboratory environment and the focus on just one task setting, the internal and external validity of the presented results is limited. As was previously discussed (Experiment 1) and has been found in related work (Barros et al., 2018), multi-task investigations certainly deliver some variation in perceived experiences and physiological patterns. Therefore, it is critical to investigate further how robust some of the presented results are, when observed in different tasks (e.g. that require different abilities), in different task formats (e.g. that are less constrained or more tailored towards mastery), and for the social context in different interaction settings (e.g. with more social information, more autonomy for the individual, or different group sizes). However, it can also be appraised that in the presented experiment, there are already multiple different manipulations involved, which is why at least some additional validity (compared to a pure DM approach) could be attested. Improvements to the experiment designs are considerable specifically for the MP condition for both DM and AM manipulations. Given that effect sizes for the manipulated flow experiences in the MP could be improved, in particular, the mechanisms for optimal difficulty setting should be refined. On the one hand, the calibrated difficulties might, for example, benefit from seeing a higher starting level or faster level increment, as the level of perceived (and optimal) difficulties were found below the SP condition reports. On the other hand, especially the AM condition should see dedicated refinements, especially as results showed the positive link between autonomy and flow that did not seem to be leveraged in the MP condition. As an adjustment, the self-selected individual optimal difficulty could be used to distribute equations based on individual preferences. In this approach, each individual would get an equation that fits their indicated preference. Alternatively, an extension of the present research could increase the complexity and freedom in the math task in general by providing more subtasks (e.g. six equations with different lengths) and longer trials (e.g. 60s), so that the self-selection possibilities are returned during an MP condition. Both design aspects would re-introduce self-selection towards more and less difficult subtasks and could, in comparison elucidate whether the intensified flow experience in social interaction is more due to an optimised sharing and allocation of difficulty-balanced task aspects. Lastly, the comparison of the social context settings with the present data would be a valuable endeavour to understand further what does or does not intensify flow in individual and group settings. The present findings do provide a useful starting point to further differentiate if and which types of signal exchange (task, affect or interpersonally-related) might be required to reap benefits of intensified flow in small group interactions. Future work could, therefore, start by manipulating the availability of interaction channels, while using the otherwise unchanged experiment design outlined (i.e. extending the present experiment to include audio- or visuo-auditive communication). In the simplest and most pressing form, a comparison of the digitally-mediated MP condition to a face-to-face setting will have to be conducted.

The second, important limitation specific to this experiment is the selection and processing of measurement instruments, both reported and physiological. For the reported measures, a limitation emerged by the selection of a short flow survey instrument. Specifically, as some findings point to the importance of understanding more nuanced facets of flow (e.g. how stress- or effortful it is), the utilisation of short is likely insufficient to develop an elaborate understanding. Future work could benefit strongly from leveraging more complex and multidimensional survey instruments (e.g. the FSS - see Jackson, Martin, and Eklund, 2008) that will provide more insight into perceived experiences. In the physiological domain, signal processing techniques can be improved further. For example, the segmentation of higher (Beta) frequencies occurred based on decisions from related work (Berta et al., 2013), not from physiological criteria (as they are for example available for the individual Alpha frequency - see Klimesch, 1999). Therefore, spectral filters might help in isolating better, which ranges provide useful information on flow and related processes. In another example, the ROI for the EEG analyses were primarily selected from theoretical arguments and previous work. However, this method does not protect from some of the inter-individual variation in the emergence of observed patterns over the scalp. Spatial filters that constitute a significant advancement in EEG research (Blankertz et al., 2016) would represent a useful advancement to understand further where flow-related effects are best observed for each individual. The importance of using such refined approaches in the future is exemplified by work that has been able to predict the difficulty level of a game task from the EEG data alone (Naumann et al., 2016). To achieve such sensitivity for the extracted EEG data could be very valuable not only in terms of practical uses (being able to observe with a much higher level of certainty if a task is too easy or too difficult for an individual) but also in terms of how theoretical knowledge is advanced. Having such high sensitivity could mean, for instance, that one could observe experts performers or people with high flow proneness or metacognitions in Mastery (MAS) paradigms (see, e.g. De Kock, 2014; Kramer, 2007) to understand better what the flow preconditions and configurations are on a neural level.

Lastly, the previous remarks on feature extraction processes highlight the specific limitation in the presented work, that primarily a priori decision-making was involved together with traditional methods of statistical inference. This approach grounds the developed knowledge in scientific best practices. Nevertheless, especially in the field of psychophysiology, it can be argued that such deterministic approaches are generally flawed when attempting to isolate patterns across domains of reported and physiological data (Bridwell et al., 2018). In principle, this is the argument that there is no reason as to why concepts from psychology textbooks should directly match on to neurophysiological data (see Section 4.1). Therefore, it is a critical limitation of the present approach to neglect data-driven approaches to identify neurophysiological patterns during flow, which has been argued for in related work (Rissler et al., 2018; Maier et al., 2019). In general, there is an increasing prevalence of using data-driven methods for feature extraction, selection and integration into classification models for the observation of mental states (Brouwer et al., 2015; Bridwell et al., 2018; Roy et al., 2019). Deep Learning methods have, for example, shown great promise to make sense of EEG signals due to their capacity to learn good feature representations from raw EEG data (Roy et al., 2019). Given that presently, a high degree of a priori decision making is involved in the analytical process in most flow neurophysiology research, more data-driven approaches, might hold the valuable potential to improve the validity and accuracy of derived physiological features and should be employed in future work - if only for feature extraction and selection.

6.5 Conclusion

The results from this experiment contribute in theoretical and practical ways to a foundation for the theory-driven development of adaptive NeuroIS. First, flow theory is advanced by testing whether or not difficulty, autonomy, and social context manipulations elicit varied flow intensities in the context of digitally mediated cooperation. In doing so, mechanisms for flow elicitation are confirmed and extended. The unexpected finding of lower flow in social interaction identifies the need for further refinement of the conceptual integration of the relationships of optimal difficulties, autonomy, and social interaction within flow theory. In particular, two important new directions for research are identified, that pertain to the questions of opportunities and autonomy as a driving factor in intensified flow in the laboratory and to the question of how digital media could impact flow experiences in groups through the exclusion of open communication and social signals. For both directions, simple alterations to the present design are outlined (inclusion of more subtasks and summative integration of member preferences for the former and comparison with face-to-face settings for the latter), that can further deepen the understanding on how to elicit more intense flow in individuals and groups.

Furthermore, the refined analyses of neurophysiological data further consolidate the empirical knowledge on how flow can be described through changes in the heart and the brain. Altogether, from the present data, it can be summarised that flow appears to be represented by moderate physiological activation (moderate HRV) and mental workload (moderate HiBeta power - and tentatively elevated frontal Theta power), and by increased attentional engagement (reduced frontal Alpha). In addition, flow appears to be represented by an absence of variation in approach-avoidance motivation or affective valence (as indicated by the absence of FAA changes). Importantly, these results emerge through the inclusion of various mechanisms for the elicitation of flow experiences in the laboratory (DM, AM, and SCM), which represents the major contribution of this work to the flow neurophysiology literature. Of particular relevance is the finding that through frequency band personalisation and sub-segmentation, some previous findings could be consolidated (specifically, frontal Alpha reduction), and some promising new directions emerged. Specifically, the frequency band segmentation highlighted the particular sensitivity of the HiBeta frequency ranges with manipulations of difficulty. The additional absence of confounds with time, and the group level influence on HiBeta levels, further indicate that these higher frequency ranges could have a valuable role for the observation of flow on the individual and group level.

The presented results need further confirmation through additional experiments with varied tasks and task formats. In addition, future work should employ a set of more refined data-driven feature extraction and selection methods. Presently, it is primarily argued that the observed patterns allow discussing flow-related changes in a refined manner and that they pose interesting alternatives for the detection of situations of optimal difficulty. These alternatives are especially highlighted for scenarios in which less information might be available than typically is in laboratory setups (i.e. with wearable EEG with fewer and unevenly distributed electrodes). When considering how adaptive NeuroIS employ thresholds to inform adaptation rules (see, e.g. Ewing, Fairclough, and Gilleade, 2016; Karran et al., 2019), features indicating maxima during EASY (e.g. frontal Alpha power) or HARD conditions (e.g. HiBeta power) could be valuable boundary condition indicators. In this regard they may be used to robustly identify when difficulty is unbalanced and flow unlikely.

Ideally by finer spectral and spatial EEG power analysis, future flow research will move even closer to identifying robust concomitants and markers of flow that can be employed in adaptive NeuroIS using portable EEG in real-world scenarios. So far the identification of sensitive features has implications for the development of flow prediction and classification models (see e.g. Rissler et al., 2018; Maier et al., 2019), as they describe sparse feature spaces that might allow to observe flow. For example, especially given the herein found usefulness of HiBeta power, EEG observation from a few select sites becomes an interesting option, at least when the goal is to observe situations of (non-) optimal difficulty with few electrodes and without confounds from time and task context. Especially since some of the more robust findings emerged from central ROI, an interesting direction would be to explore the potential of flow state prediction using ear-EEG systems (see e.g. Bleichner, Kidmose, and Voix, 2020). Such systems have recently been shown to allow mental workload classification in driving simulations (Wascher et al., 2019) and might therefore provide an unobtrusive approach to flow boundary condition detection as well. Eventually, systems able to adapt to flow intensities could reduce flow interruptions (e.g. by blocking incoming messages - see, e.g. Rissler et al., 2018) or provide feedback for flow self-regulation (e.g. by optimising arousal levels and catalysing task focus through EEG-neurofeedback - see Lux et al., 2018).

In conclusion, in this chapter, valuable contributions to flow theory and adaptive NeuroIS development have been presented through the advancement of the understanding of individual-level flow experiences in isolation and social interaction. Given the complexity of social interactions, In the following chapter, a dedicated investigation of the group-level flow dynamics is pursued to complete the study of social flow experiences.

Chapter 7

Experiment 2 - Group-Level Flow Dynamics

Contents of this section are in part adopted or taken from Knierim et al. (2019). See Section A.1 for further details.

7.1 Flow Experiences in Small Groups

To provide foundations for flow-facilitating, adaptive NeuroIS, the work in this dissertation focuses on studying flow in primarily cognitive tasks. Thereby, it is investigated how flow could be intensified in experimental research (RG2). In this approach, flow neurophysiology research converges on more Knowledge Work (KW) related scenarios and refines how neurophysiological processes related to flow can be described across situations using wearable sensor systems (RG3).

Experiment 1 and 2 showed that varying flow intensities could be elicited in laboratory research by using various manipulations. In particular, in Experiment 2, findings from Experiment 1 were integrated into alternative research designs, in particular in the form of providing increased autonomy by allowing participants to self-select an optimally challenging difficulty level for the mental arithmetic task. Furthermore, another potential approach for flow intensification in the laboratory was explored, namely, the comparison of completing a task in isolation or during social interaction. In the results for Experiment 2, only autonomy, but not social context was found as a factor that elicits deeper flow experiences. However, an interesting observation was made for the neurophysiological data. HiBeta EEG features showed group (level-2) random effects for all but the posterior Region of Interest (ROI) that are similar to the random effects structure for reported flow. These findings indicate a reciprocal influence on mental workload from other group members, that could be related to the reciprocal influences of groups on their members' flow experiences. Such an effect has not been documented in previous flow research. As already mentioned, the study of flow in social interaction has in recent years been of highlighted importance for flow scholars (Walker, 2010; Magyaródi and Oláh, 2017; Hout, Davis, and Weggeman, 2018; Tse et al., 2016). This importance is mainly due to the increased reliance on small groups to complete complex KW tasks. To follow up on the neurophysiological findings and to further explore additional compelling social interaction flow dynamics in Digitally-Mediated Communication (DMC), a series of additional analyses was pursued in this chapter.

Flow theory was initially developed with a focus on the experience of the individual (Csikszentmihalyi, 1975). One of the leading questions in Csikszentmihalyi's scholastic endeavours was to find an explanation for why some individuals would spend many hours by themselves pursuing tasks that did not show to provide much extrinsic value (Csikszentmihalyi, 1975). However, while having been conceptualised independent of social context, flow theory always recognised the potential for flow in social interaction (Walker, 2010). In recent years flow in social contexts has increasingly gained attention with competing terms such as "team flow" (Keith et al., 2014; Hout, Davis, and Weggeman, 2018), "collective flow" (Salanova et al., 2014), "contagious flow" (Culbertson et al., 2015), "shared flow" (Zumeta et al., 2016), "networked flow" (Gaggioli et al., 2017), and "social flow" (Kaye, 2016; Walker, 2010). Momentum for this emergence has built on the enticing proposition of one of the first salient articles by Walker (2010), that flow in social interaction might represent an even more fulfilling experience than flow in isolation. To date, not only the terms but also the descriptions of group-level flow are somewhat diverse. The breadth and diversity of terms (and underlying theoretical compositions) highlight that there is still much ground to be covered to understand better: (1) what causes and consequences of flow experiences in small groups are, and (2) what the qualitative character of flow experiences in groups is.

In terms of causes (or requirements) of flow in groups, it was summarised in Section 3.3 that causes are first of all described as the individual level flow preconditions (difficulty-skill balance, clear goals and feedback). Yet, two additional preconditions were identified that describe factors required for a collaborative interaction (not just mere co-presence of individuals during a task). The first is a requirement for cooperation, that includes factors as member interdependence (i.e. members are dependent on the work of others) or coordination of actions (i.e. members are able to coordinate their work with others). For example, it is considered, that when individuals in a group are merely co-present but do not have to or cannot share their work, their flow experiences are likely to occur only on the individual level. Walker (2010) describes three types of flow experiences that are solitary, co-active and interactive flow experience, wherein the latest form describes a situation in which interdependent interaction is required. If such interaction is reduced or potentially removed, the shared experience is unlikely to occur, which makes the factor of group member interdependence crucial for the concept of group-level flow. Stated otherwise, if a groups' task can be easily divided into fully independent action, group-level flow will not necessarily be achieved (Keith et al., 2016). The second precondition for flow in groups is a requirement for group member integration. Integration means that group members have well-aligned goals, abilities, roles and procedures. For example, if diverging individual goals are present, or if member abilities are not well aligned, (non-technical) conflict is more likely to occur which hinders the emergence of flow in the social unit (Hout, Davis, and Weggeman, 2018). In line with the reasoning by Walker (2010), these two additional preconditions (cooperation and integration) are herein considered as the foundational differentiating feature between flow experiences in isolation and small groups.

However, an interesting question that has not received much attention is whether or not there are group composition variables that influence the emergence of flow in small groups. Group composition variables are herein understood as the group member characteristics that make up the unique character of each social unit (see

Wildman et al., 2012). Initially, it would seem plausibly in line with flow theory, that flow intensity in a social interaction setting would be dependent on the level of diversity in members' abilities (Hout, Davis, and Weggeman, 2018). Furthermore, beyond such technical diversity, other member characteristics could be considered to impede social flow experiences. For instance, in much of the related work on small group interactions the degree of diversity in small groups (e.g. in the form of cultural, technical, or cognitive diversity) has been found to interfere with group interaction processes and to for example impact group performance and interaction satisfaction (Knippenberg and Schippers, 2007; Wildman et al., 2012). For example, in related small group research, gender diversity or diversity in emotional intelligences have been found to influence task-independent small group performances (Woolley et al., 2010; Engel et al., 2015), even in digitally-mediated interaction scenarios (Engel et al., 2014). How such social unit composition factors might influence social flow, represents a gap in the current body of knowledge.

Regarding the consequences of flow in social interaction, three central dimensions have emerged that are the performance of the social unit, satisfaction from the social interaction, and growth of the social unit (e.g. as a form of building collective efficacy, knowledge, trust, or social relationships). However, most of the present research has been conducted Face-to-face (F2F) in small groups (2-4 individuals), perhaps to explore the social interaction with higher external validity. Another possibility is the recognition that shared experiences are more likely to occur in smaller groups (Aubé, Brunelle, and Rousseau, 2014). Thus, a central gap is present, that is the study of flow in social interactions in Digitally-Mediated Communication (DMC). Digital environments are known to deprive social signals which are critical for the exchange of affective and interpersonal information (Derks, Fischer, and Bos, 2008; Chanel and Mühl, 2015). As digital environments are central to today's decentralised work environments, and as social signals may be important for flow in cooperative scenarios, there is an urgency to investigate flow in DMC small group interactions.

Beyond this consideration of causes and consequences, it is a central question in social flow research, whether or not flow in social units represents a qualitatively different phenomenon from flow in isolation (Walker, 2010). Two central lines of thought have emerged. The first perspective is rooted in traditional flow theory and understands flow in social units to be comprised of the same nine dimensions of flow experience, yet including the additional interdependence precondition (thus sometimes including communication as a precondition) (Keith et al., 2014; Keith et al., 2016), and including intensified experiential outcomes (e.g. enjoyment and growth) (Walker, 2010). The second perspective understands flow in social units as an emergent property from the social interaction, that is better described by own preconditions and state characteristics. Some researchers that follow this understanding have considered social unit flow to be operationalised by a sense of reciprocal flow influencing. Magyaródi and Oláh (2017) and Olsson and Harmat (2018) for example operationalise this social unit flow by using a flow synchronisation questionnaire that gauges the level of (individually perceived) shared fluency and absorption. Tse et al. (2016), on the other hand, follow a more established route by observing the influence of the social unit on reported flow intensities using the Intra-Class Correlation Coefficient (ICC). The ICC is a standard method in small group research to assess the amount of variance in an individuals response that can be explained through the social unit membership (Bliese, 2000). In that sense, shared flow is indicated as the individual-level experience that is however dependent on (influenced by) the social unit. Lastly, even new theoretical models have been proposed that ought to describe the emergent group-level flow phenomenon. One such perspective is brought forward by Hout, Davis, and Weggeman (2018), who outlines the preconditions of (1) collective ambition, (2) common goal, (3) aligned personal goals, (4) high skill integration, (5) open communication, (6) safety, and (7) mutual commitment, and the state characteristics of (8) sense of unity, (9) sense of joint progress, (10) mutual trust, and (11) holistic focus.

In any case (i.e. independent of a particular theoretical perspective), there have so barely been any studies that consider how neurophysiological processes might be involved in such emergent group-level flow experiences. In relation, the only two studies that provide evidence are the work by Keeler et al. (2015) and Labonté-Lemoyne et al. (2016). Keeler et al. (2015) find reduced arousal and stress as indicated by reduced adrenocorticotropic hormone levels that are thought of reflecting social flow experiences. Stated otherwise, the stress-buffering effect (see also Palumbo et al., 2017) from the social interaction is considered to be an element of an emergent group-level flow experience. Interestingly, such a finding would be convergent with the work by Tse et al. (2016) who find that flow experience is intensified in (stress-averse) dyads, especially when tasks are becoming hard. Furthermore, Labonté-Lemoyne et al. (2016) find that one group members boredom might be linked to the second group members flow experience through right frontal Alpha power differences. This finding could also indicate that a lower experience of distress in (at least some) social unit members is beneficial for the flow experience of others. However, none of this research integrates multidimensional data (reports, behaviour, physiology), with a large sample to further elaborate on possible neurophysiological patterns of flow experiences in social interaction (and in particular for the possibility of shared flow experiences). This state represents a salient gap in the research on flow in social units.

In summary, central gaps in the research on flow in social units pertain to (1) the causes of flow in terms of social unit compositions and behaviours, (2) the consequences of flow in social interaction in digitally-mediated environments, and (3) the analysis of the emergence of shared flow experiences, especially paired with the investigation of multidimensional data (report, behaviour, neurophysiology). Given that Experiment 2 in this dissertation collected exactly this data, it allowed closing these gaps. Thus, a dedicated set of exploratory analyses was conducted. The guiding Research Questions (RQ) that are investigated in this context of digitally-mediated environments are as follows:

- **RQ9** Is flow in social units influenced by group composition (diversity)?
- **RQ10** What are the relationships of flow with group performance, satisfaction, and growth?
- **RQ11** How does shared flow emerge in digitally-mediated environments?

Together, these additional analyses focusing on small group dynamics and group-level experiences contribute to the literature on flow experience and the development of adaptive NeuroIS by (1) extending the knowledge on flow in digitally-mediated social interaction through analyses of possible causes and consequences of flow in these settings, and by (2) being the first to explore influences on group-level (shared) flow experiences. Specifically, by analysing data across report, behaviour and neurophysiology domains, possibilities are identified as to when shared flow experiences may emerge (identification of boundary conditions). In addition, opportunities for the unobtrusive detection of such group-level experiences are proposed together with simple extensions to flow theory on how to integrate social dynamics into the premise of optimal difficulty as a precondition to flow experience. The following sections build on the previous documentations of related work (see Section 3.3) and leverage the data collected in Experiment 2 of this dissertation (see Chapter 6). Therefore, the following chapter only briefly appends data processing descriptions as needed for the novel analyses.

7.2 Data Processing

In this chapter, three groups of analyses were conducted following the RQs, as mentioned above. In general, the data (pre-)processing strategies were identical to those reported in Section 6.2.4. Report, behaviour and neurophysiological data were utilised after the processing mentioned previously. Wherever necessary, additional, subsequent processing was performed. The description of these steps, including their rationale, are outlined here.

The first analysis focuses on the influence of small group member diversities on flow experiences. For this purpose, a Linear Mixed Model (LMM) similar to the previously developed one in Section 6.3.2 was created. This time, however, only the MP condition data was utilised. The model development process comprised a sequential build-up, first including known random effects (level-1 and level-2) and fixed effects (difficulty and flow proneness), and then including fixed effects for reported diversity metrics (perceived diversity in member efforts, technical abilities, and general perception of group diversity) and demographic diversity metrics (age range, gender distribution, technical backgrounds). P-values were BH-corrected.

The second analysis focuses on the relationship of flow experiences with variables related to desirable group experiences, specifically regarding group performance, satisfaction and growth. The perceptions of flow and group performance were collected as a Repeated Measures (RM) after each condition. The remaining variables were collected once at the end of the experiment. The quality of the self-report instruments (internal consistencies for multi-item measures) was already reported in Table 6.3 in Section 6.2.4. To identify relationships, suitable correlation metrics were used. This means that for the RM measures (performance), RM correlations (Bakdash and Marusich, 2017; Bland and Altman, 1995) were calculated. For the remaining variables, Pearson correlations between the average flow per participant and their perceived interaction experiences were calculated. P-values were BH-corrected.

The third analysis focuses on the emergence of shared flow experiences. For this purpose, the metric called ICC was utilised that describes the amount of variance that can be explained in a person's variable outcome (report, behaviour or physiology) by the group membership. The ICC as it is used in small group research (Bliese, 2000) is determined as the variance from the group level (level-2) divided by the sum of the variance from the group level (level-2) and the residual variance in the dependent variable (or the total variance in short). The ICC1 can be considered an index of the reliability of the group means (Bliese, 2000). This property means that for high ICC values, a single rating from an individual can be considered to provide a relatively reliable rating of the group mean. However, for lower ICC levels, multiple ratings might be required to provide reliable estimates of group means (Bliese, 2000). In this research, the ICC1 was computed from an LMM with the group level (3-person unit) as a random effect variable for univariate and multivariate analyses (i.e. across all difficulty conditions and for each difficulty condition). For the univariate analysis, the within-subject level (RM) was also included as a random effect. In theory, the ICC, when calculated from an LMM can take values from 0 to 1, with 0 representing the absence of group influence, and 1 representing complete determination of the individual's response by the group. In practice, rather low values are typically found, with values from 0 to 0.4. Values of 0.5 are considered exceptional (Musca et al., 2011), and values above 0.25 are considered large effects, values between 0.25 and 0.1 moderate effects, and values above 0.05 small effects (Bliese, 2000). For simplification, a level of 0.1 means that 10% in the DV response of an individual can be attributed to the membership to a group. By this property, the ICC can be used to determine, if a DV can be considered to be aggregatable to the group level (e.g. through mean averaging) and if a shared experience is present. In short, the emergence of a shared flow experience in the present data can be considered to be present, if elevated (i.e. roughly > 0.05) ICC1 levels are found. By utilising the ICC assessment of group member reciprocal influences on flow, an advantage is that the analysis is bound to the fundaments of flow theory and the employed survey instruments.

To enable follow-up analysis on the ICCs across the different data domains (report, behaviour or physiology), it was decided to impute missing values to retain sufficient statistical power. Otherwise, the removal of missing data in each domain would have strongly narrowed the number of complete data sets (i.e. complete groups). Groups with data for at least two members were retained (again, to retain statistical power), as the LMMs handle unbalanced data well (Cnaan, Laird, and Slasor, 2005). Predictive mean matching (Vink et al., 2014) was employed for the imputation, using the repeated measures and the group ID as predictor variables. This way, only data missing at random were imputed, not entire missing sessions. Thus, primarily those data were imputed that were removed through previous outlier removal, but also instances where EEG features might be missing for an ROI where electrodes were removed in the data preparation process. In addition, potential group-level effects are already accounted for in the imputation by including the group ID in the imputation regression. Therefore, in the end, matched (i.e. available completely for each participant) variables for 101 participants (37 groups in total, 27 of them with three members, 10 with two members) were available.

Given that the ICC metric is a variance-based point estimate, for which no general distribution is known, comparison of ICC values requires the use of resampling methods. Initially, bootstrapping with 5000 repetitions (with replacement on the group level) was employed to determine ICC1 Confidence Interval (CI)s. This resampling approach returned very large CI estimates with ranges up to .311, which would mean, that the present effect sizes would be likely to occur completely at random (i.e. an effect per condition could just as well be none as it could be exceptional). Given that bootstrapped CIs from small samples show a substantial variation (Hesterberg, 2015), the complicated statistical distribution of the data underlying the ICC1 (Ren, Yang, and Lai, 2006), and the finding that bootstrapping ICC CIs performs poorly for small cluster sizes (Wang et al., 2019), the unintuitive CI results here were considered to be too conservative. In related work, ICC-specific bootstrapping algorithms have for such purposes been explicitly developed for the cases of generalised linear models with binary outcomes (Ren, Yang, and Lai, 2006) or non-linear mixed models (Demetrashvili and Van den Heuvel, 2015). Furthermore, in a related manner, more straightforward, leave-one-out resampling methods have been found to perform better than bootstrapping approaches for the CI-based comparison of different clustering coefficients (Severiano et al., 2011). Due to this understanding, as a compromise between the overly conservative bootstrapping results and too optimistic leave-one-out CIs, a leave-3-out resampling approach was chosen to estimate the CIs for the ICC1 values in this experiment and, more specifically, for the pairwise differences amongst them. Leaving out three groups per estimation, a total number of $\binom{37}{3} = 7770$ resampled estimates were computed.

Following the ICC evaluations for the identification of reciprocal (flow) influences, a set of follow-up analyses were conducted. These approaches have not been previously pursued in (social) flow research and represent a major contribution of this work. Specifically, through a combination of median splits with the aforementioned resampling methods (to compare sub-sampled ICC1 estimates), it is explored which variables indicate a relationship with the strength of flow group dependence. It is important to note, that these analyses provide insight not on the per-group intensity of a shared flow experience, but rather on conditions under which a reciprocal influence on flow experiences emerges in groups in general. The process for these analyses is as follows. First, the designated explanatory variable of interest (always the mean average of the group) is used to split all the groups into sets with lower and higher levels along the median (excluding groups that are on the median). Second, the flow ICC1 metrics are calculated for the lower and higher DV expression sub-groups. In step three, the previous process is repeated for each sample of a leave-three-out resampled data subset to estimate CIs of differences between the two sub-groups (lower and higher variable levels). However, given the already small data set, a median split approach comes with a severe limitation in statistical power. Given the potential that this can introduce sporadic results, an additional sensitivity analysis was conducted to further assess the likeliness and robustness of the median split results. For this sensitivity analysis, ICC1 levels are estimated for sequentially shortened data, in which lower and higher percentiles of the data are removed (in steps of 1 percentile). This process means, for example, that in a first step, samples in the 99th percentile are removed from the data, in the second step the groups in the 98th percentile, and so forth until the data contains only the data up to the 50th percentile (excluding the median again). The samples thus represent instances that are step by step closer to the below-median sub-group, or stated otherwise, the samples represent instances where the highest-scoring groups are sequentially eliminated. Then this process is also completed by starting to remove data from the 1st percentile up to the 50th percentile (the lowest scoring groups are stepwise removed). This process allows inspecting if a result found in the median split is following a trend or rather represents a sporadic result. This process is visualised in Figure 7.1. Given the different frequencies of collected data, these analyses are either conducted for the univariate ICCs (across all difficulty conditions - for variables collected at the end of the experiment) or for the multivariate ICCs (for variables collected after each condition). Through this approach, there are no aggregation mismatches created for the interpretation (separation of general perception levels and conditional perceptions). Furthermore, for the multivariate ICC-split analyses, the two conditions with optimal difficulties (CAL and AUTO) were used only. This step was taken for two reasons. First, it was considered that multiple evaluations would be useful to deflate the previously outlined risk of false positives and negatives. Second, given that the border conditions (EASY and HARD) have shown apparent differences in various variables to the two optimal difficulty conditions, they were excluded from this evaluation (see Section 6.3).

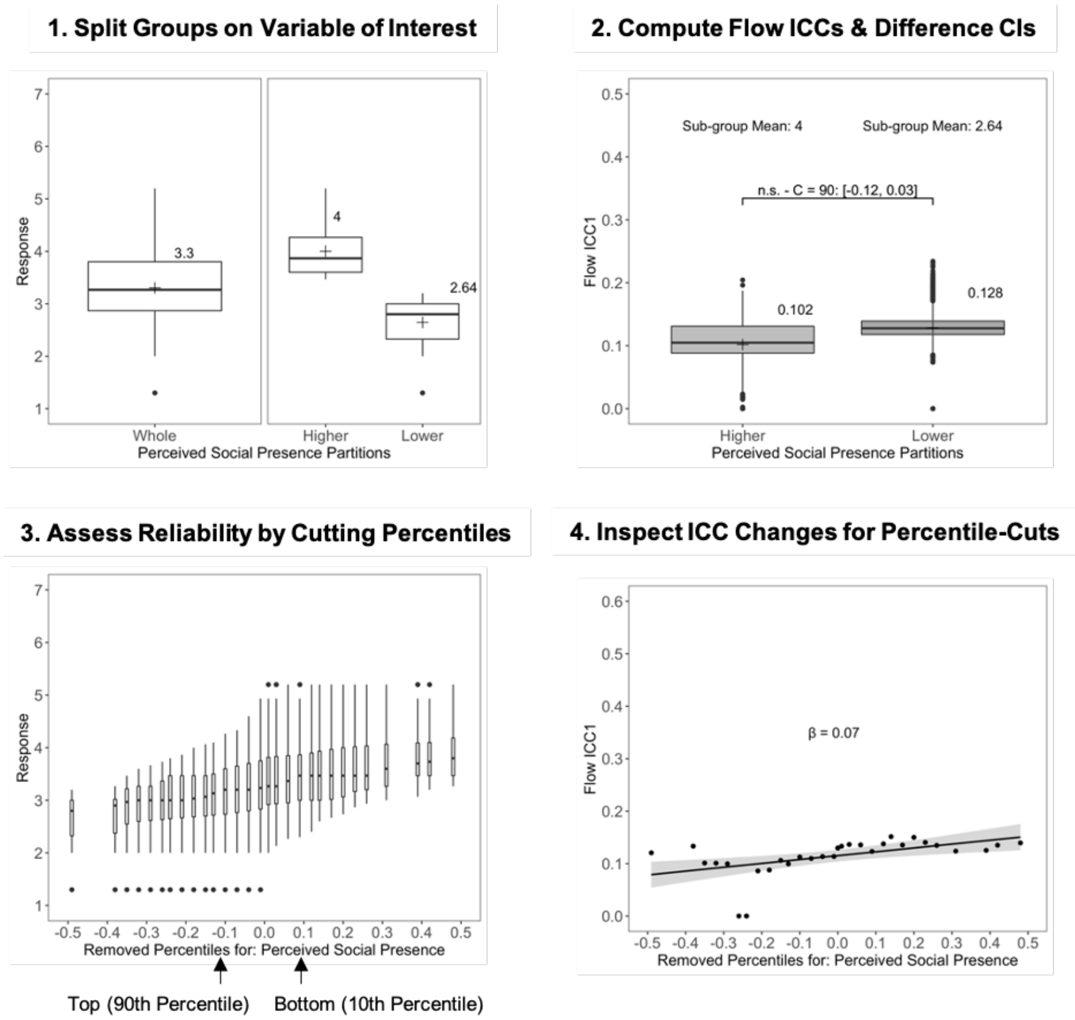

FIGURE 7.1: ICC Influence Analysis Process.

Given that the previously mentioned analyses do not allow to infer single-group flow clustering levels (due to their high aggregation level), in a final analysis step, variables with similar ICC patterns are used for regression analyses that include mean-aggregated variables, whenever possible (i.e. for conditions in which ICC levels > 0.05 are found - see Bliese, 2000). In this approach, it is especially of interest, whether or not data with higher temporal availability (here: neurophysiological - precisely, EEG data) shows predictive power for mean flow levels in groups.

7.3 Results

7.3.1 Flow & Small Group Diversity

The first analysis focuses on the influence of small group member diversities on flow experiences. For this purpose, an LMM similar to the previously developed one in Section 6.3.2 was created. This time, however, only the MP condition data was utilised. The model with the best fit includes individual (level-1) and group (level-2) random effects, and fixed effects for difficulty, flow proneness, and reported diversity metrics (perceived evenness in member effort, integration of abilities and expertise, and general perception of group diversity). Demographic diversity metrics (age ranges, gender distribution - as the number of women per group, and technical backgrounds - as the number of different study majors) were not included in the final model as the did not improve model fit. The final model is shown in Table 7.1. The results show that besides the already identified fixed effects, only perceived diversity in member effort has a significant, positive influence on perceived flow (trend level). The latter result means that flow experiences increased when individuals perceived all group members to put forward similarly high levels of effort. In contrast, other indicators of diversity (perceived diversity, integration of abilities, age range, gender, and technical diversity) show no relationship with flow in this experiment.

Metrics	Flow (MP)
Fixed Effects	
Difficulty^F	25.9192***
Flow Proneness^C	0.4887 (0.1656)*
General Group Diversity^C	0.0873 (0.0684)
Even Effort^C	0.1274 (0.0571)^t
Skills Integration^C	-0.0773 (0.0725)
Random Effects	
Level 1 Resid. (Indiv.)	0.1213***
Level 2 Resid. (Grp.)	0.0045**
Goodness-of-Fit	
AIC	902.8918
BIC	946.2915
Marginal R^2	0.195
Conditional R^2	0.359

Notes: ^{t}p <.1; *p <.05; **p <.01; ***p <.001; ^{F} = Factor Variable (Showing F-Test Results); ^{C} = Continuous Variable (Showing Beta-Coefficients with SE); Random Effects Shown as Variance; P-Values are BH-Corrected.

TABLE 7.1: LMM for Flow in the MP Condition Including Diversity.

7.3.2 Flow & Small Group Experiences

The second analysis focuses on the relationship of flow experiences with variables related to desirable group experiences, specifically regarding group performance, satisfaction and growth. The results for the RM variables (performance indicators) are shown in Figure 7.2. The results for the group satisfaction and growth variables

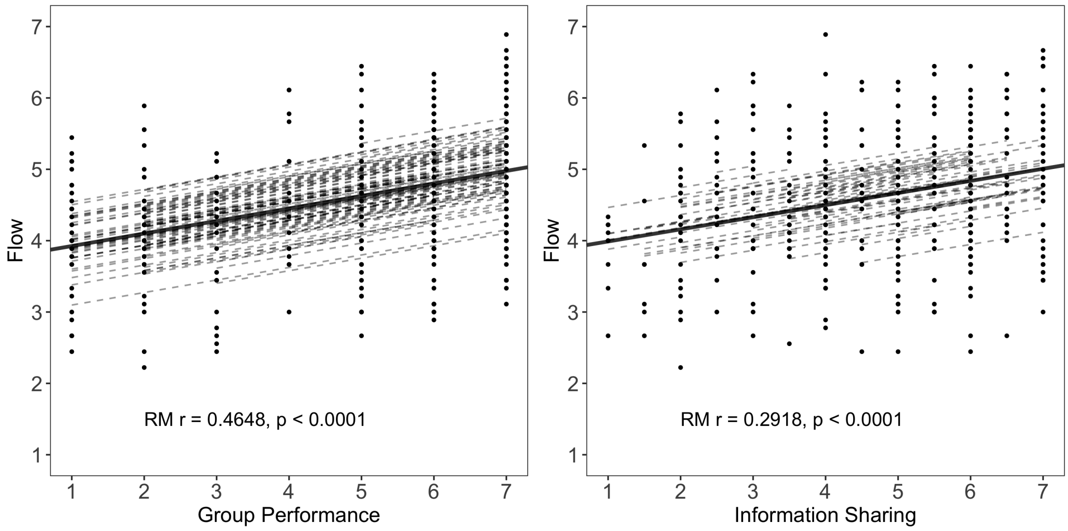

FIGURE 7.2: RM Correlations of Flow Reports with Perceived Group Performance Reports in the MP Condition in Experiment 2. P-Values Are BH-Corrected. The Fit Line is Created From an LMM. Dashed Lines Represent Predicted Values Per Subject With Random Intercept.

are shown in Figure 7.3. For performance, both the perception of group performance and the perception of the amount of information sharing - a measure that has been previously used as an indirect measure of performance in small groups (Aubé, Brunelle, and Rousseau, 2014; Heyne, Pavlas, and Salas, 2011) confirm this relationship through RM correlations. By analysing the mean flow experiences together with general session perceptions from the end of the experiment, further group-level results are found. Firstly, significant, positive small to moderately sized correlations are found between flow and variables indicating a sense of satisfaction with the group like a sense of enjoyable group member interactions and relationships, general satisfaction with the group work, and a sense of affective involvement with the group work. Further, significant, positive and moderate correlations are found between flow and perceptions of individual growth through the interaction with the group, perceptions of collective efficacy (a sense of growth of the group as a unit), and perceptions of identity fusion (a sense of identification with the group).

Thus, across numerous variables, flow experiences show positive correlations with indicators for group performance, satisfaction and growth. Previous research shows fitting results (see Table 3.1), but show for the first time, that these relationships are found even in digitally mediated-interaction scenarios. This finding is especially important as the previous chapter has shown that a too restrictive interaction format might impede flow experiences in small groups. Nevertheless, such limitations to flow emergence do not seem to impede the relationship with these desirable group interaction experiences. Such results highlight the potential that flow support might be a useful vehicle to elicit or a representative indicator to assess desirable experiences in knowledge work for individuals and particularly for groups.

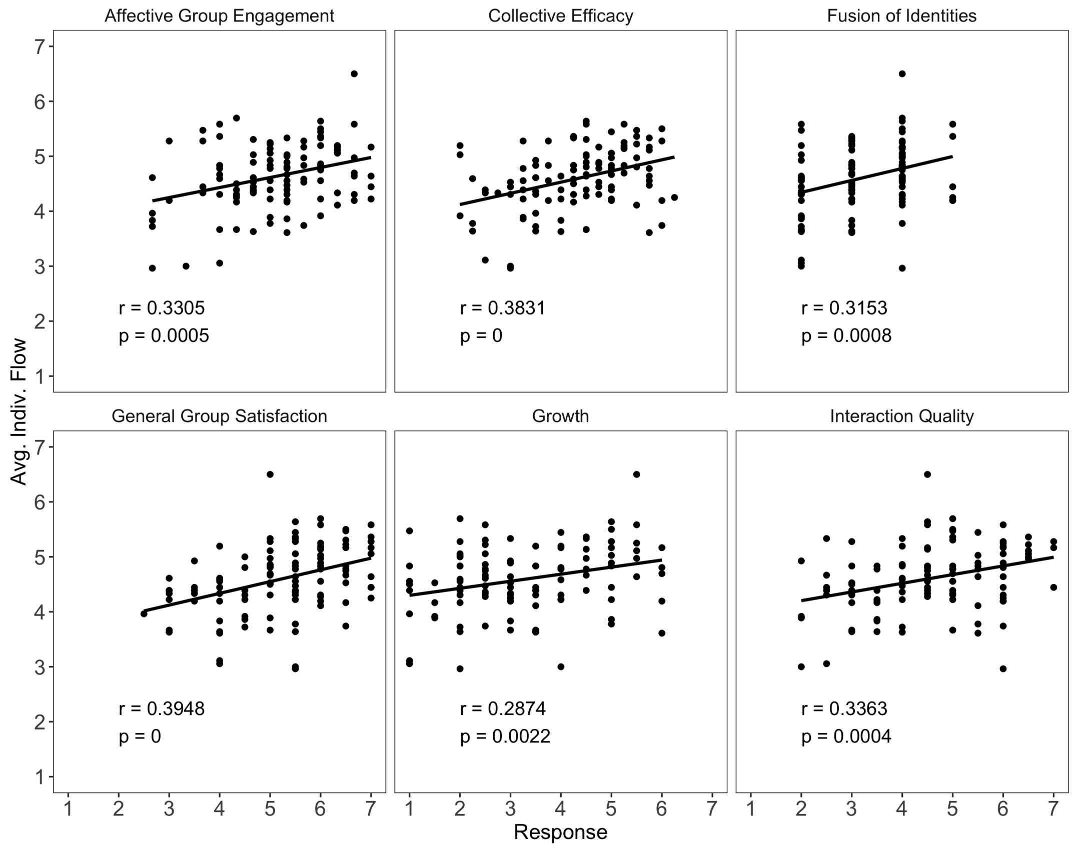

FIGURE 7.3: Correlations of Avg. Flow Reports per Participant with Group Interaction Experience Variables in the MP Condition in Experiment 2. P-Values Are BH-Corrected. The Fit Line is Created From a LM.

7.3.3 Shared Flow Experience Patterns

Flow ICCs

In social flow research (see, e.g. Magyaródi and Oláh, 2017; Heyne, Pavlas, and Salas, 2011; Hout, Davis, and Weggeman, 2018) it is proposed by some researchers, that the experience of flow in social units represents a shared, group-level experience. Some researchers also propose that the group-level flow experience is phenomenologically different from individual flow (Hout, Davis, and Weggeman, 2018). Some of this research has already attempted to explore the possibility of a shared group level experience. For example, Tse et al. (2016) find that individual flow experience reports in dyadic face-to-face interactions (using a DM paradigm) are highly influenced by other group members (ICC of .447) and are thus validly aggregatable to the group level. Other research that has argued for group-level aggregation of flow experiences has so far not provided evidence for this aggregation potential (see, e.g. Heyne, Pavlas, and Salas, 2011; Magyaródi and Oláh, 2017). Given that the present research represents an interesting variation to this related work (i.e. being situated in a digitally-mediated interaction setup with a cognitive task and the AM as a novelty in a small group DM approach), the presence of shared flow experiences was explored.

To assess the emergence of shared flow experiences in this experiment, univariate and multivariate ICC1 coefficients were computed. For the univariate ICC1, in addition to the group ID as a random effect, the individual ID was added as a random effect to account for individual changes in flow reports across the difficulty conditions (similar to the approach by Tse et al., 2016). For a visualisation of the underlying data, see Figure 7.4. Figure 7.5 shows the ICC estimates, including the recommended interpretations for effect sizes by (Bliese, 2000) and Figure 7.6 shows the multivariate resampled ICC distributions, including pairwise differences. For the evaluation of the results, the absolute values of the point estimates along the recommended effect sizes are interpreted. The overall level of group influences on flow reporting is moderate (univ. ICC1 = .130) and substantially lower than in the only comparable study available (Tse et al., 2016 report an ICC1 of .447 in a dyadic, face-to-face puzzle task with varied difficulties). Therefore, while the data suggest that flow experiences are reciprocally influenced in groups (i.e. that there is some possibility for the presence of a "shared" flow), it appears to be much lower than in related work. Furthermore, from the assessment of multivariate ICCs across difficulty conditions, the present data suggest that the potential for such a "shared" flow appears to disappear, when a task becomes (too) hard (ICC = 0 in HARD in contrast to moderate effect levels in the other difficulty conditions).

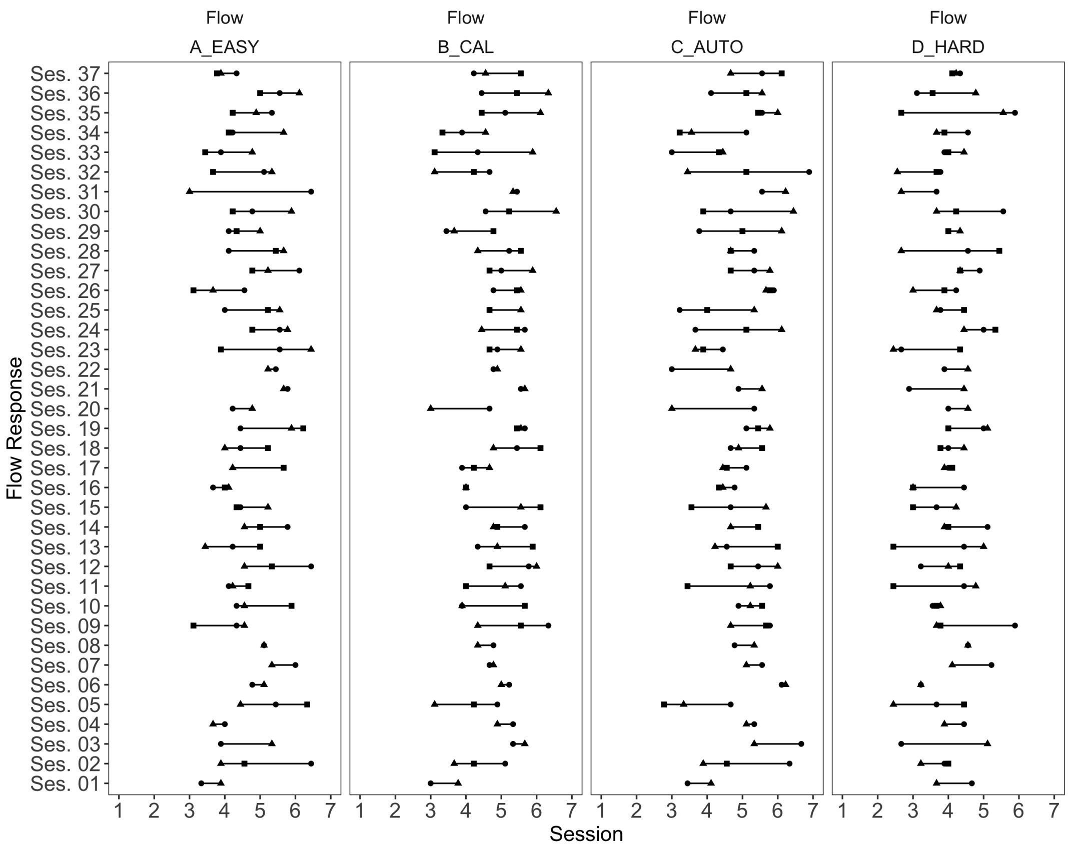

FIGURE 7.4: Flow Reports Per Group and Difficulty Condition in the MP Condition in Experiment 2. Shapes Represent Group Members.

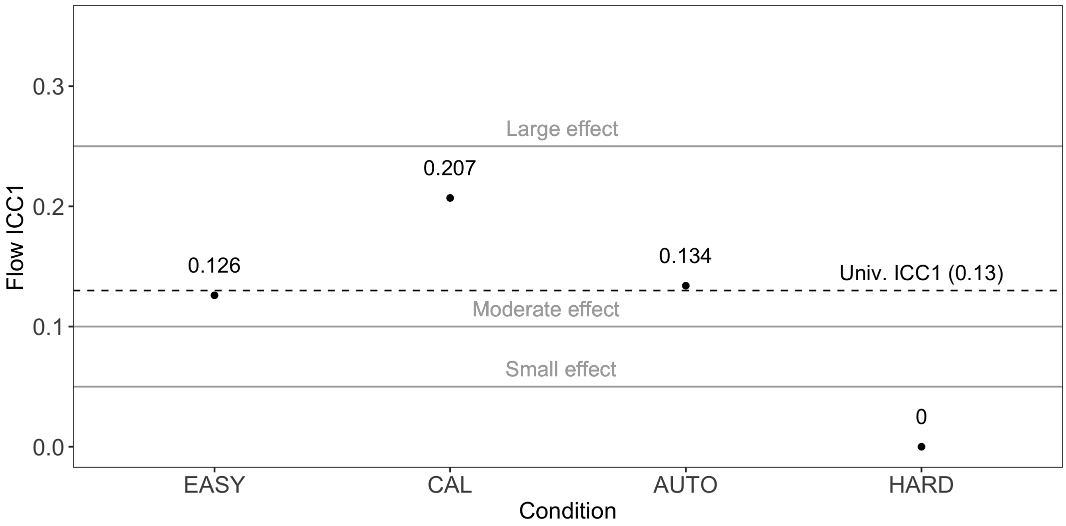

FIGURE 7.5: Univariate and Multivariate Flow ICC1 Point Estimates Incl. Recommended Effect Size Thresholds. The Black Dashed Line Represents the Univariate Flow ICC1.

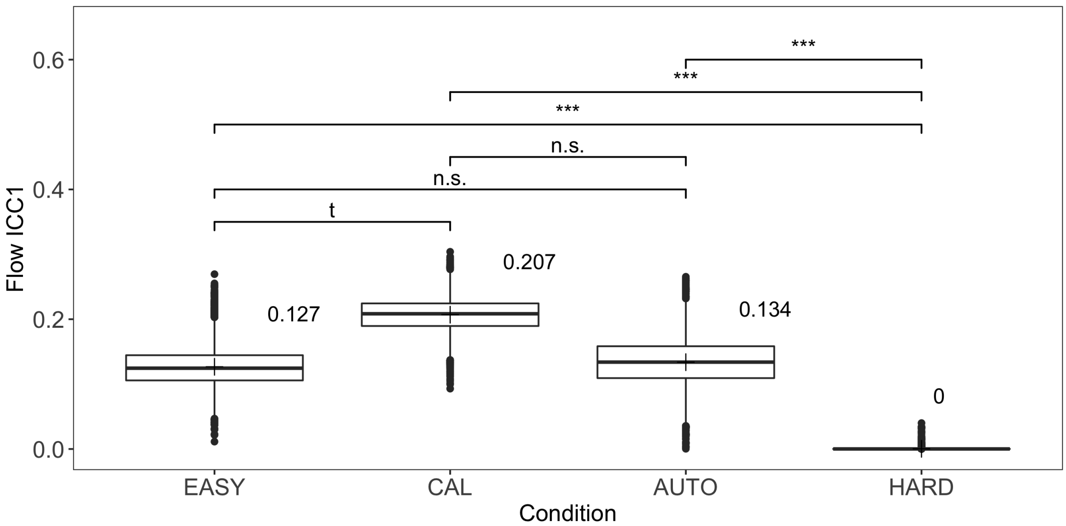

FIGURE 7.6: Leave-Three-Groups-Out Resampled Multivariate ICC1 Distributions Incl. Significance Levels of Pairwise Differences. Crosses and Numbers Next to Them Represent Resampled Means.

Possible Explanations for Low Flow Clustering

Overall, the ICC results indicate a present, yet only moderately strong influence of the social unit on individual flow reports. This observation is in stark contrast to previously reported findings (Tse et al., 2016). The reasons for this difference could lie in peculiarities of the experimental designs and cultural backgrounds. Tse et al. (2016) use a full Within-Subject (WS) design that includes conditions with and without social interaction for each participant. Experiencing the contrasts of these situations directly might be causing different strengths of reciprocal flow influence to emerge (remember: flow could also be getting worse in the social interaction - perhaps especially in simple experiment tasks). Also, Tse et al. (2016) work with an Asian sample. Due to stronger collectivist cultural orientations (Hofstede, 1984), it could be possible that a stronger influence of other group members on individual flow emerges by tendencies to integrate the actions of other's more strongly into one's actions. Furthermore, as shared flow has been suggested to more readily emerge in smaller sized groups (Armstrong, 2008), merely the fact that Tse et al. (2016) observed dyadic interactions could be a central cause for the higher ICC levels in that study. More experiments are required to elaborate if these possibilities represent valid explanations for different intensities of reciprocal flow influencing in small groups.

However, given the previously documented results on lower flow intensities in social interaction in this experiment (see Section 6.3.2), it is also possible that the lower ICC levels might be related to other factors that can be further investigated given the present data. Importantly, these factors relate to the previous observation, that the MP condition represents a rather restrictive environment in terms of social information, task degrees of freedom, and integration of individual difficulty preferences. It is considered possible, that the emergence of reciprocal flow influences, might be limited by these factors: First, a lack of social signal information might be reducing the potential to experience shared flow by limiting processes like emotional contagion or stress-buffering, that might be involved in the emergence of shared flow experiences (see, e.g. Labonté-Lemoyne et al., 2016). Given that the experiment in Tse et al. (2016) was conducted in a F2F setting, the ICC difference to this study could also be explained by the reduced social information factor. Second, a lack of degrees of freedom to select sub-tasks that are optimally difficult might also be reducing the possibility for shared flow emergence, as individual strengths, weakness and preferences cannot be usefully integrated to leverage the potentials of acting in a small group. Again, comparing the present design to that by Tse et al. (2016) in which dyads solve puzzles with many more sub-tasks (puzzle tiles), this explanation for the ICC differences is similarly plausible. Given that the present experiment includes variables for perceptions of social presence and autonomy, and recorded preferences for optimal task difficulty of each group member, these two possibilities were evaluated using median splits.

As perceptions of social presence were collected at the end of the experiment, data splits were performed on the group average perceived social presences, and univariate flow ICCs were estimated afterwards. Figure 7.7 shows the results of the flow ICC estimations. The paired difference CIs indicate that no significant differences are found after the social presence median split. Also, the LM in the sensitivity analysis further indicates the absence of such an effect. Therefore, it is considered that differences in social presence perceptions are not related to the emergence of shared flow. However, it should be noted that the absolute values for perceptions of social presence are rather low. Also, as was discussed in the previous chapter, the social presence variable cannot be seen as a representative for the wide variety of social dynamics that could be possible to occur and interact with the emergence of group influences on flow experiences. Therefore, more research is needed to substantiate if the current result representatively describes the independence of social information and "shared" flow.

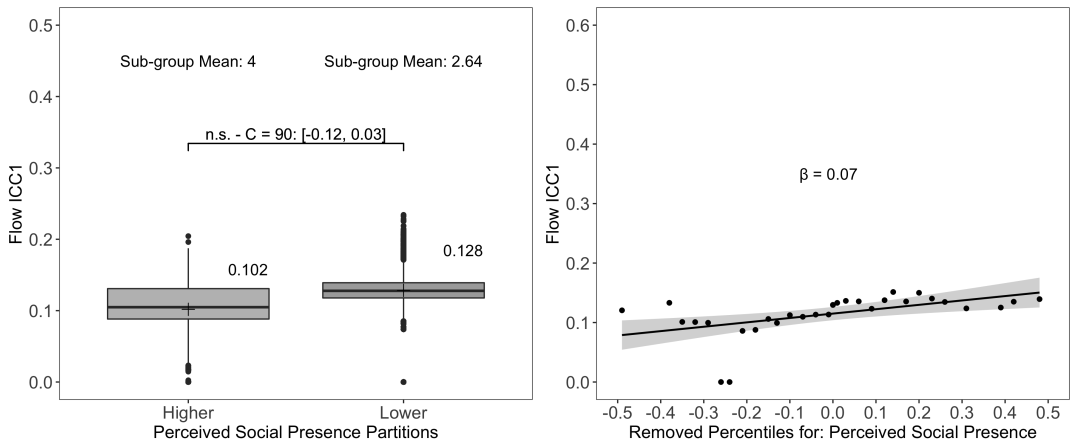

FIGURE 7.7: Flow ICC1 Estimates After Median Splits on (Group Average) Perceived Social Presence in Experiment 2. C = Confidence Level; Left: Leave-Three-Groups-Out Resampled ICC Estimates. Right: Flow ICCs After Stepwise Removal of Lowest and Highest Percentiles.

Preferences for optimal difficulty were collected once per participant and group at the start of the AUTO condition. Data splits were performed on the range of difficulty preferences, and multivariate flow ICCs were estimated afterwards, in this instance, for the AUTO condition as it reflects the time of preference elicitation. Figure 7.8 shows the results of the flow ICC estimations. The paired difference CIs indicate a (trend level) significant difference for the median split, with higher flow ICCs when difficulty preferences are more similar. The LM in the sensitivity analysis further indicates that such a trend might be present. Therefore, it is considered that similarity or diversity in preferences for optimal difficulty are weakly related to the emergence of reciprocal flow influences.

Lastly, perceptions of autonomy were collected after each difficulty condition. Therefore, data splits were performed on (group average) perceived autonomy and multivariate flow ICCs were estimated afterwards. Results are presented for the optimal difficulty conditions CAL and AUTO, as these are considered to be the most representative, natural situations in the experiment (see Section 7.2). Figure 7.9 shows the results of the flow ICC estimations. The paired difference CIs indicate significantly higher flow ICC in the CAL condition, but not the AUTO condition. However, in the latter condition, the a difference is similarly visible, and the sensitivity analysis suggests a trend in both conditions. Therefore, it is considered that differences in autonomy perceptions are related to the strength of group influences on flow. These results lend first evidence for the proposition that increased autonomy in groups influences the reciprocal flow influence of group members. Yet, the effect size might be moderated by the task difficulty, dissipating with more difficult tasks.

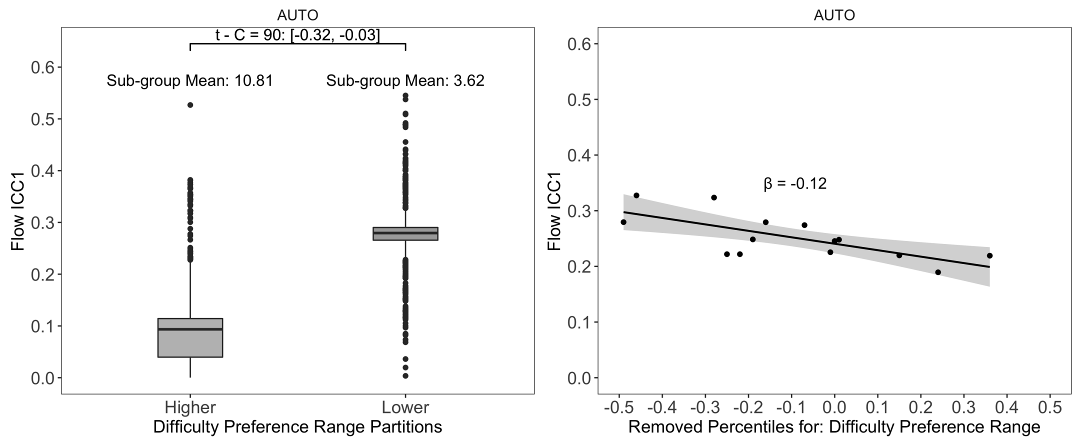

FIGURE 7.8: Flow ICC1 Estimates After Median Splits on (Group Range) Preferences for Optimal Difficulty in Experiment 2. C = Confidence Level; Left: Leave-Three-Groups-Out Resampled ICC Estimates. Right: Flow ICCs After Stepwise Removal of Lowest and Highest Percentiles.

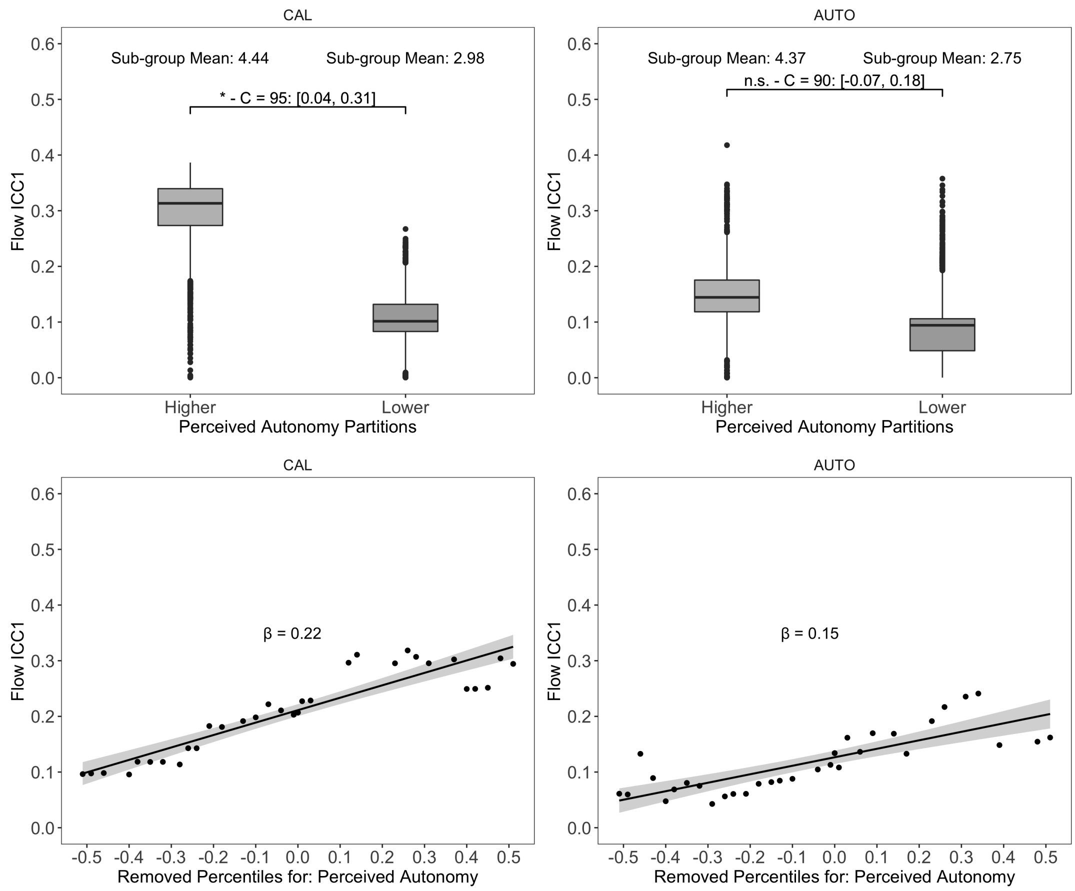

FIGURE 7.9: Flow ICC1 Estimates After Median Splits on (Group Average) Perceived Autonomy in CAL and AUTO Optimal Difficulty Conditions in Experiment 2. Left: Leave-Three-Groups-Out Resampled ICC Estimates. Right: Flow ICCs After Stepwise Removal of Lowest and Highest Percentiles for Social Presence.

In a last set of analyses, it was assessed further if variables related to the diversity of group members show a relationship to flow ICC levels. Specifically, perceived diversity metrics (general group diversity, even efforts, skills integration and average flow proneness) were assessed. The demographic group composition variables (e.g. gender, technical background) that were analysed previously were excluded from this analysis as their distributions were not suitable for the median split analysis. The results of the analyses are summarised in Table 7.2. No relationships between univariate flow ICC levels and these perceived diversity metrics were found. This result indicates that the emergence of flow group-level influences of flow is independent of such group member diversity factors. This pattern is further in line with the general impression from the previous flow-diversity analyses, that flow in groups (at least in this setting) might not be so dependent on diversity factors, or that these variations were not strong enough to show as significant effects in this experiment.

Variable	Means	ICCs	Sig.	CI	β
Flow Proneness	4.1 / 3.76	0.123 / 0.090	n.s.	C = 90 [-0.04, 0.10]	0.04
Gen. Grp. Diversity	6.10 / 4.72	0.148 / 0.079	n.s.	C = 90 [-0.09, 0.17]	0.05
Even Effort	6.06 / 4.62	0.106 / 0.111	n.s.	C = 90 [-0.10, 0.05]	0.00
Skills Integration	5.52 / 4.23	0.091 / 0.143	n.s.	C = 90 [-0.14, 0.01]	-0.03

Notes: First Value is for Upper Sub-Group, Second Value is for Lower Sub-Group; β Represents the LM Coefficient From the Sensitivity Analysis.

TABLE 7.2: Univariate Flow ICC Relationships to Diversity Metrics in Experiment 2.

Altogether, these follow-up analyses provide first empirical evidence for some of the developed propositions that reciprocal flow influence might be dependent on task degrees of freedom, or integration of individual difficulty preferences. These results do not disconfirm other propositions and are insufficient to rigorously substantiate how the emergence of shared flow is determined. Yet, they provide a starting point for the development of dedicated experiments that can provide better means to confirm the results. Presently, a hypothesis that can be derived as a starting point is that higher levels of autonomy are needed for the emergence of a shared flow experience. This result means in contrast, that when group members feel that they cannot act autonomously (i.e. that their actions are determined externally), they will less likely experience shared flow.

Shared Flow Changes with Difficulty

Beyond the assessment of flow report clustering over the whole experiment, clustering indices were also analysed at a more granular level. Multivariate (i.e. difficulty condition-specific) flow ICCs have not yet been reported in the related literature. However, such elaborations could help in explaining what makes a shared flow experience happen. In the present experiment, moderate ICCs are visible for flow in all difficulty conditions except HARD (see Figure 7.6). To further assess the robustness of this finding, additional indicators for difficulty from three data domains (report, behaviour, and physiology) were utilised to conduct follow-up analyses. Specifically, (group average) reported difficulties, objective task difficulties (task difficulty levels), and HiBeta EEG powers (the most sensitive mental workload feature in the present neurophysiological data) were used in median split analyses. The results for reported difficulties are shown in Figure 7.10, for task difficulty levels in Figure 7.11, and HiBeta (at all four ROIs) in Figures 7.12 to 7.15. For all three data domains, the results further confirm the initial assessment, that with higher difficulties flow ICC levels are reduced. This confirmation is the case in both optimal difficulty conditions. For the HiBeta features, this effect is most visible in the CAL condition at all ROIs, and visible in both CAL and AUTO conditions at Central scalp electrode positions. It is important to note that the HiBeta feature at Central sites has in previous flow-EEG analyses also shown the strongest effect sizes of these four ROIs (see Section 6.3.3), which further indicates a particular relevance of this cortical site for the identification of flow boundary conditions.

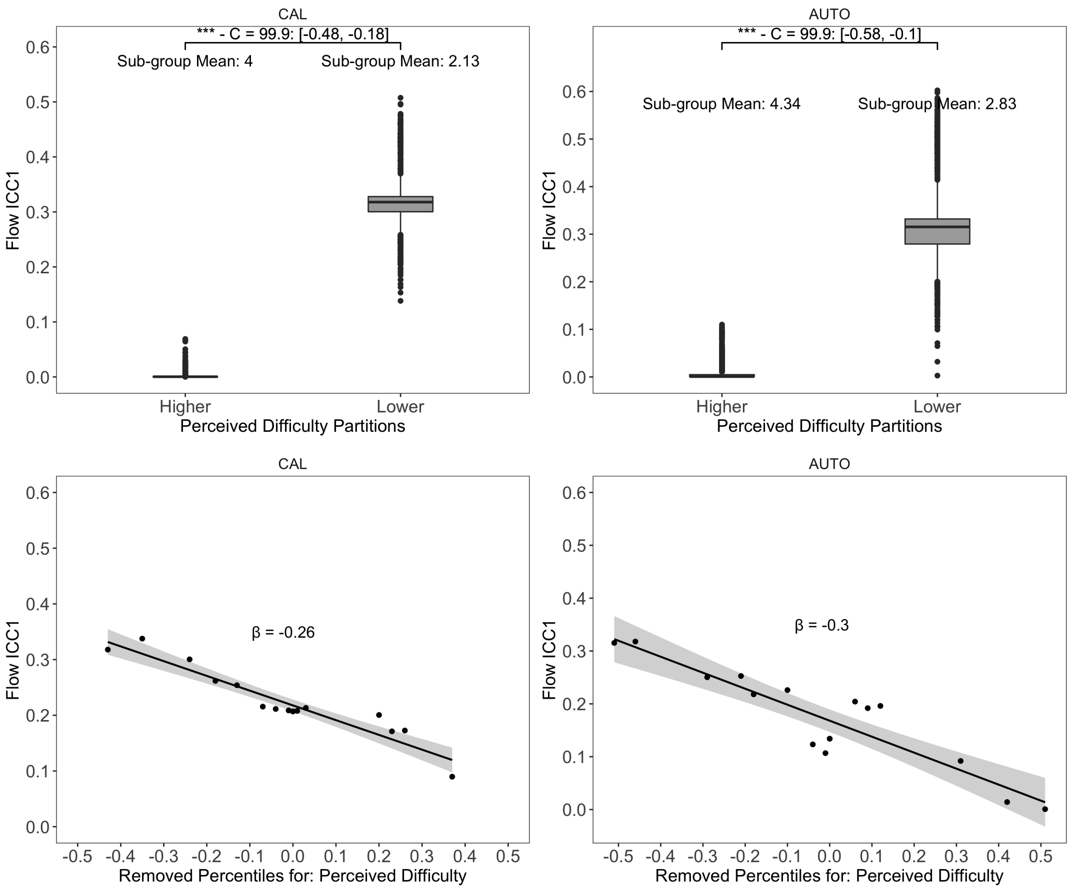

FIGURE 7.10: Flow ICC1 Estimates After Median Splits on (Group Average) Perceived Difficulty in CAL and AUTO Optimal Difficulty Conditions in Experiment 2. Left: Leave-Three-Groups-Out Resampled ICC Estimates. Right: Flow ICCs After Stepwise Removal of Lowest and Highest Percentiles.

The finding that metrics that assess difficulties from various domains unanimously support the dissipation of reciprocal flow influences with high difficulties shows the robustness of this observation. A possible explanation for this phenomenon is that with too high task difficulties, group members turn their attention to themselves and away from other group members. Such a disengagement from the group (a form of isolation) could explain why shared flow experiences are no longer emerging, as essentially every group member might be acting by themselves alone. Thus, in too hard tasks, despite formal interdependence, only a form of co-located flow (see Walker, 2010) can emerge. Note again, that in this situation, it is still possible

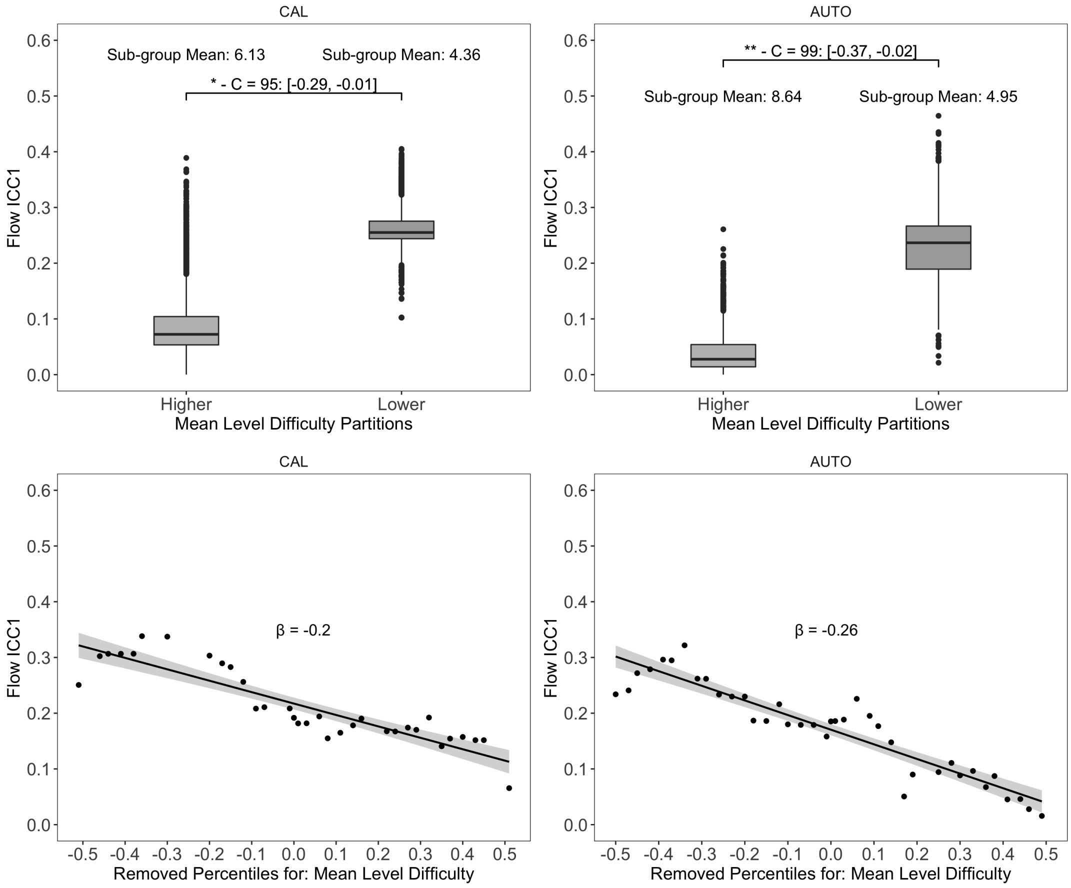

FIGURE 7.11: Flow ICC1 Estimates After Median Splits on (Average) Task Difficulty Levels in CAL and AUTO Optimal Difficulty Conditions in Experiment 2. Left: Leave-Three-Groups-Out Resampled ICC Estimates. Right: Flow ICCs After Stepwise Removal of Lowest and Highest Percentiles.

that individuals might experience some form of flow by themselves, it is just no longer influenced by other group members. The observation that such instances might be readily identifiable using a neurophysiological marker with high temporal resolution (that is less intrusive than reports and more task-independent than a task difficulty level feature) provides an exciting opportunity for future research and the development of adaptive NeuroIS. Given sufficient calibration of a respective system, HiBeta EEG power could be used to identify situations when tasks are becoming too difficult for group members and shared flow experiences are no longer possible. Such EEG features that are reflective of mental workload states could also be employed in efforts to identify situations of too low individual difficulties. Therefore, they appear to possess the potential to calibrate, for example, task automation or feedback systems that inform individuals and groups about the distribution of workloads. Lastly, another possibility was previously indicated, that is the emergence of group-level influences on HiBeta features in contrast to most other neurophysiological features. The relevance of this finding is further pursued in the next section.

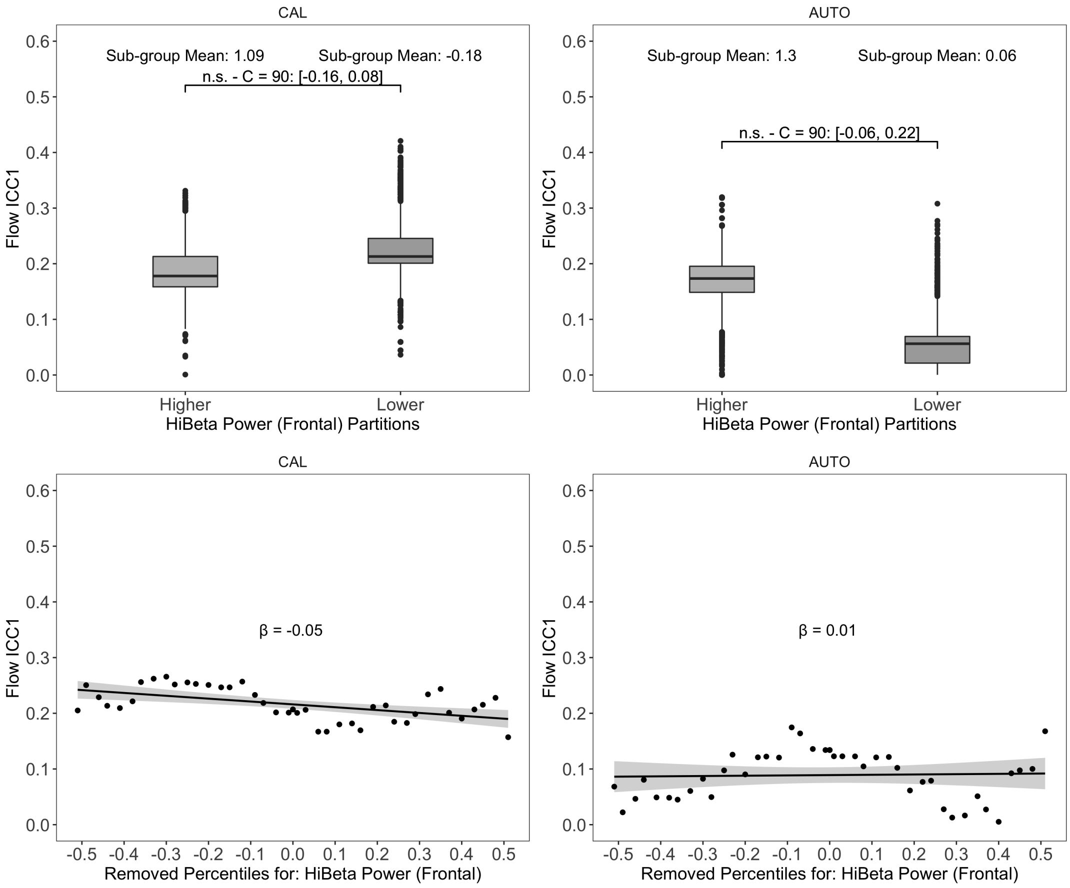

FIGURE 7.12: Flow ICC1 Estimates After Median Splits on (Group Average) HiBeta at Frontal Electrode Positions in CAL and AUTO Optimal Difficulty Conditions in Experiment 2. Top: Leave-Three-Groups-Out Resampled ICC Estimates. Bottom: Flow ICCs After Stepwise Removal of Lowest and Highest Percentiles.

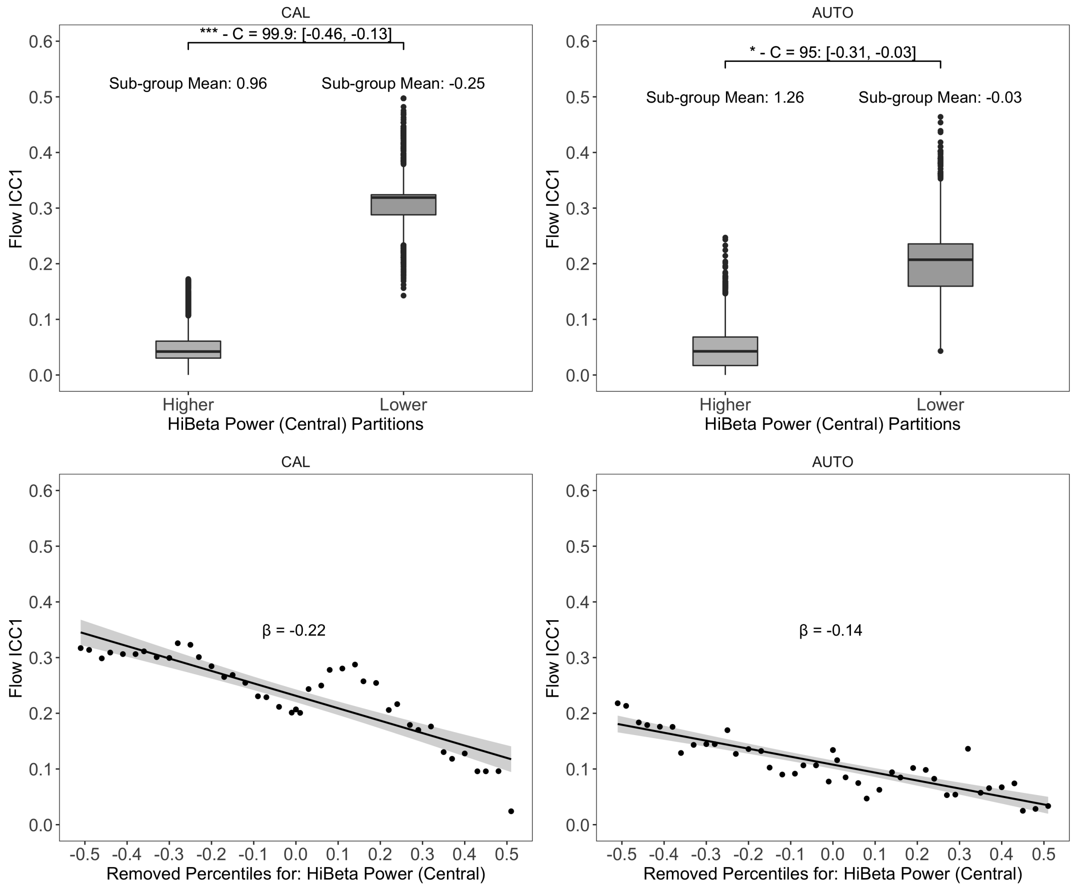

FIGURE 7.13: Flow ICC1 Estimates After Median Splits on (Group Average) HiBeta at Central Electrode Positions in CAL and AUTO Optimal Difficulty Conditions in Experiment 2. Top: Leave-Three-Groups-Out Resampled ICC Estimates. Bottom: Flow ICCs After Stepwise Removal of Lowest and Highest Percentiles.

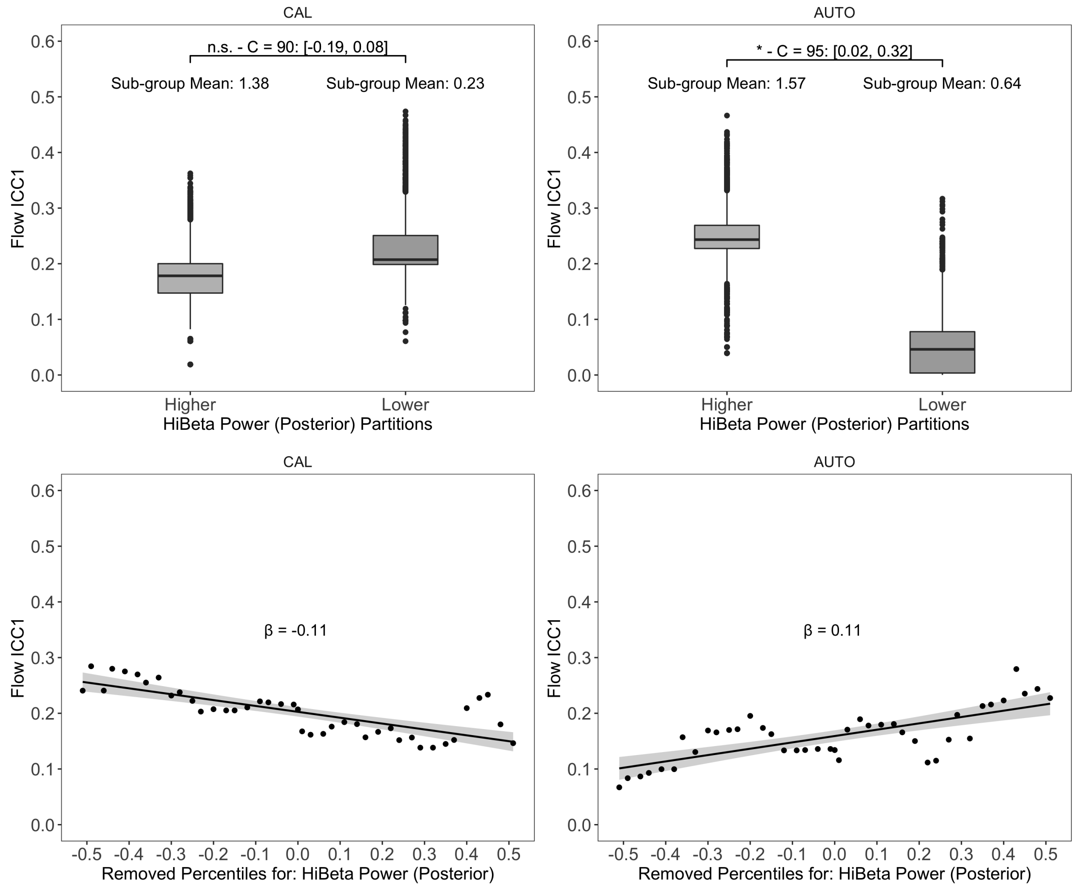

FIGURE 7.14: Flow ICC1 Estimates After Median Splits on (Group Average) HiBeta at Posterior Electrode Positions in CAL and AUTO Optimal Difficulty Conditions in Experiment 2. Top: Leave-Three-Groups-Out Resampled ICC Estimates. Bottom: Flow ICCs After Stepwise Removal of Lowest and Highest Percentiles.

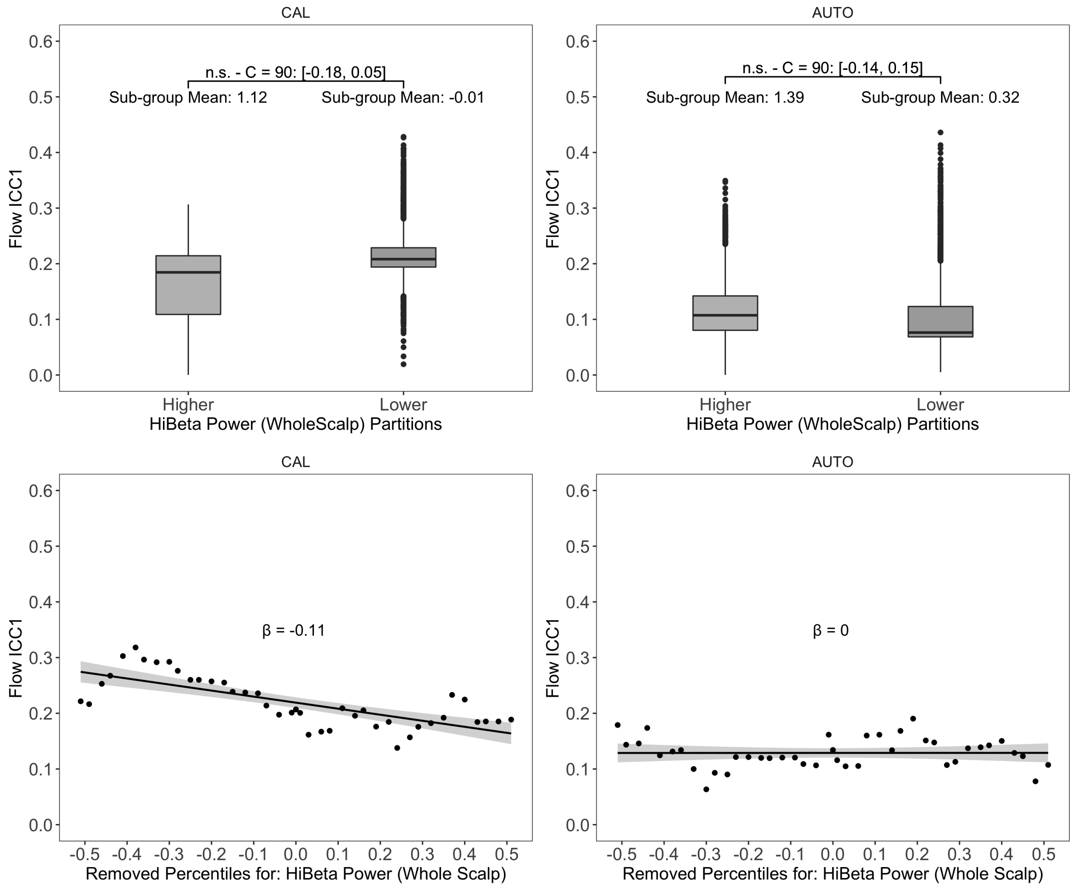

FIGURE 7.15: Flow ICC1 Estimates After Median Splits on (Group Average) HiBeta Over the Whole Scalp in CAL and AUTO Optimal Difficulty Conditions in Experiment 2. Top: Leave-Three-Groups-Out Resampled ICC Estimates. Bottom: Flow ICCs After Stepwise Removal of Lowest and Highest Percentiles.

Shared Flow Prediction

The observation of group influences on Frontal, Central, and Whole Scalp HiBeta features (see Section 6.3.3, Table 6.20) brings with it interesting possibilities. The emergence of group-level influences for this sensitive and specific mental workload feature suggests that such reciprocal influences could represent an influence on mental workload by other group members. As flow theory is firmly rooted in the argument that (elevated) optimal difficulty is required for more intense flow experiences (see Chapter 2), the theory could be extended by the proposition that (optimal) reciprocal influences of difficulty act similarly as a precondition for the emergence and intensification of shared flow experiences. However, more evidence is needed to substantiate such a proposition. To further investigate a possible relationship, two follow-up analyses were conducted. First, multivariate ICC patterns for the HiBeta features were evaluated for indirect inference. Second, based on this first analysis, a direct analysis is pursued using (group) mean aggregated flow reports as DV and HiBeta power as IV in an LMM analysis on group flow - group HiBeta relationships.

The results of the first analysis (the HiBeta ICC patterns per condition) are shown in Figure 7.16 and Figure 7.17. Both figures indicate that HiBeta levels show a similar progression as flow ICCs, namely a dissipation at high(er) task difficulty levels. Again, the Central ROI is found as the most related feature, together with the Whole Scalp HiBeta feature. These findings further indicate that flow clustering and HiBeta clustering might be related.

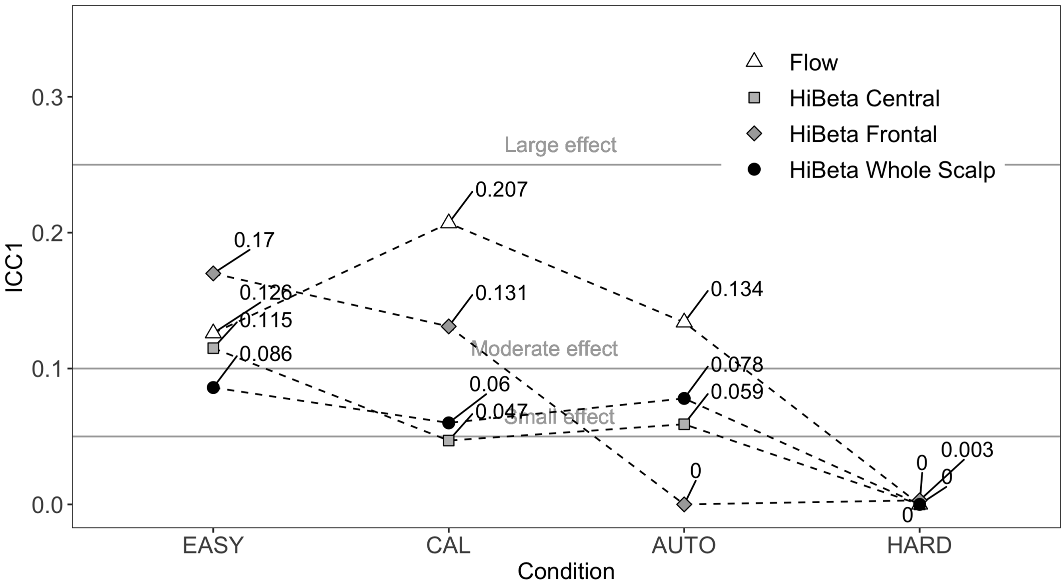

FIGURE 7.16: Multivariate Flow and HiBeta ICC1 Point Estimates Incl. Recommended Effect Size Thresholds.

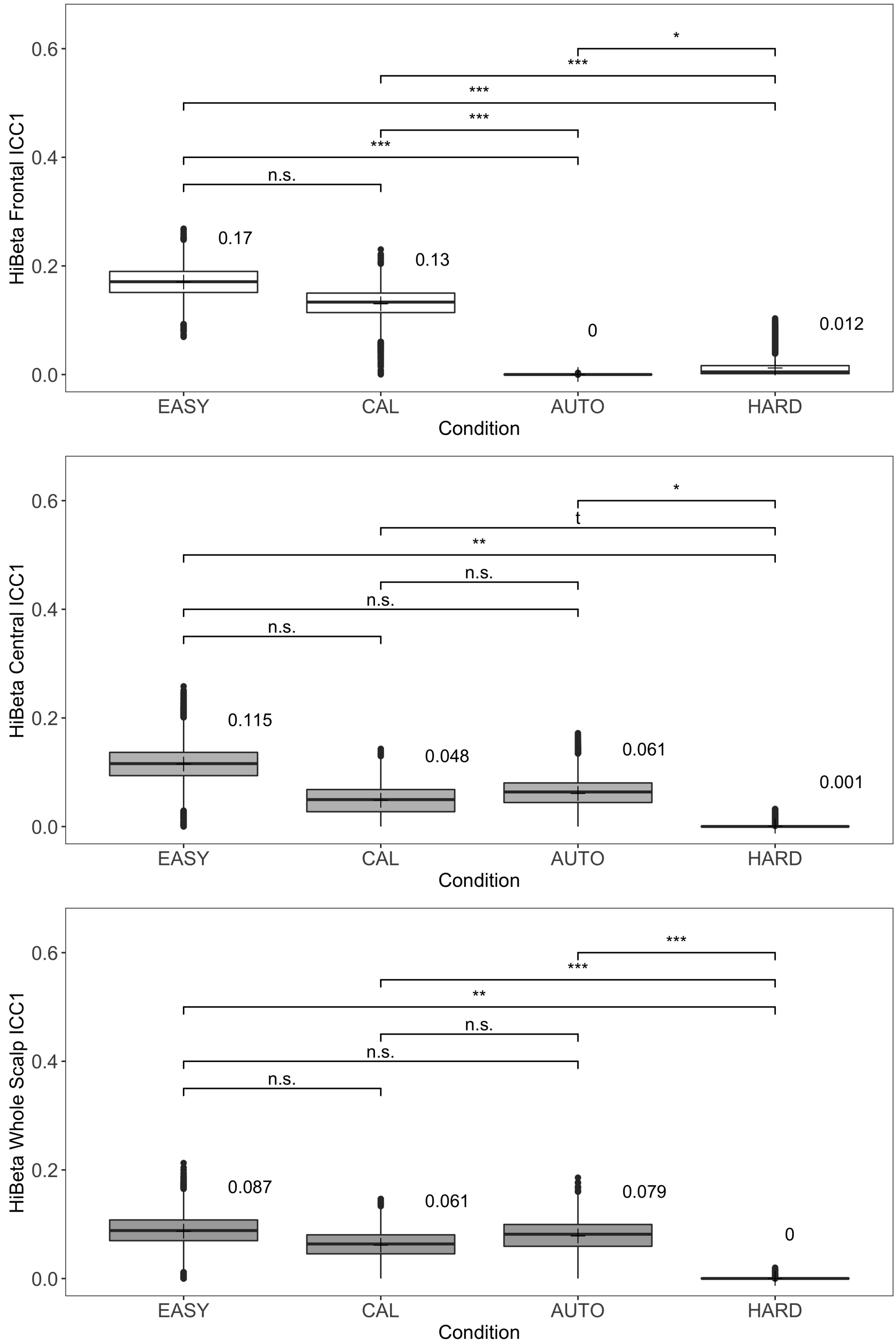

FIGURE 7.17: Leave-Three-Groups-Out Resampled Multivariate ICC1 Distributions Incl. Significance Levels of Pairwise Differences for HiBeta Features at Frontal, Central, and Whole Scalp ROIs. Crosses and Numbers Next to Them Represent Resampled Means.

Given that these first results indicate the suitability of flow and HiBeta pooling, (group) mean aggregations were computed for both types of data for all conditions in which the ICC levels were above the small effect size threshold of 0.05 (see Bliese, 2000). Afterwards, separate LMM models were developed to assess direct relationships (see Table 7.3). For Frontal and Whole Scalp ROIs, LMMs include linear (group mean) HiBeta power as predictors of (group mean) flow. For the Central ROI, a better model fit was indicated for the inclusion of orthogonal quadratic HiBeta power. All models also include group (level-2) random effects. The importance of this analysis is, that in contrast to the previous ICC analysis, analysis with group mean aggregations allow to infer the absolute level of shared flow experiences. Initially, only for the quadratic group mean Central HiBeta power, a significant relationship was found with group mean flow. However, follow-up sensitivity analyses that included only the quadratic term in the model did no longer find a significant relationship. Therefore, so far, there is no sufficient support that group mean flow can be predicted using group mean aggregations of EEG mental workload indicators.

Metrics	Group Mean Flow ~ HiBeta Frontal (excl. AUTO)	HiBeta Central	HiBeta Whole Scalp
Fixed Effects			
Intercept	4.7475 (0.0821)***	4.8183 (0.0613)***	4.7881 (0.0830)***
Power	0.0384 (0.0927)	0.4734 (0.5989)	0.0474 (0.0768)
Power^2	-	-1.3349 (0.5765)*	-
Random Effects			
Level 2 Resid. (Grp.)	0.1116*	0.0653*	0.0918**
Goodness-of-Fit			
AIC	131.9562	196.3044	204.0962
BIC	141.3314	210.0723	215.1106
Marginal R^2	0.003	0.054	0.004
Conditional R^2	0.376	0.256	0.284

Notes: ^{t}p <.1; *p <.05; **p <.01; ***p <.001; Fixed Effects Shown as Beta-Coefficient with SE; Random Effects Shown as Variance; P-Values are BH-Corrected.

TABLE 7.3: Mean-Flow Mean-HiBeta LMM in Experiment 2.

Given the previous results, it is still an interesting endeavour to investigate this possibility further. The observed ICC levels for both flow and HiBeta in this experiment are reasonably low. Therefore, intensifications of this clustering strength could bring more robust findings to light. Currently, the low ICC levels come with a lack of reliability of the group means (Bliese, 2000) that can only be overcome through such intensification. A simple approach would be to compare the digitally mediated setting to a more open F2F setting with more complex tasks, as this seems to have lead to stronger flow report clustering in previous work (see Tse et al., 2016). Together with the other findings on the relationships of flow and HiBeta, this would currently appear to be one of the best candidates to indicate such a relationship. Furthermore, it was previously observed that the present data does not include one of the more robust indicators of mental workload, that is frontal Theta power (likely due to the absence of missing midline electrodes). Thus, the proposed relationship could be investigated in such situations through additional features in other EEG measurements. The same can generally be said for the observed HiBeta features, namely that higher density EEG recordings together with spectral and spatial filters, could further identify if the suggested relationship exists. The likeliness of a relationship of reciprocal workload influences with reciprocal flow influences represents a proposition that is close to original flow theory, and has herein received initial evidence. This relationship can well be tested in future research both by improving effect sizes (and thus reliability of ICC metrics) and sensitivity of EEG feature extraction. Therefore, the derived hypothesis that shared workload might be involved with shared flow as a precondition or experiential component represents a valuable contribution to the related literature and a central one from this second Experiment in this dissertation.

7.4 Discussion

7.4.1 Central Findings

In these analyses, central gaps in the research on flow in social units were addressed that pertain to (1) the causes of flow in terms of social unit compositions and behaviours, (2) the consequences of flow in social interaction in digitally-mediated environments, and (3) the analysis of the emergence of shared flow experiences, especially paired with the investigation of multidimensional data (report, behaviour, neurophysiology). Concerning the derived research questions (that specifically focus on the observation of flow in digitally-mediated small group collaboration), novel insights emerged that need to be critically appraised.

RQ9 - Flow & Small Group Diversities

Research Question 9 asked if group composition (diversity) variables influence flow in social units in digitally-mediated interaction. Theoretical work has highlighted that dimensions like even group effort, skill integration and member diversity could be of importance for the emergence of flow in social interaction (Hout, Davis, and Weggeman, 2018; Walker, 2010). For this reason, relationships of these variables with flow experiences in small groups were expected, yet were not found (except for even effort). On the one hand, this confirms previous propositions to some degree by showing that perceptions of uneven efforts of group members appear to be linked to lower flow intensities. Two possible reasons for this are considered. One, individuals might disengage from the task when noticing that other's do not put forward similar amounts of efforts as themselves. Two, uneven effort levels might be related to imbalances in task difficulties, as some members might have to take on more (too much) task load while others might take on less (too little) of that task load. In both cases, lower flow intensities would be likely to emerge.

On the other hand, the absence of flow-diversity results is rather unexpected. Given the simplistic task and the restrictive experimental design, it is possible, that some of these variables reflecting group diversities do not play a sufficiently strong role in this situation. Consider for example, that each group member in this sample, is likely sufficiently equipped to complete the mental arithmetic task (this was an explicit consideration in the design of the experiment). Therefore, a variable like skill integration (see Hout, Davis, and Weggeman, 2018), might exhibit ceiling effects. Such effects are not present in more naturalistic, real-life settings where complex tasks require diverse and scarce abilities, and where group members must find ways to integrate them in a more dedicated manner. This coordination process is not required in the present experiment (given more similar sub-task difficulties and member abilities). However, such coordination processes are also prone to interference from technical or emotional conflict (Clarke, 2010). Thus, while the present results suggest that flow experiences in groups could in some circumstances be less related to member and skill diversities, it is also likely, that there involvements influences flow in groups in more complex, natural settings. For these reasons, it needs to be studied further when and how such relationships exist and under which circumstances their influence might be altered to achieve desirable outcomes for Knowledge Work (KW) groups. Such desirable small group interaction outcomes are in this work further confirmed to be related to flow experiences.

RQ10 - Flow & Group Interaction Experiences

Research Question 10 asked what the relationship of flow is with group performance, satisfaction and growth in digitally-mediated interaction. The results from the correlation analyses support the connection of flow to these positive experiences. Thus, the relevance to advance experimental flow research is highlighted further as these correlations position flow at least as a proxy, a representative measure for all these desirable outcomes in the KW context. Researchers have proposed, to use flow experiences as a metric for high quality work experiences (Quinn, 2005). Such a proposition receives additional support through this multitude of correlative flow relationships. This observation is in line with previous studies on flow in social interaction (see Table 3.1). It is essential to see that the digital mediation scenario does not impair the general connection to performance, satisfaction and growth outcomes for groups. However, as is the downside of correlation analyses, it further needs to be elaborated what the nature of the relationship of flow to these dimensions is. Whether or not, for example, flow is a direct cause of these dimensions would be important to identify to refine (social) flow theory further. Also, for practitioners and the development of adaptive NeuroIS, the identification of causal relationships (and possible moderating and mediating influences) remains a gap that needs to be closed to improve the quality of recommendations and to substantiate actions for behavioural changes.

RQ11 - Group-Level Flow Experiences

Research Question 11 asked how shared flow emerges in digitally-mediated interaction. First, it is found that the clustering of flow experiences (in small groups) is present at a much lower level than in related work. This finding indicates that the potential of a shared flow experience is reduced in this experiment. As potential explanations, the lack of degrees of freedom in the task processing, but also the lack of social signals and communication are identified and explored further. While the latter proposition is not supported, this does not mean that other social dynamics might not be related to shared flow experiences. Candidates for such dynamics are emotional contagion (Labonté-Lemoyne et al., 2016) and stress-buffering (Tse et al., 2016; Palumbo et al., 2017) that might be suppressed in the present experiment design, as hardly any form of task-unrelated communication could take place.

In contrast, support is found for the former proposition, as flow Intra-Class Correlation Coefficient (ICC) coefficients appear to increase when the perceived autonomy of group members increases and when the self-selected optimal difficulty is closer to all members preferred optimal difficulty. Two possible reasons are considered to explain this pattern. First, given that autonomy is known to strongly impact individuals' task engagement (a central finding in the Self-Determination Theory in social psychology - see Deci and Ryan, 1985; Ryan and Deci, 2002), lower levels of autonomy might be leading to a reduction in engagement with the group. This process would mean that individuals essentially become less interactive, and a situation of co-presence emerges, where individuals might experience more or less flow by themselves, but where they no longer influence each other and therefore not each other's flow experiences. In the sense of Walker (2010)s categorisation of (social) flow experiences, situations with lower or higher member autonomy might thus be differentiated as situations of co-located (lower autonomy and lower group engagement) or shared flow (higher autonomy and higher group engagement). Second, the previously found relationship between autonomy and optimal difficulty lends itself for an alternative explanation of the autonomy-shared flow relationship (see Section 6.3.2). Specifically, the influence of autonomy on reciprocal flow influences might be mediated by optimisation of difficulty for team members. Such difficulty optimisation could be caused on the one hand by group members that allow others more to self-select the sub-tasks that they find ideal for themselves. More freedom in optimal task selection is also evident in the groups with lower ranges of preferred difficulty. Tasks with high difficulty cause direction of attention to the self and one's task (see, e.g. Tozman and Peifer, 2016; Fairclough et al., 2013). It is, therefore, plausible to assume, that especially tasks that are too hard, might lead to isolation from the group. Again, such a shift in the interaction pattern might explain why lower autonomy might lead to a dissipation of reciprocal flow influences, as more situations with (too) hard tasks might emerge that cause group members to disengage from the interaction with the group.

The latter proposition is particularly supported by the findings on flow ICC changes with task difficulty. Across multiple analyses and data domains (reports, behaviour, and neurophysiology), it is observed that shared flow experiences appear only possible when the task is not too hard. Consequently, when a task becomes too hard, individuals are likely no longer engaging in the interaction with the group but focus solely on themselves and their task. While this means that they still can experience flow individually, the shared part of the experience is no longer possible. Not only the robustness of the finding, but the finding and the approach itself is a novelty in the related work and represents a central contribution to (social) flow theory. These findings implicate lower task difficulties as necessary preconditions for the emergence of shared flow experiences. A unique contribution is the inclusion of neurophysiological indicators of mental workload in the derivation of this hypothesis. This feature is important because it identifies cognitive dynamics related to the task difficulty variable, and because it represents a way to measure and (dis-)confirm this hypothesis in future work. That being said, it needs to be appraised that the robustness of the HiBeta feature for showing flow ICC changes was different from the other data domains. Why this is the case needs to be investigated further in future research. While the potential is indicated (especially for HiBeta power at Central electrode locations), refined extraction of frequency ranges and spatial locations should be performed to confirm the utility of EEG measures to differentiate situations of (too) high difficulty in small groups.

Beyond the observation of shared flow boundary conditions, another interesting observation emerged from this (first) analysis of neurophysiological relationships to reciprocal flow influences. In particular, ICCs of flow and HiBeta (found and known as a sensitive indicator of mental workload - see Section 6.3.3 and Michels et al., 2010) showed similar changes with manipulated difficulty. Furthermore, while the significance could not be confirmed here, a direct (quadratic) relationship was suggested between group average flow reports and group average HiBeta power at Central electrode locations. Again, given that the evidence is only indirect so far the proposition emerges, that shared flow experiences might be related to shared workload levels. More specifically, beneficial reciprocal influencing of workload levels could be related to more intense shared flow experiences and undesirable workload influencing to a reduction in (shared) flow intensities. Beneficial and undesirable workload influencing are herein considered in relation to the well-established relationship of individual difficulty and flow (Nakamura and Csikszentmihalyi, 2009), namely that when a group member optimises another's load (difficulty) their flow experience should intensify. An example for this could be that a group member for whom the task is currently too easy, decides to take a part of the workload for another group member for whom the task might currently be too hard. Resulting from this interaction, both group members flow experiences would intensify, and the act of this load optimisation through others could be part of the qualitative difference in perception of flow (i.e. because others foster one's flow and vice-versa, the situation is experienced as a shared phenomenon in the group).

While more evidence is needed to support this proposition, its plausibility is strengthened by the conceptual closeness to the original arguments in flow theory. Another strength of this proposition is also that it can be further investigated using neurophysiological measures. Therefore, it represents the first proposition of how shared flow experiences might be eventually observed using neurophysiological methods. Furthermore, this finding extends the flow theory proposition of balanced difficulty-skill entry conditions to the group level by adding the proposition that functional (i.e. towards optimal difficulty), and reciprocal workload management might represent a precondition for (stronger) shared flow experiences.

7.4.2 Limitations & Future Directions

General limitations of the experiment (i.e. sample composition, sample size, or sensor quality) have been previously discussed, as have been specific methodological issues for the physiological feature extraction (see Section 6.4.2). As the presented analyses in this chapter build on the same data, these limitations apply equally.

Therefore, the more specific limitations for these analyses are more directly related to the used methodology. First of all, it needs to be appraised, that for multiple analyses (e.g. flow with diversity metrics, flow with group experience metrics, and flow ICC median-split follow-ups), a mismatch is present for some variables in the frequency of response elicitation (i.e. some measurements were taken after each condition and others only at the end of the experiment). This mismatch required the aggregation of flow reports per participant, and thus brings the possible limitation of introducing spurious results (see Bakdash and Marusich, 2017). Future work should find a way to circumvent this issue by also collecting general flow perceptions at the end of such an experiment, together with some of the diversity metrics. On the other hand, metrics such as group effort could also be collected with a higher frequency (after each task). These adaptations may strengthen the robustness of related findings.

Limitations with aggregation levels also apply generally to the ICC analysis. The ICC is an aggregate measure that reveals information about the individual response only by comparison of all groups. While the approach of using median splits allowed to generate valuable insights into which groups might be experiencing stronger or weaker reciprocal flow influences, the metric does not allow to infer what the strength of these influences is for any single group. To overcome this limitation means to observe interacting groups with repeated measurement (one could consider a moving window repeated measures ICC assessment per group), which quickly becomes difficult using self-reports. However, self-reports are currently still required as the measure of ground truth for flow. Nevertheless, such an approach might have to be undertaken at least until more unobtrusive and more highly-frequent flow measurement options become available.

A related direction that could be explored in this regard is the research on physiological synchronisation (see, e.g. Palumbo et al., 2017). Since it has been observed, that physiological changes can synchronise amongst small group members and have been found in relation to group performances (see, e.g. Stevens et al., 2012; Stevens, Amazeen, and Likens, 2013; Berka and Stikic, 2017), it would be interesting to analyse, if, for instance, HiBeta synchronisation is related to increases in flow synchronisation (i.e. correlation over time) or clustering (i.e. repeated measures ICC). For flow, this will likely require a higher specificity in the used physiological features, which is still something that future research will have to identify (see Section 6.4.2). In the simplest form, using the ICC median split approach presented here, one could study if the synchronisation of neurophysiological indicators of mental workload is likely to be related to changes in cross-group ICCs. This idea means that physiological synchronisation could be used to perform data splits again to see first if the two metrics identify similar patterns (i.e. groups with higher synchronisation should also be the groups with higher ICCs). A final proposition for the experimental development and validation of an adaptive NeuroIS could be to combine a set of controlled and more open task situations as in Experiment 1 (e.g. letting small groups complete reference conditions with EASY, AUTO, and HARD task conditions first, followed by work on a more complex, naturalistic task). Based on the collected data from the first (restrictive) part of such an experiment, classification models could be developed that attempt to infer instances of within-group workload and flow. Such approaches can also be seen in work with (individual) meditation state staging (Hinterberger, Kamei, and Walach, 2011). In this work, the cognitive experience is first observed in a controlled form. Then experiential states are staged post-hoc for an open meditation session using highly relevant EEG features and Machine Learning (ML) models.

Finally, a central limitation for most of these analyses is the lack of experimental manipulation (except the multivariate ICC analysis by DM). This limitation means that the herein presented results are of a correlational or quasi-experimental nature (the ICC median-split analyses), which is why they need to be followed up upon by structured manipulation to confirm the results and to learn about causality directions in the identified relationships. Several manipulations would appear useful. Future work should compare more complex interaction settings to the present setting, to analyse what intensifies reciprocal flow influences. This direction can either mean the inclusion of more communication channels (to test if social dynamics like emotional contagion or stress-buffering - or other process criteria like strategising influence shared flow emergence) or presentation of a more complex task (to test if higher levels of autonomy allow more optimal difficulty selection in the small group and thus influences shared flow emergence).

To test the proposition that shared flow could be related to shared workload, an additional direction could be to manipulate if and how workloads can be shared structurally. To do this, one could maintain the present mechanisms of (partially) private information together with final answers from all group members, and could vary mechanisms for task sharing (e.g. sending and requesting vs no sharing - i.e. pre-determined sub-tasks). Importantly, these experimental manipulations could be successful in eliciting higher reciprocal flow influences, which would improve the reliability of group mean responses (Bliese, 2000). As the previous results indicated but did not sufficiently confirm that group mean flow levels might be related in a non-linear (quadratic) form to the group mean HiBeta levels, the increased mean reliability would help to evaluate this finding further. Altogether, such experiments could help to uncover not only how to theoretically refine the dynamics of flow experiences in groups (e.g. extending shared flow preconditions) but also to advance the development of adaptive NeuroIS capable of advising on how to enable or maintain the flow experiences of small groups.

7.5 Conclusion

This research is the first to investigate group-level (i.e. a shared) flow in a digitally mediated environment. In doing so, important new findings emerge that advance flow theory. First, it is found that reciprocal flow influences do emerge, but less strongly so than in related work with dyadic Face-to-face (F2F) interactions. Possible reasons for this are discussed, and initial evidence is found, that higher levels of perceived autonomy (that have been previously linked to optimal difficulty calibration) and actual better optimal difficulty calibration for group members (indicated by more similar difficulty preferences) lead to stronger reciprocal flow influencing. This effect is hypothesised to be caused by maintenance of lower levels of mental workload, which in turn might facilitate reciprocal influencing and monitoring of group member actions. The latter aspect of the hypothesis is grounded in the robust finding (emerging across difficulty manipulations, and reported, behavioural, and neurophysiological difficulty-related measurements), that with too high task difficulties (and thus too high mental workload), reciprocal flow and workload influences dissipate. Thus, when a task becomes too hard, individuals are likely no longer engaging in the interaction with the group but focus solely on themselves and their task. While this means that they still can experience flow individually, the shared part of the experience is no longer possible.

Not only the robustness of the finding, but the finding and the approach itself is a novelty in the related work and represents a central contribution in this work. These findings have important implications for flow theory, namely that preconditions of shared flow experiences could be considered to be represented by lower difficulties. Altogether the herein outlined analyses also represent the first attempt to identify neurophysiological features that might be linked to shared flow experiences. The possibility indicated from the similar patterns of flow and mental workload indicators (HiBeta), that shared flow might be related to shared workload represents an interesting new specification to the theory and warrants further research. Opportunities to conduct this additional research are outlined in the form of manipulations of social information, task complexity, and workload sharing mechanisms. These extensions can further elaborate on the ways of how future adaptive NeuroIS might be operationalised that can foster shared flow experiences.

The correlation analyses confirmed the positive, relationships between flow experience and desirable group interaction outcomes (perceived group performance, satisfaction and growth) even in the context of digitally mediated interaction. Presently, this emphasises the potential for practitioners to use flow experience surveys as a simple proxy to capture these interaction outcomes within their work teams, and to capitalise on the teambuilding boosting qualities of shared flow experiences in simple, difficulty-calibrated collaborative tasks. Altogether, these results highlight the positive impact that flow-fostering adaptive NeuroIS could have in the KW environment. Therefore, the relevance to further advance this research across additional settings and with refined methods is evident. In the following chapter, a discussion of how such future research can be supported and advanced is therefore presented to derive the essential learnings from this dissertation.

Chapter 8

General Discussion

To experience more flow has been related to individual and collective benefits, as flow is strongly linked to general well-being in life, and the strengthening of social relationships (Keeler et al., 2015; Tse, Nakamura, and Csikszentmihalyi, 2020). In the work domain, flow experiences have been related to better job performances (through increased productivity and creativity) and more worker satisfaction, leading to reduced employee turnover and shielding from burnout (Fullagar and Delle Fave, 2017; Yotsidi et al., 2018). Similarly, research on the social dimension of flow at work highlights the positive links of flow to workgroup performances, interaction satisfaction and collective efficacy development (Keith et al., 2016; Zumeta et al., 2016; de Moura Jr and Bellini, 2019). The main goal of this dissertation is to contribute to the facilitation of flow experiences by advancing the foundations for adaptive NeuroIS. This approach leverages the increasing feasibility of unobtrusive state observation through neurophysiological sensors (Blankertz et al., 2016; Seneviratne et al., 2017; Krol, Haselager, and Zander, 2019). Building on the basis of flow theory, a series of experiments were conducted that focused on overcoming central limitations in experimental flow elicitation (i.e. shallow flow contrasts through Difficulty Manipulation - DM paradigms), while observing experiences using wearable ECG and EEG sensors. Thus, the accompanying research goals and questions focused primarily on intensifying flow in controlled settings and on the possibility of combining available neurophysiological knowledge with diversified observations to identify flow from continuous physiological measures across tasks and paradigms, for individuals and small groups. To this regard, two Structured Literature Reviews (SLR) (on Peripheral Nervous System - PNS and EEG observation of flow) and two experiments (with mental arithmetic and scientific writing tasks - and four manipulations of difficulty, naturalism, autonomy, and social interaction) were conducted, and the cumulative results are herein critically reviewed. To do so, first, the insights on flow intensification are discussed, followed by an integration of insights on flow observation from wearable ECG and EEG sensors. Each section briefly discusses its background and integrates findings, and then discusses limitations and directions for future related research. Afterwards, in two sections, the broader avenue of developing adaptive NeuroIS is discussed, rooted in the current state of knowledge and appraising the ethical limitations that such adaptive systems could face in the future.

8.1 Intensification of Flow in the Laboratory

Beyond Difficulty Manipulation

The experiments in this dissertation have centrally focused on intensifying flow in laboratory settings further to enhance approaches for unobtrusive, continuous flow detection. DM, controlled Experience Sampling (cESM), Autonomy Manipulation (AM), and Social Context Manipulation (SCM) paradigms have been explored as these were found to represent the most salient candidates for flow intensification in the related work. DM was included as the current best practice (Keller, 2016), and SCM was included as perhaps the most recent candidate for flow intensification. Both approaches worked moderately well, particularly so because they elicited contrasts between lower and higher flow experiences. The reasons as to how these approaches are limited have been extensively discussed in the previous chapters. For the DM approach, the major limitation lies in the challenge of calibrating optimal difficulties that take individual preferences (for slight underload or overload) into account (Løvoll and Vittersø, 2014; Fong, Zaleski, and Leach, 2015). For the SCM approach, major limitations were likely found in the restriction of social information and the complexity of the task, both factors that are typically much more strongly present in real-world situations, even those with digitally-mediated communication. Both DM and SCM approaches have herein likely shown limitations through the more controlled experiment setup. Such control in experiments is generally seen as an advantage for the increase of internal validity of Dependent Variable (DV) measurement (Bless and Burger, 2016). With flow research, however, it has also previously been considered to hinder the emergence of deep flow experiences (Hommel, 2010; Delle Fave, Massimini, and Bassi, 2011). In both Experiment 1 and Experiment 2 of this dissertation, it appears so, that relaxation of these constraints has helped to intensify flow experiences. Specifically, in comparison, the approaches centrally characterised by the variation of autonomy (cESM in Experiment 1, AM in Experiment 2), have shown the most promising flow intensification potential.

Including Autonomy for Intensified Flow

Recommendations for future flow research herein centrally propose an integration of higher autonomy for the experiment participants. Such autonomy would ideally be integrated into an otherwise controlled environment. The two approaches that have been shown here are: (1) to allow participants to self-calibrate a task's difficulty towards an optimal level, and (2) to allow participants to bring a task that is naturally relevant to them and to self-select a challenging task goal. Similar options have already been employed in related work. For example, Barros et al. (2018) allow to self-select the difficulties in game tasks in a DM paradigm. Shearer (2016) allows participants to select from a palette of five pre-selected video games at the start of an experiment that is conducted over two sessions to include individual preferences. Manzano et al. (2010) and Harmat et al. (2011) observe musicians as they perform well-known pieces that were brought to the experiment by the participants. When higher control is required, an interesting option for flow intensification research could thus be to integrate recommendations for task complexity with participant autonomy. In some related work, the herein termed Mastery (MAS) paradigm has been employed (De Kock, 2014; Kramer, 2007; Manzano et al., 2010). The MAS paradigm focuses on repeated task execution so that participants can build expertise or so that through the repetitive process, the flow-entering process is facilitated. A significant advantage of this design is the repeated measurement itself which is a useful feature for neurophysiological variables as it capitalises on the observation of intra-individual variance over inter-individual variance. To include more expertise variation, participants could be sampled so that they come with different pre-existing expertise levels. An example of such an approach could be seen in the observation of chess players (Tozman, Zhang, and Vollmeyer, 2017). For the Knowledge Work (KW) context, this could be similarly performed with undergraduate students, graduate students, and more experienced scholars in a scientific writing task, or with computer science students and experienced programmers. Including autonomy in a MAS paradigm could be achieved by allowing participants to bring their work, or to repeatedly calibrate the difficulty of the task after each trial or interruption. For example, with the chess players, they might be allowed to self-select a weaker or stronger opponent after each game, dependent on how they feel that such an opponent best suits their preference for optimal difficulty. In the KW context, such adaptation might be included through revising self-selected goals after each trial. Such approaches not only integrate more individual preferences and are therefore likely to foster optimal difficulty calibration (see Section 6.3.2) and intrinsic motivation (autonomy is known as a strong driver of intrinsically motivated behaviour in Self-Determination Theory - see Deci and Ryan, 1985; Ryan and Deci, 2002). Also, naturalistic (i.e. closer to real-world) approaches increase the external validity of the research findings, which in itself is an important goal that flow (neurophysiology) research should focus on.

Increasing Naturalism for Intense Flow

The focus of the presented experiments on (mostly) simple laboratory tasks is thus still amongst the limitations of this dissertation. As was previously emphasised, the results stretch across experimental paradigms, yet they are centrally bound to highly simplified tasks (except for the writing task in Experiment 1). While, for example, results from flow reports have been found to converge reasonably well and have demonstrated the feasibility of these experimental paradigms, a larger variety of findings is present for neurophysiological patterns. This variety is why a central recommendation for future work on flow neurophysiology must not only be the inclusion of more externally valid scenarios but must also include more cross-task or cross-situational research. As flow is highly volatile in workplace environments (Ceja and Navarro, 2012) and as cross-task DM research has found substantial variation in physiological features across tasks (Barros et al., 2018), a conceivable path is to stepwise increase complexities and freedom in flow experiment designs. Initially, the inclusion of self-calibration of optimal difficulty in DM experiments that use more complex tasks in laboratory environments could be a promising start. Afterwards, moving towards more cESM-like research approaches that focus on (some) environmental control (e.g. using writing tasks such as in Experiment 1 or programming tasks as in Müller and Fritz, 2015) could further help to integrate previous findings and increase external validity. Eventually, such research can then be integrated with field studies (more traditional Experience Sampling - ESM) that continuously observe flow experiences using physiological measures (see, e.g. Gaggioli et al., 2013). For these approaches, it needs to be outlined, that a sufficiently high pre-existing expertise level of participants will be required to enable more intense flow elicitation. The inclusion of expertise is likely to allow for more implicit task processing which is considered to facilitate the emergence of perceptions of fluency and task absorption (see, e.g. Ullén et al., 2010; Manzano et al., 2010). This requirement has been at the top of the lists of recommendations for what to considerate to elicit highly intense flow experiences, and therefore converges well with the goals of increasing both internal and external validity of flow (neurophysiology) research.

Intensifying Flow in Small Groups

In the direction of social interaction flow research, the recommendations of increased autonomy and naturalism for flow intensification apply similarly. As was discussed before, the restrictive format of the design in Experiment 2 is likely to have limited the emergence of flow experiences both on the individual and the group-level. It will be critical for social flow research to identify further if either task-related processes (e.g. coordination or monitoring of actions) or socio-affective processes represent vital factors for the emergence of flow in social interactions. This distinction is important as the digital interaction formats that are pervading today's workplace environments might especially come with limitations to the latter kind of processes (Derks, Fischer, and Bos, 2008; Chanel and Mühl, 2015). Therefore, research on flow in digitally-mediated interactions will likely also benefit from including more naturalistic observation scenarios, such as the cooperation of knowledge workers in synchronous communication media. An exemplary approach could be to observe small groups that are working on a joint research project (e.g. working simultaneously on a document on a digital platform similar to Google Docs) and to interrupt them repeatedly to "catch group-level flow in the act". Such an approach could be integrated well with the research design in Experiment 1. If even in such more open and natural interaction formats, flow experiences are found to be similarly intense as in Experiment 2 (i.e. lower than in related work - both on the individual and group level), it could be concluded that social information that is only present through Face-to-face (F2F) interaction is an essential requirement for social flow experiences. Subsequently, the structured manipulation of socio-affective information transmission (e.g. through the inclusion of text-, voice-, or video-messaging) could then provide vital answers on how to intensify flow experiences in digitally-mediated social interaction. Such intensification is particularly required for the development of adaptive NeuroIS that are capable of fostering group-level flow experiences. The ideally accompanying measurement and analysis methods are discussed in the next section.

8.2 Neurophysiological Observation of Flow

Electrophysiological Possibilities for Flow Detection

In this dissertation, research using (wearable) neurophysiological sensors was conducted to consolidate the state of flow neurophysiology research and identify which observations can be robustly identified across experimental paradigms. The focus on (mostly) wearable sensors was placed as these are the candidates likely to be used in the KW scenarios of the future (see, e.g. Lance et al., 2012; Blankertz et al., 2016). As the basis for these observations, two SLRs consolidated the highly fragmented state of research on flow neurophysiology and highlighted that HRV measures are a highly used measure of choice for flow observation, given its relation to physiological activation and indication of parasympathetic Autonomous Nervous System (ANS) branch activity. Nevertheless, three competing hypotheses for physiological (parasympathetic) activation during flow were found (moderate activation, high activation, non-reciprocal co-activation). Similarly mixed results were identified for EEG features, namely competing views on frontal Theta, Alpha, and Beta activation that were considered to relate to moderate or high mental workload (Theta and Beta) and low or high frontal downregulation (Alpha). To the regard of these mixed results, the two presented experiments have led to a series of consolidating results and have highlighted novel directions that future research can build upon. Throughout the experiments, no features (neither in the ECG or the EEG) were found that exactly or closely mimic the variations in reported flow. Given that related work has also not uncovered such marker-features (see Chapter 4), the present results suggest, that either such markers are not identifiable with the used feature spaces (extracted and possible), or that flow elicitation is still too weak to allow for neurophysiological contrasts to emerge in these feature spaces. Both propositions are equally plausible, given the youth and limitations of current flow elicitation paradigms and measures of ground truth (flow self-reports might not yet be wholly and robustly collecting the experience). Also, the currently explored feature spaces can be considered to be at a fairly high level. This state means that features such as HRV and Theta, Alpha, and Beta frequency band power have been implicated in a wide array of more abstract concepts as attention, activation, and others (Brouwer et al., 2015; Bridwell et al., 2018; Blankertz et al., 2016). Therefore, it is questionable if they provide the necessary specificity to represent flow as individual markers.

An Evidence-Based Description of Flow in the Body and the Brain

That being said, from the present data it can be summarised that flow appears to be represented by (1) moderate physiological activation (moderate HRV - not low or high HRV as some researchers have suggested - see Harmat et al., 2011; Keller et al., 2011), (2) moderate mental workload (moderate HiBeta power - and tentatively elevated frontal Theta power - not maximal as some research has suggested - see Ewing, Fairclough, and Gilleade, 2016), and (3) by increased attentional engagement (reduced and stable frontal Alpha - not increased frontal Alpha as some research has suggested - see Léger et al., 2014; Labonté-Lemoyne et al., 2016). In addition, flow appears to be represented by an absence of variation in approach-avoidance motivation or affective valence (as indicated by the absence of Frontal Alpha Asymmetry - FAA changes). The latter is an interesting finding, as it indicates that flow is more likely a state of affective or motivational neutrality, likely explained by the absence of self-monitoring and self-evaluative processes (see, e.g. Sadlo, 2016; Harris, Vine, and Wilson, 2017b). It is possible, however, that affective connotations only arise after the task is completed and self-evaluative processes emerge again. To study these temporal dynamics of FAA could, therefore, be an exciting direction for flow EEG research.

Altogether, it is important to note that the aforementioned neurophysiological results emerge through the inclusion of various mechanisms for the elicitation of flow experiences in the laboratory (DM, cESM, AM, and SCM), which represents the major contribution of this work to the flow neurophysiology literature. Of particular relevance is the finding that through frequency band personalisation and sub-segmentation, some previous findings could be consolidated (specifically, frontal Alpha reduction), and some promising new opportunities emerged. Specifically, the frequency band segmentation highlighted the particular sensitivity of the HiBeta frequency ranges with manipulations of difficulty. The additional absence of confounds with time, and the group level influence on HiBeta levels, further indicate that these higher frequency ranges could have a valuable role for the observation of flow on the individual and group level. While a connection of Beta powers to flow is not entirely new, its sensitivity and emergence over a wider area of the scalp make it a promising feature to be leveraged in eventual adaptive NeuroIS (see next section). Importantly, these ranges were here found for the first time to show an indirect relationship to shared flow experiences (possibly as indicators of reciprocal workload influences of group members). Further evaluation of this relationship will likely have to venture into the domain of neurophysiological synchronisation research (see, e.g. Palumbo et al., 2017; Stevens, Amazeen, and Likens, 2013; Berka and Stikic, 2017). In synchronisation research, is has been observed that neurophysiological signals of multiple individuals occasionally show correlations over time and that such synchronisations can relate to superior group performances and social processes like emotional contagion (Labonté-Lemoyne et al., 2016) or stress-buffering (Tse et al., 2016). To enable such synchronisation analyses with EEG data, hyperscanning protocols will be required that allow the time-locking of collected signals in a highly precise manner (on the level of milliseconds - see, e.g. Toppi et al., 2016).

Presently, it is primarily argued that the observed patterns allow discussing flow-related changes in a refined manner and that they pose interesting alternatives for the detection of situations of optimal difficulty. The potential is especially given in scenarios where less information might be available than typically is in laboratory setups (i.e. with wearable EEG with fewer and unevenly distributed electrodes). Ideally, by finer spectral and spatial EEG power analysis, future flow research will move even closer to identifying robust concomitants and markers of flow that can be employed in adaptive NeuroIS using portable EEG in real-world scenarios. Nevertheless, the presented results need further confirmation through additional experiments with varied tasks and task formats. Besides, future work should employ a set of more refined data-driven feature extraction and selection methods.

Data-Driven Methods for Future Work

In this work, primarily a priori decision-making was involved together with traditional methods of statistical inference. While this approach grounds the developed knowledge in scientific best practices, especially in the field of psychophysiology, it can be argued that such deterministic approaches are generally flawed when attempting to isolate patterns across domains of reported and physiological data (Bridwell et al., 2018). Therefore, it is a limitation of the present approach to neglect data-driven approaches to identify neurophysiological patterns during flow, which has been argued for in related work (Rissler et al., 2018; Maier et al., 2019). In general, there is an increasing prevalence of using data-driven methods for feature extraction, selection and integration into classification models for the observation of mental states (Brouwer et al., 2015; Bridwell et al., 2018; Roy et al., 2019). Deep Learning methods have for example shown great promise to make sense of EEG signals due to their capacity to learn good feature representations from raw EEG data (Roy et al., 2019; Bridwell et al., 2018). Given that presently, a high degree of a priori decision making is involved in the analytical process in most flow neurophysiology research, more data-driven approaches might hold the valuable potential to improve the validity and accuracy of derived physiological features and should be employed in future work - if only for feature extraction and selection. In related approaches, Machine Learning (ML) methods have shown promising accuracies for the staging of different mental states. For instance, Hinterberger, Kamei, and Walach (2011) induced and classified different meditation states by using principal component analysis together with Fisher linear discriminant analysis classifiers. For the task- and person-independent classification of mental workload levels, Radüntz (2017) and Radüntz (2020) use a novel method called dual-frequency head mapping together with support vector machine classifiers. Deep learning methods have also been employed increasingly to varied affective experiences (Rouast, Adam, and Chiong, 2018). Also, convolutional neural networks, in particular, have shown promising results for the extraction of good feature representations from (often raw) EEG signals, for example, to detect workload levels (Zhang et al., 2018), depression (Acharya et al., 2018), or sleep stages (Sors et al., 2018). Leveraging these data-driven methods will likely be a critical step in the development of adaptive NeuroIS that could eventually be able to foster flow in the context of KW.

8.3 Towards Adaptive Systems

Non-Optimal Difficulty Based Adaptation

Based on theoretical and empiric research on flow elicitation, several directions are conceivable for the development of flow-facilitating adaptive systems. However, flow neurophysiology research has primarily focused on manipulations of task difficulties (see Chapter 4), as has the work in this dissertation, and has not yet provided highly diagnostic findings (or explanations on which processes underlie the emergence of flow). Thus, the most likely approaches for unobtrusive, real-time data-driven flow facilitation are systems that can either (1) propose adaptations of task difficulty/load, or (2) try to maintain flow experience, while it is ongoing. As a basis for such systems, the previously identified features could be used. For example, the combined observations of higher HRV, increased frontal Alpha power, and low HiBeta power could provide promising indicators of underload experiences. In the opposite direction, observations of reduced HRV, low frontal Alpha power, and high HiBeta power, could provide promising indicators of overload experiences. ECG, EDA, and EEG-feature based ML models have already provided initial support for the possibilities to differentiate situations of underload, balanced load, and overload (Chanel et al., 2011; Berta et al., 2013; Chatterjee, Sinha, and Sinha, 2016; Sinha et al., 2015), albeit with moderate accuracies. Given that future work increases the accuracies and robustness of such classifiers, mechanisms to increase or decrease task load (dependent on the task and user-situation) are conceivable.

Underload-Adaptive Systems

If the task or the environment allows, underload-adapting systems could attempt to directly increase the difficulty of the task (in a similar manner to how games adapt the difficulty for a player that might be performing too well - see, e.g. Ewing, Fairclough, and Gilleade, 2016). A work-context example for such adaptation can be seen with air traffic controllers (Prinzel et al., 2000), or with invoice processing situations (Rissler et al., 2018), where a system could increase the task difficulty or load by increasing the task throughput (how many subtasks a user has to handle in a given time frame). Similarly, underload-adaptive systems could signal to group members that capacities are available so that work subtasks can be directed to the individual with idle capacities. In the realm of KW, tasks and environments are highly complex and are not easily integrated into a simplistic IT-based system that can adjust difficulties, which makes the latter approach a more likely scenario. However, beyond placing the control into other users, underload-adaptive systems could challenge the system user or user groups to increase the difficulty or the load of the task themselves (see, e.g. Ewing, Fairclough, and Gilleade, 2016; Gilleade, Dix, and Allanson, 2005). For example, systems could recommend the setting of more challenging task goals (for oneself and or others). Such a message could recommend perspective-taking to urge the user(s) to consider the current task to be more challenging in some way. Systems could also provide the user with more freedom or autonomy in using the system. By making more complex tools available or by hinting that devising more intricate means of how to complete a task (perhaps instead of doing something manually, a user might want to learn how to write a simple program that takes on the task for him), the user might more readily experience to have to stretch their capacities to cope - one of the central characteristics of experiencing flow (Csikszentmihalyi, 1975). A particularly valuable benefit of the latter approach would be the supported competence development of the system user.

Beyond what is known directly from DM-based flow research, additional adaptive functionalities could be to increase arousal levels. This proposition is based on the findings, that flow is typically associated with at least a moderate (physiological) arousal, which often coincides with moderate task difficulties (Peifer, 2012; Tozman, Zhang, and Vollmeyer, 2017). If increased arousal (perhaps in the form of a proper warm-up or task preparation that is documented in the sports domain - see Swann et al., 2012), indeed represents another facilitator to flow experiences, underload--adaptive systems could attempt to increase experienced arousal by for example automatically providing more upbeat and exciting background music. Alternatively, they might recommend to the system user, that a small degree of physical exercise or drinking some coffee might increase their arousal states. Similarly, alertness increases could be a potential facilitator of flow when underload is experienced, as it could interact with abilities to concentrate and (self-directed) adaptations of the load. Systems could, for example, try to automatically increase alertness through increasing a computer screen brightness or colour composition, as more blueish light is found to increase alertness in digital device users (Wolska et al., 2019). However, the effect that such an intervention has on flow should be evaluated beforehand.

Both the arousal and alertness increasing approaches will have to take other variables into account. For example, during the mid-day, due to circadian rhythms, it is unlikely that arousal increases are feasible as effectively as during other times of the day (see, e.g. Debus et al., 2014 with the finding that flow intensities are lower around noon). Also, in the evening, the increase of alertness through light stimulation or coffee could interfere with natural sleep patterns, which should be avoided. At this point it should be stressed, that in particular in the case of underload-adaptive systems, it needs to be made sure, that the recommended increase of difficulty and workload come only in instances of prolonged underload, as some simple task processing is probably desirable in balance with challenging task situations (Debus et al., 2014; Engeser and Schiepe-Tiska, 2012).

Overload-Adaptive Systems

In the second approach for difficulty or load adaptation, overload-adaptive systems could assist (Ewing, Fairclough, and Gilleade, 2016; Gilleade, Dix, and Allanson, 2005) the system user by directly decreasing the task load or difficulty if possible. Analogous to the previously discussed options for underload-adaptive systems, the decrease of subtasks an individual has to process could be re-directed by the system to computerised agents, to other workers or to a later point in time. Again, what appears most likely usable for the moment would be that systems enable users to self-regulate the perceived task difficulties/load. Systems could recommend that a restructuring of goals might be in order to better cope with the task. Such recommendations are made in adaptive meta-cognitions research (Wilson and Moneta, 2016). For example, a user that finds a task much too difficult could benefit from the proposition that breaking the task into smaller pieces will make it much more manageable. Such a flexible goal structuring has also been proposed in research on how writers experience flow as a, particularly useful flow-facilitating mechanism (Flower and Hayes, 1981). Similar mechanisms could be proposed to work groups. Systems could also attempt to recommend more preparation or utilisation of simpler tools for the task, a way to build competence first to facilitate flow both during this preparatory phase and the postponed task completion later. This idea is derived from the sports literature, where preparation in the form of increased pre-competition exercise is found to increase flow during the preparation and the competition phase (Brunner and Schueler, 2009). While this requires a supportive environment (where towering deadlines do not impede growth), the benefit could be very high in the long-term, as again, not only is worker well-being enhanced by reducing adverse/stressful events, but by fostering competence development and growth.

Again, with the often referenced relation between arousal and flow, related flow facilitation options seem plausible. However, at this point, a particular subtlety factors in, that is the overload-arousal relationship. Some research points to arousal being high during overload instances, for example, when an individual experiences stress and threat and find it hard to cope with the present situation. Such considerations emerge from integrations of flow and stress research (Peifer, 2012; Tozman, Zhang, and Vollmeyer, 2017). The detriments of increased arousal are also documented in Peifer et al. (2015), where study participants received intranasal injections of cortisol, which reduced their flow experiences in comparison to a control group. On the other hand, in situations where task difficulties are excessively high, the individual might also detach from the task and experience even low arousal, when it is felt that tasks demands cannot be met. This disengagement is described in the integrations of flow theory with the shark-fin curve of the motivational intensity model (Ewing, Fairclough, and Gilleade, 2016). Considering it a progression of high, and stressful demands, to excessive and unmanageable demands, only the former position would be a sensible aspect for flow facilitation approaches. If high arousal indeed represents a limitation to experience flow, overload-adaptive systems could attempt to reduce experience arousal, by, for example, playing soothing background music, or by recommending that the user takes a short break to relax. Another option could be to integrate real-time interventions for relaxation, for example, breathing or meditation techniques, supported by physiological measurement (see, e.g. Chin and Kales, 2019). In this regard, it has been found, that HRV-biofeedback (Loudon and Zampelis, 2017), and EEG-neurofeedback (Gruzelier et al., 2010) targeted at eliciting calmer states can lead to more intense flow experiences. To induce relaxation to facilitate flow in instances of overload, would also appear to be supported by findings that more recovered individuals find it easier to experience flow throughout the workday (Debus et al., 2014). As bio- and neurofeedback methods are more becoming readily implemented in NeuroIS (Lux et al., 2018), such approaches could be implemented soon, even for small groups (Knierim et al., 2017a).

For overload-adaptive systems, situations in which thinking and alertness are rather high, yet not channelled towards the main task (e.g. during rumination and multi-tasking) could also be an interesting flow-facilitation approach. It would seem plausible that such situations could coincide with situations of high task difficulty/load, but more research is needed for that. It has however been found, that unfinished tasks, which can lead to increased rumination, appear to decrease flow experiences (Peifer and Zipp, 2019), as does an increased tendency for multi-tasking (Peifer et al., 2019). If such instances can be detected, overload-adaptive systems might, for instance, assist the user by (1) isolating a single task on a user interface and blocking other task instances or sources of information (for example the blocking of incoming messages, especially when not task-relevant has been found to increase flow - see Rissler et al., 2018), and perhaps (2) locking on to a current task (that can be completed in a short time), to facilitate task concentration and completion. The latter approaches could also conceivably directly reduce perceived task difficulty and load when it is, for example, caused by attempts to manage too many things at once.

Similar to underload-adaptive systems, overload-adaptive systems, will have to work with sensitive thresholds and will, in particular, have to work with personalised information (a promising option are therefore dynamic thresholds - see Karran et al., 2019). This personalisation is especially important as the nature of the perceived difficulty-skill balance is sensitive to individuals' preferences for slight underload and overload to experience flow (Fong, Zaleski, and Leach, 2015; Løvoll and Vittersø, 2014). Again, some overload could be acceptable in some instances and might occasionally be needed to cause action. However, especially as ailments such as stress and burnout appear to be on the rise in recent years (Berg-Beckhoff, Nielsen, and Ladekjær Larsen, 2017), it would appear particularly fruitful if adaptive systems could provide a way to less excessive task demands, that foster the growth, performance, and well-being of knowledge workers in the short-term and long-term.

Flow-Stabilising Systems

In the last direction to flow-adaptive systems, some aggregated findings from this dissertation and related literature point to potentials to isolate and detect flow experience while they are most likely occurring. For example, observations of moderate HRV, low frontal Alpha power and moderate HiBeta power could point to situations of optimal task difficulties. In a related manner, DM-paradigm based ML research finds electrophysiological data to allow for the detection of flow experiences (contrasted with underload or overload), with moderate accuracy (Chanel et al., 2011; Berta et al., 2013; Chatterjee, Sinha, and Sinha, 2016; Sinha et al., 2015). Also, ML models from studies that manipulate interruptions (Rissler et al., 2018), observe work progress (Müller and Fritz, 2015), or gaming experience (Shearer, 2016), provide reasonably accurate predictions of lower or higher flow intensities. These models hardly allow for explanations or recommendations for system users on how to alter their situation to experience more flow (except for the interruption research). However, all of these models provide opportunities to stabilise ongoing flow, which is a major challenge on its own, as, flow is often found to be a fleeting phenomenon, that can occur and vanish from one moment to the next, particularly in work environments (Ceja and Navarro, 2012). It still stands to assess empirically, how a flow-stabilising system should correctly be used to keep a flow experience alive. The difficulty is that flow experiences might be entangled with many different personal, situational, and environmental factors. Reduction of interruptions by decreasing message throughput might be useful in instances where a worker is operating in isolation. However, a group that is currently working in a highly communicative process might be severely disrupted in their flow experiences when alterations to the message throughput occur.

A general recommendation for a flow-stabilising system would therefore presently be two-fold. First, a flow-stabilising system should take care, to perturb the active configuration of the system as little as possible. This mechanism means that it should shield from the introduction of outside variables. With the context of the two aforementioned examples, a flow support system might want to block e-mails for the worker operating in isolation, and might for the teamwork only block messages from outside-team members. Second, as clear feedback is a central precondition to flow experience, a flow-stabilising system might facilitate flow through providing positive, reinforcing feedback, that flow is currently occurring. By such means, the system user might not only receive additional feedback, that he appears to be on the right track of how he is doing things, but could also be subtly made aware that he should just continue onwards, that is, not to perturb the present configuration themselves. Potential for such feedback transmission would be through channels that are otherwise not as often occupied, for example through a haptic stimulation pattern of a smartwatch (see, e.g. Shull and Damian, 2015; Azevedo et al., 2017 with recent work on how subtle haptic feedback can provide a stress relief or information sub-consciously).

Technical Foundations of Flow-Adaptive NeuroIS

Regardless of the mechanisms of how flow experiences are supposed to be fostered, a flow-facilitating system must take the user experience into account. IS research has well established, that inconvenient IT-based systems (e.g. with low ease-of-use) are unlikely to be adopted by users (Davis, 1985). This convenience factor is a central reason, why wearable sensors are of high interest to IS researchers, given that they promise low intrusion while delivering highly frequent and (ideally) highly relevant information for system adaptation (Seneviratne et al., 2017). In the present research on flow experience observation, EEG measures have been identified as most useful for the sake of observing flow boundary conditions for individuals and teams. However, acquisition of robust, high-quality EEG data is presently still a challenge for many wearable devices (Lance et al., 2012; Blankertz et al., 2016), despite considerable advances in this direction (see, e.g. Guger et al., 2012; Mullen et al., 2015). Motion artefacts represent a complication, as they can easily lead to shifts in electrodes, which worsens signal-to-noise ratios substantially, mainly when dry electrodes are used. For this reason, gel-based electrode systems are still a standard recommendation for the work with EEG data (Cohen, 2014; Teplan, 2002).

In the context of KW, in some instances, for example, when knowledge workers perform their tasks at a desk in isolation, this complication might be less of an issue. This pragmatic circumstance is why it was previously discussed that cESM flow research is promising, as it converges with these measurement requirements. However, as situations become more complex (e.g. work environments where interactions with colleagues take place), the data quality limitation becomes more severe. In such situations, another problem for EEG data collection might emerge, that is the form factor of EEG systems. In many instances, EEG systems come with large, not particularly aesthetically pleasing electrode caps. This nuisance can become a major issue as individuals might not feel they want to wear such a large contraption at work at all. Therefore, wearable, appealing EEG systems are being designed that target a broader consumer audience (David Hairston et al., 2014; Ratti et al., 2017).

In the present work, one such EEG device - the Emotiv Epoc+ was utilised to collect data in a scenario that might suit future workplace settings. Future work might want to take this a step further while maintaining research-grade data quality. Such an endeavour is currently supported by a novel system for EEG recording around the ear (Debener et al., 2015). The benefits of this system are, that it can be used dry or with the application of gel and that it is usable for multiple-hour recordings without participants remarking about major discomfort (Debener et al., 2015). The comfort is realised by a flexible, printed EEG electrode array that can be attached around both ears, and that fits individual physique (see Figure 8.1). The small form factor also highlights that the electrodes can be easily concealed, due to the c-shaped form around the ear (hence the term cEEGrid). This ear-EEG system has been found to also reliably differentiate posterior Alpha changes during resting states (Debener et al., 2015), to capture eye blinks and movements (Bleichner and Debener, 2017), heart rate components (i.e. an ECG recording within the EEG signal) (Bleichner and Debener, 2017), and to differentiate between different levels of mental workload in prolonged car driving simulations (Wascher et al., 2019). Especially the two latter findings, therefore, make the c-Shaped Electro-Encephalpography Grid (cEEGrid) system a promising candidate for the continuous observation of flow boundary conditions. Consider, that Experiment 2 in this dissertation has particularly found workload indicators (HiBeta power at FC and T electrode sites) to differentiate these boundary conditions and to be directly related to reported flow (see Section 6.3.3). The electrode positions of the cEEGrid system could therefore well be able to collect similar signals given the closeness to temporal Regions of Interest (ROI). Together with the concealed form factor, this ear-based EEG system might, therefore, have the potential to make continuous, unobtrusive EEG recordings of mental workload socially available and functional for flow-adaptive systems. Future work will, however, have to evaluate this potential, ideally first by employing it in cESM designs and working with elaborate feature extraction and ML methods to ascertain its utility.

FIGURE 8.1: cEEGrid - A Flexible EEG Array With Ag/AgCl Electrodes That Can Be Placed Around the Ear Using Adhesives.

Outlook on Flow-Adaptive Systems

This paragraph concludes the recommendations that presently seem most justified by the empirical research basis on (neurophysiological) flow detection based on task difficulties and loads. Indeed, with more emerging research that is based on other experimental paradigms and that can incorporate more measures on flow and related constructs, other opportunities for flow-adaptive systems will be conceivable. For example, as there are recommendations in the flow literature, that cognitive-affective strategies like the facilitation of concentration or positive thinking could facilitate flow (Swann et al., 2012), systems that specifically observe such dimensions (i.e. attention and emotion dynamics), could additionally implement interventions.

Lastly, it should also be highlighted, that only in recent years, the first studies using neurostimulation methods have emerged, a method that could readily become a part of flow-facilitating systems. So far, research using transcutaneous vagus nerve stimulation (Colzato, Wolters, and Peifer, 2018) has found that flow can be inhibited through neurostimulation, can have no effect (when using transcranial magnetic stimulation - see Ulrich et al., 2018), or can increase flow experience in both novices and experts when compared to sham stimulation (when using transcranial direct current stimulation - see Gold and Ciorciari, 2019). The latter appears especially promising, as transcranial direct current stimulation has so far been found to cause no adverse effects in humans in large-scale meta-analytic studies (Aparicio et al., 2016; Bikson et al., 2016; Nikolin et al., 2018). However, given the invasiveness of neurostimulation and intricacies of neural processes (Ulrich et al., 2018), flow intervention research might for now want to leave the neurostimulation approaches to scientists using it as a tool to improve our understanding of flow neurophysiology. In contrast, flow intervention research should primarily focus on how adaptive systems interact with the self-regulation processes of individuals, beyond prescribing how tools should be used.

In general, it is highly likely that due to inter- and intra-individual variance, and due to contextual factors contributing to flow experiences, flow-adaptive systems should be continuously adapted and calibrated by the respective user. For example, interactive ML processes would appear useful, that achieve higher system efficacy by being tuned in a reciprocal information provision - parameter adjusting manner together with the system user (Yannakakis, Cowie, and Busso, 2018). In the effort to achieve a future of KW, that more often provides highly satisfying experiences and peak performances, facilitating flow seems highly desirable. In this chapter, directions for how digital systems that utilise unobtrusive means of automated data collection and processing, have been outlined that could be used to facilitate or stabilise flow. However, as desirable as the continuation of this work might appear, it also needs to be discussed, what the limitations of flow facilitation are. For this reason, in the next section, these limitations, in terms of physiological and ethical ramifications of flow facilitation systems, are addressed.

8.4 Ethical Limitations of Fostering Flow

The Morality and Dark Potentials of Flow

When discussing the facilitation of flow, the scholarly tenor appears to be, that increases in daily flow experience intensity and frequency are desirable goals. This valuation is particularly fuelled by the highly salient early work of Csikszentmihalyi, which describes peak experiences, with optimal performances and high degrees of self-satisfaction, that scholars and laypeople can relate to (Engeser and Schiepe-Tiska, 2012; Schüler, 2012). This view is epitomised by flow states being named "optimal experiences" (Engeser and Schiepe-Tiska, 2012). Csikszentmihalyi stated that: "Experiencing it [flow] more often does make life more satisfying, and should prevent a person from living a dull life. It is a thrill that stands out from routine and uneventful times." (Csikszentmihalyi, 1997, p. 97 - words in brackets are added here for better understandability). However, it should be noted, that while this tendency to promote flow is dominant (Engeser and Schiepe-Tiska, 2012; Schüler, 2012), there is neither a normative nor should there be a positive perspective that flow is entirely desirable in every given situation or to achieve every conceivable goal (Engeser and Schiepe-Tiska, 2012). This neutrality is because flow is without morality - it is in itself neither good nor bad (Partington, Partington, and Olivier, 2009). Therefore, a discussion regarding the facilitation of flow should not be one-sided in favour of the reported flow benefits only. Instead, critical appraisals of some of the dark potentials of flow are due (Engeser and Schiepe-Tiska, 2012; Schüler, 2012; Wilson, 2016). Flow can greatly improve human lives, but it also has the potential to make them worse.

The so-called dark potentials of flow revolve about the characteristic of the deep involvement and intrinsic reward of flow. Schüler (2012) discusses in detail why these characteristics also have the potential to cause adverse effects. For example, by focusing solely on a task, and subsequently losing self-awareness, conflicts can arise between pursuing the task and one's own or other people's goals, needs, and values (Schüler, 2012). Therefore, flow has been discussed to lead to increased risk-taking (Stranger, 1999; Partington, Partington, and Olivier, 2009; Schüler and Pfenninger, 2011), sometimes at the cost of the individuals physical, psychological and social functioning (Keller et al., 2011; Schüler, 2012) to the degree of addictive behaviour (Partington, Partington, and Olivier, 2009; Dixon et al., 2019; Ross and Keiser, 2014), and even to a fostering of anti-social behaviour (Harari, 2008). An extreme example of the latter is the proposition that underlying some fascist regimes, is the provision of a game plan that sets clear goals and clarifies feedback. In doing so, a renewed involvement with life is fostered that many followers of such regimes may find to be a relief from prior anxieties and frustrations (Engeser and Schiepe-Tiska, 2012). In a similar direction, flow experience in soldierly combat has been documented as a thrilling experience that could be facilitating survival, but also the killing of others (Harari, 2008). In another dangerous activity domain, namely extreme sports, flow has been discussed as a reason for why individuals might pursue even greater risks to increase the challenge level required to enter more intense flow (Stranger, 1999; Partington, Partington, and Olivier, 2009; Schüler and Pfenninger, 2011). This rewarding characteristic of flow could, therefore, attract psychological and especially physiological damage (Partington, Partington, and Olivier, 2009). Observations for this potential connection of flow to risk-taking have been made in the case of big-wave surfing (Partington, Partington, and Olivier, 2009), white-water kayaking (Schüler and Pfenninger, 2011), and motorcycling (Rheinberg, 1991; Sato, 1988). Furthermore, by the experience of high control and the absence of anxiety, risky decisions can be made due to the overestimation of one's abilities.

The matter of physiological damage has also been discussed by flow physiology research. As has been previously discussed, some authors find high levels of physiological activation (in the form of strongly decreased HRV - e.g. Keller et al., 2011), which could indicate, that flow places a particular physiological burden on the body. Such an observation is deemed plausible, as the mere challenge-skill balance precondition of flow describes the necessity to stretch one's skills to grow and match the task demands that are just below what the individual can usually cope with. While several other studies refute this proposition arguing that flow is more strongly represented by moderate physiological demand (Peifer et al., 2014; Peifer et al., 2015; Harmat et al., 2011; Manzano et al., 2010; Harmat et al., 2015), the question as to whether flow can also lead to exhaustion is yet to be answered. Importantly, many of these studies typically employ relatively short experimental designs, in which tasks only last for up to five minutes per condition. Also, in the discussion of problem gambling and gaming, it has been discussed that flow could lead to a neglect of social and physiological needs (eating, sleeping) (Partington, Partington, and Olivier, 2009; Schüler, 2012; Murch, Chu, and Clark, 2017; Dixon et al., 2019). In this regard, the important question is being asked, whether or not flow can be related to addictive behaviour, with some research finding indirect connections through personality variables (Partington, Partington, and Olivier, 2009), others finding a direct positive correlative connection (Lee, Aiken, and Hung, 2012; Murch, Chu, and Clark, 2017; Dixon et al., 2019), and yet others proposing that flow and addictions could be inversely related (Ross and Keiser, 2014; Wan and Chiou, 2006). So far, it is not yet evident, if a relationship between flow and addictions exists, or what its nature is (Wilson, 2016; Schüler, 2012). These initial reports still only represent a very sparse theoretical and empirical basis (Schüler, 2012), and given the divergence of findings, it is yet to be seen, how the dark potentials of flow manifest. However, such work importantly highlights, that flow experiences can also have downsides, which need to be appraised and researched, especially when one is considering to develop flow-facilitating interventions (Schüler, 2012; Wilson, 2016).

Flow Diversity and Developing Flow Metacognitions

At the very least, caution must be exerted moving research into flow-facilitation forward. First of all, more (and rigorous) research is needed that produces theoretical extensions and models, explaining how the issues mentioned above (in particular risk-taking, exhaustion, and addiction) are connected to flow experiences. A central limitation in much of the existent research is the conceptual ambiguity and the operationalisation of constructs. It has not always the case (nor the goal), that behavioural or self-report instruments captured the breadth and complexity of flow and other phenomena (Schüler, 2012). However, such rigour will be critical in such a delicate matter, the determination of which harmful effects might be fuelled by fostering flow. In particular, the study of flow neurophysiology has for this purpose deemed an important direction (Schüler, 2012), as it could allow to more directly integrate knowledge on flow with - for example - biomarkers of addictive tendencies. Second of all, three recommendations can already be made for the facilitation of flow that integrate the dark potentials. Regarding the prevention of exhaustion, interventions should make sure to balance the time spent in flow with time spent in relaxation, or other cognitive experiences (Engeser and Schiepe-Tiska, 2012; Peifer, 2012; Wilson, 2016). To this point, there is some evidence, that without relaxation, flow experience is less likely, perhaps even impossible. In a study with software engineers, Debus et al. (2014) find that flow during the workday is at any time lower for individuals who indicated lower relaxation at the start of the day. Regarding the prevention of (extreme) risk-taking and promotion of addictive tendencies, it has been argued, that flow interventions should in the case of experts focus on task diversification (Schüler, 2012). Flow diversification means, that for individuals that already possess a high level of proficiency in a particular task, and that seek flow, it might be wiser to focus on building new skills instead of further deepening old ones to experience flow (Schüler, 2012). In this sense, another goal to flow facilitation should be, to promote flow diversity over (extreme) flow intensity. Besides, it must be a high priority to ensure that flow-adaptive NeuroIS follow the principles outlined in research on non-addictive information systems (Kloker, 2020) to prevent adverse effects of automated flow facilitation. Lastly, regarding the prevention of other adverse consequences and the utilisation of flow in ethically questionable circumstances, flow interventions should focus on building (dark) flow meta-cognitions. This recommendation is based on Csikszentmihalyi's contention, that individuals should learn when to - and when not to - experience flow, to not only blindly follow the desire to experience the rewarding activity (Csikszentmihalyi, 1975). Therefore, a goal for flow facilitations should be to empower individuals by developing awareness and knowledge about what the bright and the dark consequences of flow can be.

Ethical Considerations for Adaptive NeuroIS

Lastly, as flow intervention work is operationalised through adaptive NeuroIS, one more ethical issue needs to be addressed, that is data privacy. Ethical issues in IT systems are not just a general issue that society is presently concerned with (see, e.g. Martin and Murphy, 2017; Mehmood et al., 2016), but one that is particularly relevant in the case of flow. Individuals' (or teams') flow (or the absence of it) could be exploited by peers, leaders or organisations (Schüler, 2012). Instances could be peers that pressure an individual to utilise his flow experience to engage in risky behaviour, or it could be a manager that decides to fire employees who show low levels of flow experience throughout their workday. It could just be the members of a work team that blame the "weakest flower" during a cooperative session for hindering the progress of the team. To address these issues, both data accessibility and data access security need to be considered (see also Knierim et al., 2017b), two dimensions that are, for example, found in generalised data quality frameworks (Wang and Strong, 1996). Accessibility refers to the rights of who can inspect collected data and control the data collection process itself (Wang and Strong, 1996). Access security then refers to the degree of the safety mechanisms in place that guarantee that the accessibility rights are upheld (Wang and Strong, 1996). To guarantee data access security, flow facilitation designers need to devise prototype-based strategies and recommendations. For IT-based systems, access security concerns are probably best handled by software engineers, knowledgeable in cryptographic best practices.

Data accessibility, on the other hand, is a more complicated issue. First of all, given the dangers of flow data misuse, the data collection, storage, processing, and review, can be subject to misuse and represent steps in the pipeline where accessibility is to be defined. On the one hand, it stands to reason that the complete pipeline and the access to it should be in the control of the end-user. This control, about which data is collected, how it is transformed and represented, and even when and by whom it is being reviewed, should empower the individual using the flow intervention and prevent misuse. Furthermore, directing this control to the end-user should foster the adoption of the technology, as control over it has been found to increase trust (Peters, Calvo, and Ryan, 2018). However, a variety of reasons complicate this simple solution that are: (1) A recommendation for some oversight through others to prevent adverse/dark flow consequences (especially as long as systems are fairly simple and cannot autonomously derive recommendations), (2) In some instances the need to share flow information with others (e.g. work teams), and (3) access of system developers to the pipeline for maintenance and research. These factors all pose some need for others to have at least occasional access to the flow data. Still, the privacy and the control of the end-user over their flow data is of the utmost importance. Therefore, a compromise is recommended here. It could be a feasible solution, to require a user, to share his/her flow data with an oversight instance (and system developers) at a regular interval in time. But, to favour the needs of the individual, he/she could retain control over what or who that instance is, for example, trusted peers, managers, or third-party data scientists and physicians. In the case of teamwork, where flow data sharing is required, the individual might retain control over how his/her data is represented within the group. For example, if he/she accepts that his/her individual information to be shown in isolation or whether he/she would rather have the data represented in an aggregated form (e.g. as the average of a group). In summary, to bridge the gaps, users could be required to accept some access to their data but could retain control over the abstraction of it and the receiving entities, thus also taking into account inter-individual preferences. This way, both the requirement to shield individuals from the dark sides of flow can be balanced with the upsides, while the individual retains most of the control about how this process is carried out.

Ultimately, this chapter is not supposed to advise against working on flow facilitation, especially as an increase of the experience in everyday lives is likely a good contribution to general well-being (Tse, Nakamura, and Csikszentmihalyi, 2020; Nakamura and Csikszentmihalyi, 2009; Moneta and Csikszentmihalyi, 1996). Instead, this chapter points out the limitations and dangerous potentials, so that they are integrated with the positive goals in future research.

Chapter 9

Conclusion

The experience of flow is described as a unique experience of complete task immersion, in which action and awareness merge, concentration feels effortless, and that is accompanied by peak performances and exhilarating satisfaction (Csikszentmihalyi, 1975). Due to numerous beneficial relationships of flow to life and work experiences (both individual and social), the facilitation of flow represents a desirable goal for scholars and organisational practitioners (de Moura Jr and Bellini, 2019). However, the facilitation of flow experiences still represents a significant challenge. The situational requirements for flow are complex and rooted in the cognitive-affective dynamics of the individual. The facilitation of concentration, alertness and recovery, the shielding from self-criticism and the balance of workload, are amongst some of the flow requirements, that are difficult to manage in today's hectic workplaces (Ceja and Navarro, 2012; Spurlin and Csikszentmihalyi, 2017; Peifer et al., 2019). Furthermore, developments such as self-directed work, and mixtures of co-present and virtual collaborations of small groups further extend this Knowledge Work (KW) complexity (Spurlin and Csikszentmihalyi, 2017; Bakker and Woerkom, 2017; Keith et al., 2016). These trends are accompanied by problematic phenomena like information overload (e.g. through high frequencies of electronic messaging), or increases in professional ambiguities due to requirements of more self-organisation (Bakker and Woerkom, 2017). These phenomena stand in contrast to flow experience requirements as they represent attention-competing stimuli, unclear goals, a lack of feedback, and the elicitation of frustration or anxiety (Spurlin and Csikszentmihalyi, 2017; Bakker and Woerkom, 2017). Altogether these developments mean, that comprehensive flow facilitation at work must revolve around a person-, task- and situation-independent approach. One such approach is the development of adaptive NeuroIS. This dissertation set out to advance the foundations for flow-facilitating adaptive Neuro-Information Systems (NeuroIS). This approach leverages the increasing feasibility of unobtrusive state observation through neurophysiological sensors (Blankertz et al., 2016; Seneviratne et al., 2017; Krol, Haselager, and Zander, 2019), yet recognises that more research is required that bridges fundamental and applied settings. To overcome this gap, a combination of more and less controlled cognitive task settings, and wearable sensor systems was put in place over two experiments. More specifically, the guiding Research Goals (RG) for this dissertation were stated as:

- **RG1:** Integrate the present body of knowledge on how neurophysiological data can be used to detect flow experiences.
- **RG2:** Identify how flow experiences can be intensified in the laboratory in cognitive tasks.
- **RG3:** Consolidate which neurophysiological patterns of flow can be detected with wearable sensors across different situations, including simplistic, naturalistic, and social interaction scenarios.

At this point, it ought to be critically reviewed if and how these goals have been achieved and what the major contributions of the presented work are.

RG1 - Integration of Knowledge

In this dissertation, two Structured Literature Reviews (SLR) have been conducted that summarise the present state of knowledge on what is known on Peripheral Nervous System (PNS) and Central Nervous System (CNS), specifically EEG observation of flow. For the PNS, it is found that a majority of work has used time- and frequency domain Heart Rate Variability (HRV) features (more specifically: parasympathetic HRV-indicators), and that diverse propositions have emerged from high HRV during flow (a sign of strong calming physiological influences during a configuration of non-reciprocal co-activation of Autonomous Nervous System - ANS branches), to moderate HRV during flow (a sign of moderate physiological activation), to low HRV during flow (a sign that flow is as similarly demanding as are stress experiences). For EEG work (more specifically: frontal regions), it was found that even higher diversity in the findings regarding flow is present, with propositions of increased or maximal frontal Theta levels (that point to moderate or high levels of mental workload), low or high frontal Alpha (that point to attentional engagement or frontal downregulation in the sense of the Transient Hypofrontality Theory - THT), and low or high frontal Beta (that were considered to point to low or high arousal or mental workload levels) during flow. In addition, it was observed that this diversity in findings is likely driven by a low degree of integration of scholarly work, as very low degrees of cross-citation are present. This fragmentation is why the structured integration of this literature is a significant contribution to flow neurophysiology literature from this dissertation.

Together with the added findings for a lack of neurophysiological study of flow experiences in small groups, and the consolidating results from two multi-paradigm experiments (see below), RG1 is considered to have been achieved. However, as a limitation to this approach, the focus on electrophysiological methods in this work certainly represents its own limitation to a feature space that may or may not hold the informative potential to identify the neurophysiological configuration during flow. For the PNS, at least more indicators of sympathetic ANS activity ought to be integrated. For the CNS, to understand which processes take place in the brain during flow, insights from hemodynamic imaging studies ought to be more extensively integrated into future reviews of the present literature.

The presented consolidation of the related work shows that the development of flow-facilitating adaptive NeuroIS must conduct more refined and integrated research. Only through the comparison of findings across measurement instruments and feature spaces, and the inclusion of results in theoretical frameworks and empirical models, a stable basis for continuous flow detection can be established. Single studies will have difficulties to achieve this goal by themselves, which is why a particular emphasis should be placed on the development of comprehensive reviews and neurophysiological data repositories. Such repositories have been created for more fundamental psychophysiological research, for example, in the form of EEG recordings during affective experiences (see, e.g. Babayan et al., 2019). Also, research in the IS discipline has highlighted the utility of integrating single studies into publicly available databases for increasing relevance and robustness of findings (Dann et al., 2019). Lastly, scholars in the field of neurophysiological signal processing have recommended the sharing of not just results, but also processing pipelines to enable continuous revision and integration of experiment results when improved methods become available (Bigdely-Shamlo et al., 2015). For flow research as well, more of such integrative efforts may fast-track the development of adaptive NeuroIS.

RG2 - Intensification of Flow in the Laboratory

Based on a summary of the best practices and latest developments in experimental flow research, multiple approaches have been pursued to achieve an intensification of flow experiences, beyond what is currently possible from the established Difficulty Manipulation (DM) paradigm. To overcome the central limitation of DM that supposedly is the solicitation of only shallow flow experiences in too deterministic (in terms of difficulty calibration), simplistic and artificial task scenarios (in terms of tasks that do not require high expertise or that do not include social interactions) (Hommel, 2010; Delle Fave, Massimini, and Bassi, 2011), a more naturalistic KW laboratory observation approach was developed that integrates controlled environments and the original flow field research method, the Experience Sampling Method (ESM). This controlled Experience Sampling (cESM) approach in which participants continued a personal, ongoing scientific writing project, indicated, that flow experiences might intensify in this task when compared to an established mental arithmetic DM paradigm. As a potential reason for this, the higher level of freedom to configure the task to one's own needs and preferences (including the setting of an adequate - i.e. optimally challenging - task goal) was identified. To follow up on this potential, in the second experiment, a manipulation of autonomy (AM) was included in the form of self-selecting one's optimal task difficulty during a similar mental arithmetic task. Furthermore, this second experiment followed up on another presently interesting direction to intensify flow in the laboratory, that is the inclusion of social interaction (SCM) (Magyaródi and Oláh, 2017; Tse et al., 2016; Walker, 2010). Hence, improvements to internal validity were considered, but also to external validity as flow experiences in KW are not only likely to occur in isolation, but also often in small groups that increasingly represent today's workplace configurations (Wuchty, Jones, and Uzzi, 2007; Keith et al., 2016). In this second experiment, the AM was found as a more effective catalyst to the intensification of flow than objective difficulty calibration approaches. This consolidated the results from the first experiment that autonomy is likely a valuable factor in intensifying flow in laboratory conditions.

This finding represents the second major contribution of this dissertation and confirms that RG2 was achieved. However, it should also be noted that the absence of flow intensification from SCM represented a novelty in the body of social flow research. Especially due to the repeated finding of autonomy influences on flow, the question arises whether or not some part of the repeatedly reported flow intensification in social interaction could be partially driven by higher levels of autonomy that might accompany the typically more complex tasks in social interactions.

For the facilitation of flow experiences, these findings highlight the need to employ more comparisons of highly controlled laboratory research with more ecologically valid scenarios. On the one hand, this direction is essential as the eventual development of a flow-adaptive NeuroIS will have to be able to function in more ecologically valid scenarios, but might require initial calibration from less confounded settings. On the other hand, the results indicate that with more autonomy, individuals appear to experience intensified flow. The elicitation of intensified flow is the cornerstone for the calibration of such systems. Altogether, it appears highly valuable to increasingly pursue the research of more naturalistic (i.e. closer to real-world) tasks in controlled settings - that is the cESM approach. For instance, experts such as designers, programmers, engineers or scholars could be observed while working on their own projects either in isolation or in groups (e.g. when using digital platforms similar to Google Docs). Doing so, controlled environments can be combined with higher task diversity and autonomy, while maintaining the benefit of eliciting intensified flow. This combination represents an efficient and feasible way to combine closer to real-world flow research with the utilisation of neurophysiological sensors.

RG3 - Flow Neurophysiology in Wearable Sensors

Across the experimental paradigms and tasks, a particular emphasis of the work in this dissertation has been on identifying potentials to unobtrusively observe flow experiences using wearable neurophysiological sensors. The focus on wearable sensors was placed as these are the candidates likely to be used in the KW scenarios of the future (see, e.g. Lance et al., 2012; Blankertz et al., 2016). To that regard, the amalgamation of findings has led to a series of consolidating and novel results that aid in formulating an overview of the brain and heart configuration during flow. From the present data, flow appears to be represented by moderate physiological activation (moderate HRV) and mental workload (moderate HiBeta power - and tentatively elevated frontal Theta power), and by increased attentional engagement (reduced and stable frontal Alpha). In addition, flow appears to be represented by an absence of variation in approach-avoidance motivation or affective valence (as indicated by the absence of Frontal Alpha Asymmetry - FAA changes).

Importantly, these results emerge through the inclusion of various mechanisms for the elicitation of flow experiences in the laboratory (DM, AM, and SCM), which represents the major contribution of this work to the flow neurophysiology literature. Of particular relevance is the finding that through frequency band personalisation and sub-segmentation promising new potentials emerged. Specifically, the frequency band segmentation highlighted the particular sensitivity of the HiBeta frequency ranges with manipulations of difficulty. An additional absence of confounds with time, and a group level influence on HiBeta power, further indicate that this higher frequency range could have a valuable role for the observation of flow on the individual and group level. While a connection of Beta powers to flow is not entirely new, its sensitivity and emergence over a wider area of the scalp make it a promising feature to be leveraged in adaptive NeuroIS in the future.

In summary, particularly the high-frequency EEG features are found to provide a diagnostic potential to not only unobtrusively identify boundary conditions for individual flow, but also for shared flow experiences of small groups - all given using wearable sensor systems. For these reasons also RG3 is considered to have been achieved. However, it should be noted that many of these results emerged from a priori determined Region of Interest (ROI) and frequency band ranges. Given that spectral and spatial features have in recent work been found to strongly contribute to more robust and personalised feature extractions (Blankertz et al., 2016; Zhang et al., 2019), future work can further improve the quality and stability of the aforementioned findings using more data-driven feature extraction and selection methods. In addition, such features can then be more readily integrated into Machine Learning (ML)-based flow classifiers, that could form the core of flow detection in adaptive NeuroIS.

Group-Level Flow Experiences

As a final contribution, it was also explored, whether or not shared flow experiences emerge in the instantiated digitally-mediated interaction format in Experiment 2, and what its drivers and mechanisms for observation are. It was found not only, that shared flow experiences do emerge in this setting (operationalised as the amount of variance in individual flow experiences that can be attributed to influences from other group members), but that they appear to be weaker than in related work that used Face-to-face (F2F) interaction formats. As the most plausible reason for this observation, disengagement from the group interaction is considered that may either be driven by crowding out of motivation or by unbalanced (i.e. non-optimal) difficulties. The latter is especially considered as it was found that reciprocal influences on flow disappear, as task difficulties become (too) hard, a finding that is visible through the experimental manipulation of difficulty, and also through median splits of report, behaviour and neurophysiological data. Most likely, hard tasks are considered to lead to a direction of attention to the self and one's task and therefore, away from interacting with the group. In a sense, group members can become isolated, which is why shared flow experiences can no longer occur. The observation that neurophysiological data might be usable to infer these situations is especially valuable as it might allow inferring robustly when a task is becoming too hard for an individual and a group, that is when flow can no longer emerge for the individual and on the group level. Furthermore, with the same neurophysiological features (HiBeta power at Central ROIs), an indirect relationship was identified between shared flow and what is most likely an indication of shared mental workload. These results are used as initial evidence, to derive a theoretic extension to flow theory, that (optimal) reciprocal influences of workloads could signify a precondition for (optimal) reciprocal influences on flow. This relationship means that actions from group members that balance out the workloads to near-optimal levels for each group member could be the driver of higher intensity shared flow experiences, a reasoning that is close to the foundations of flow experience in isolation.

Therefore, the observation of these group-level dynamics and the accompanying theoretic propositions represent a separate, major contribution from this dissertation. It should be noted, that beyond the considerations for workload and difficulty influences on reciprocal flow, it is also possible that a lack of social information might have inhibited other processes such as emotional contagion (Labonté-Lemoyne et al., 2016) or stress-buffering (Tse et al., 2016) that could be involved with the emergence of such shared flow experiences. Future work will have to elaborate on these possibilities. Importantly, for the development of adaptive NeuroIS, an exciting direction emerges through these results, that is the observation of synchronisation of neurophysiological signals for the detection of shared flow experiences. Such synchronisations have previously been reported in relation to performances of or leadership emergence in small groups (see, e.g. Stevens et al., 2012; Stevens, Amazeen, and Likens, 2013; Berka and Stikic, 2017). As KW is increasingly conducted in small groups, the integration of such multi-person data into a flow-detection system seems promising. The results presented here can be considered a starting point for this endeavour. Nevertheless, more precisely synchronised data collection with high temporal resolution (e.g. EEG hyperscanning protocols) will have to be added to consolidate these results and to enable the continuous detection of shared flow experiences.

Concluding Thoughts

In closing, this dissertation contributes to the literature on (social and neurophysiological) flow theory, and to the efforts of providing knowledge that bridges fundamental and applied settings for the development of flow-facilitating adaptive NeuroIS. Together with the discussions on how such systems may be implemented and which limitations they must acknowledge, hopefully, this work contributes a piece to the larger puzzle that is the facilitation of flow, for the benefit of positive individual, organisational and societal developments.

Appendix A

Appendix

A.1 Disclosure of Own Contributions

It is common in academic research that in the continuous discussion and exchange with other researchers, a research project is improved, shaped, evaluated, and even sometimes completely inverted. Research is to considerable extent teamwork, as in many cases, a single person would not be able to perform the data collection alone or know all available literature by heart. The same is true for presented research, many hands and heads shaped the results to small or sometimes even significant extents. These persons were attributed an appropriate credit, sometimes also resulting in a co-authorship of those papers on which this thesis is based. This section intends to constitute in detail which of the parts were performed by the author of this thesis and which parts were a joint work, in order to help the reader assess the efforts and achievements of the author's work.

Knierim et al. (2017c) is a joint article with Dr. Raphael Rissler, Prof. Dr. Verena Dorner, Dr. Mario Nadj, and Prof. Dr. Christof Weinhardt, published as a Full Paper in the Proceedings of the Ninth Retreat on NeuroIS 2017. My contributions were:

- The formulation of the proposal and research questions.
- The literature review - in particular the selection of relevant work and extraction and integration of results.
- The writing of all sections.
- The presentation of the article.

Knierim et al. (2018a) is a joint article with Dr. Raphael Rissler, Dr. Anuja Hariharan, Dr. Mario Nadj, and Prof. Dr. Christof Weinhardt, published as a Full Paper in the Proceedings of the Tenth Retreat on NeuroIS 2018. My contributions consisted of:

- The literature review.
- The formulation of the proposal and research questions.
- The theoretical foundations to suggested approaches.
- The development and implementation of the experiment.
- The data collection in the laboratory.
- The processing and analyses of the self-report data.
- The analyses of the ECG data.
- The interpretation of the results.
- The writing of all sections.
- The presentation of the article.

Knierim et al. (2018b) is a joint article with Dr. Mario Nadj, Dr. Anuja Hariharan, and Prof. Dr. Christof Weinhardt, published as a Full Paper in the Proceedings of the Fifth International Conference on Physiological Computing Systems 2018. My contributions consisted of:

- The literature review.

- The formulation of the proposal and research questions.
- The theoretical foundations to suggested approaches.
- The development and implementation of the experiment.
- The data collection in the laboratory.
- The processing and analyses of the self-report data.
- The processing and analyses of the EEG data.
- The interpretation of the results.
- The writing of all sections.
- The presentation of the article.

Knierim, Nadj, and Weinhardt (2019) is a joint article, published as a Full Paper in the Proceedings of the Third International Conference on Computer-Human Interaction Research and Applications 2019. My contributions consisted of:

- The literature review.
- The formulation of the proposal and research questions.
- The theoretical foundations to suggested approaches.
- The development and implementation of the experiment.
- The data collection in the laboratory.
- The processing and analyses of the self-report data.
- The processing and analyses of the EEG data.
- The interpretation of the results.
- The writing of all sections.
- The presentation of the article.

Knierim et al. (2019) is a joint article with Maximilian Xiling Li, Dr. Mario Nadj, and Prof. Dr. Christof Weinhardt, published as a Full Paper in the Proceedings of the Fortieth International Conference on Information Systems 2019. My contributions consisted of:

- The literature review.
- The formulation of the proposal and research questions.
- The theoretical foundations to suggested approaches.
- The development and implementation of the experiment.
- The data collection in the laboratory.
- The processing and analyses of the self-report data.
- The interpretation of the results.
- The writing of all sections.
- The presentation of the article.

A.2 Additional SLR Data

Reference Key	Reference Full	In-Degree	Out-Degree
Theory			
Mar'01	Marr (2001)	0	0
Die'03	Dietrich (2003)	2	0
Die'04	Dietrich (2004)	4	0
Web'09	Weber et al. (2009)	2	0
Van'10	Heerden (2010)	0	0
Ull'10	Ullén et al. (2010)	0	0
Die'10	Dietrich et al. (2010)	2	0
Peif'12	Peifer (2012)	0	0
Wes'12	Westcott-Baker and Weber (2012)	0	0
Cher'16	Cheron (2016)	0	0
Sad'16	Sadlo (2016)	0	0
Web'16	Weber, Huskey, and Craighead (2016)	1	0
DiD'17	Di Domenico and Ryan (2017)	0	0
Har'17	Harris, Vine, and Wilson (2017b)	0	0
EEG			
Kra'07	Kramer (2007)	3	0
Nac'10	Nacke, Grimshaw, and Lindley (2010)	1	0
Cha'11	Chanel et al. (2011)	2	0
Ber'13	Berta et al. (2013)	1	1
Fai'13	Fairclough et al. (2013)	1	0
DeK'14	De Kock (2014)	0	4
Li'14	Li et al. (2014)	0	0
Leg'14	Léger et al. (2014)	1	1
Sol'14	Soltész et al. (2014)	0	2
Bey'15	Beyer et al. (2015)	0	0
Joh'15	Johnson et al. (2015)	0	1
Sin'15	Sinha et al. (2015)	1	3
Wol'15	Wolf et al. (2015)	0	0
Cha'16	Chatterjee, Sinha, and Sinha (2016)	0	2
Ewi'16	Ewing, Fairclough, and Gilleade (2016)	0	2
Lab'16	Labonté-Lemoyne et al. (2016)	0	1
She'16	Shearer (2016)	0	4
Kla'17	Klarkowski (2017)	0	3
Bom'18	Bombeke et al. (2018)	0	3
Kat'18	Katahira et al. (2018)	0	3
Other Neuroimaging			
Kla'11	Klasen et al. (2011)	2	0

Reference Key	Reference Full	In-Degree	Out-Degree
DeM'13	Manzano et al. (2013)	0	0
Afe'14	Afergan et al. (2014)	0	0
Ulr'14	Ulrich et al. (2014)	2	0
Yos'14	Yoshida et al. (2014)	1	0
Har'15	Harmat et al. (2015)	1	0
Ulr'16a	Ulrich, Keller, and Grön (2016b)	0	0
Ulr'16b	Ulrich, Keller, and Grön (2016a)	0	0
Bar'18	Barros et al. (2018)	0	0
Hus'18	Huskey et al. (2018)	0	0

TABLE A.1: Underlying Data for the Flow Neurophysiology Study Historiograph Shown in Section 4.3 in Figure 4.6.

A.3 Questionnaire Instruments

Item Wording(s)	Answer Options
Age (No Reference) What is your age?	Number Input
Gender (No Reference) What is your gender?	Female / Male
Handedness (Picton et al., 2000) Which one is your dominant hand?	Left / Right / Both (Ambidextrous)
First Language (No Reference) Which is your first language (that you grew up with)?	List of 144 Languages
English Language Proficiency (No Reference) 1.Please indicate the level of your English language proficiency.	(English Basic User) A1 Beginner / (English Basic User) A2 Elementary English / (English Independent User) B1 Intermediate English / (English Independent User) B2 Upper-Intermediate English / (Proficient English User) C1 Advanced English / (Proficient English User) C2 Proficiency English
2. My comprehension of the English language is advanced enough that I find no difficulty in understanding this sentence. Therefore I will click on the very first response option (at the far left) for this question.	Far left - Far right (7p)
Study Major (No Reference) In which field is your study major?	List of Options 42 options according to the OECD Fields of Science (FoS)

TABLE A.2: General Demographic Survey Items Used in Experiment 1 (Pre-Experiment Invitation Survey) and Experiment 2 (Initial Survey During the Experiment).

Item Wording(s)	Answer Options
Thesis Project (No Reference) 1. Are you currently working on a bachelor or master thesis?	Bachelor Thesis / Master Thesis
2. Are you currently working on a different type of thesis?	Seminar Thesis (Bachelor Level) / Seminar Thesis (Master Level) / Dissertation / Neither (No Thesis) / Other (Text Field)
Thesis Writing Demands (Engeser and Rheinberg, 2008)	

Item Wording(s)	Answer Options
For me the current demands in writing my thesis are ...	Much Too Low - Just Right - Much Too Bigh (7p)
Flow Proneness - Domain Specific (Moneta, 2017) When I am writing something for my thesis ... 1. ... I get so involved that my concentration becomes like my breathing ... I never think of it. 2. ... I become so absorbed that I am less aware of myself and my problems. 3. ... I am so involved in it that I don't see myself as separate from what I am doing.	Never - Rarely - Sometimes - Often - Everyday, or almost everyday (5p)

TABLE A.3: Specific Demographic Survey Items Used in Experiment 1 Only (Pre-Experiment Invitation Survey).

Item Wording(s)	Answer Options
Difficulty - Skill Balance (Engeser and Rheinberg, 2008) How did you feel about the last task round overall? For me personally, the task demands were ...	Very low - Very high (7p)
Optimal Difficulty (Ulrich et al., 2014) During the last task round overall ... ... the task demands were well matched to my ability.	Not at all - Very much (7p)
Affect (Valence & Arousal) (Bradley and Lang, 1994) 1. How unpleasant/pleasant were you feeling overall during the last task round? 2. How calm/aroused were you feeling overall during the last task round?	SAM Images (9p)
Task Performance (Hart and Staveland, 1988) During the last task round overall, how... ... successful do you think you were in accomplishing the goals of the task?	Exp. 1: Very little - Very much (7p); Exp. 2: Not at all - Very much (7p)
Flow (FKS) (Engeser and Rheinberg, 2008) During the last task round overall ... 1. ... I felt just the right amount of challenge. 2. ... my thoughts ran fluidly and smoothly. 3. ... I didn't notice time passing. 4. ... I had no difficulty concentrating. 5. ... my mind was completely clear. 6. ... I was totally absorbed in what I was doing. 7. ... the right thoughts occured of their own accord. 8. ... I knew what I had to do each step of the way. 9. ... I felt that I had everything under control. 10. ... I was completely lost in thought.	Not at all - Very much (7p)

Item Wording(s)	Answer Options
Stress (Tams et al., 2014) During the last task round overall ... 1. ... I felt strain due to the task demands. 2. ... I felt emotionally drained. 3. ... I felt used up due to the task demands. 4. ... I felt fatigued due to the task demands. 5. ... I felt burned out from working on the task.	Exp. 1: Strongly disagree - Strongly agree (7p); Exp. 2: Not at all - Very much (7p)

TABLE A.4: Round Survey Items Used in Experiment 1 (After Each Mental Arithmetic Task Condition and Writing Task Interruption) and Experiment 2 (After Each Difficulty Condition Both for the SP and MP Conditions).

Item Wording(s)	Answer Options	Condition
Autonomy (Sheldon and Hilpert, 2012) During the last task round overall ... 1. ... I was free to do things my own way. 2. ... I could express my "true self". 3. ... I was really doing what I wanted to do.	Strongly disagree - Strongly agree (7p)	SP, MP
Effortless Concentration (Harmat et al., 2015) During the last task round overall, how ... 1. ... well were you concentrating during the task? 2. ... hard was it to concentrate during the task?	Not at all - Very much (7p)	SP, MP
Difficulty - Skill Balance (Group) (Engeser and Rheinberg, 2008) How did you feel about the last task round overall? For us as a group, the task demands were ...	Very low - Very high (7p)	MP
Task Performance (Group) (Hart and Staveland, 1988) During the last task round overall, how... ... successful do you think your group was in accomplishing the goals of the task?	Not at all - Very much (7p)	MP
Information Sharing (Aubé, Brunelle, and Rousseau, 2014) During the last math task round overall, we ... 1. ... shared useful information with each of the team members. 2. ... made sure we correctly understood our co-workers' contributions.	Not at all - Very much (7p)	MP

TABLE A.5: Round Survey Items Used in Experiment 2 Only.

Item Wording(s)	Answer Options
Task Importance (Engeser and Rheinberg, 2008) In general, during this whole task (all the task rounds) ... 1. ... something important to me was at stake. 2. ... I was careful to not make mistakes. 3. ... I was worried about failing.	Strongly disagree - Strongly agree (7p)

TABLE A.6: Task Survey Items Used in Experiment 1 (After Completion of the Mental Arithmetic and the Writing Task) and Experiment 2 (After Completion of Each Difficulty Condition Both for the SP and MP Conditions).

Item Wording(s)	Answer Options
Group Interaction Quality (Wageman, Hackman, and Lehman, 2005) After this whole task (all the task rounds), I feel that ... 1. ... there was a lot of unpleasantness among members of this group. 2. ... the longer we worked together as a group, the less well we did. 3. ... working together energised and uplifted members of our group. 4. ... every time someone attempted to correct someone elses' solution, things seemed to get worse rather than better.	Not at all - Very much (7p)
Group Relationship Quality (Wageman, Hackman, and Lehman, 2005) After this whole task (all the task rounds), I feel that ... 1. ... my relations with other group members were strained. 2. ... I very much enjoyed working with my group. 3. ... the chance to work together was one of the best parts of working with this group.	Not at all - Very much (7p)
Group Effort (Wageman, Hackman, and Lehman, 2005) After this whole task (all the task rounds), I feel that ... 1. ... group members demonstrated their commitment by putting in a lot of effort to help us succeed. 2. ... everyone in this group was motivated to have the group succeed. 3. ... some members of our group did not carry their fair share of the overall workload.	Not at all - Very much (7p)
Group Diversity (Wageman, Hackman, and Lehman, 2005) In general, during this whole task (all the task rounds) ... 1. ... members of this group were too dissimilar to work together well. 2. ... this group did not have a broad enough range of experiences and skills to accomplish its objectives. 3. ... this group had a nearly ideal "mix" of members — a set of people who bring different experiences and skills to the task.	Strongly disagree - Strongly agree (7p)
Group Skills (Wageman, Hackman, and Lehman, 2005) In general, during this whole task (all the task rounds) ...	

Item Wording(s)	Answer Options
1. ... members of this group had more than enough talent and experience for the kind of task that we did. 2. ... everyone in this group had the skills that are needed for the group's work. 3. ... some members of this group lacked the knowledge and skills that they needed to do their parts of the group's work.	
Group Interdependence (Wageman, Hackman, and Lehman, 2005) In general, during this whole task (all the task rounds) ... 1. ... members of this group had their own individual jobs to do, with little need for them to work together. 2. ... achieving the objectives of this group required a great deal of communication and coordination among members. 3. ... members of this group had to depend heavily on one another to get the group's work done. (**Only this third item was used in analyses as poor internal consistency was indicated for the whole construct.*)	Strongly disagree - Strongly agree (7p)
Group Size (Wageman, Hackman, and Lehman, 2005) In general, during this whole task (all the task rounds) ... 1. ... this group was larger than it needed to be. 2. ... this group had too few members for what it had to accomplish. 3. ... this group was just the right size to accomplish its objectives. (**Only this third item was used in analyses as poor internal consistency was indicated for the whole construct.*)	Strongly disagree - Strongly agree (7p)
Communication Means (No Reference) In general, during this whole task (all the task rounds) ... ... it was sufficiently possible to coordinate our work.	Strongly disagree - Strongly agree (7p)
Collective Efficacy (Zumeta et al., 2016) After this experiment (all the task rounds) I feel that our group ... 1. ... can show better skills than other groups in tasks like these. 2. ... is effectively prepared for such tasks in the future. 3. ... has improved abilities to overcome problems in such tasks. 4. ... can perform such tasks better than other groups.	Strongly disagree - Strongly agree (7p)
Identity Fusion (Swann et al., 2009) Please chose the diagram that best describes the relationship between you and the other group members (now after all the task rounds).	Pictograms of two more or less overlapping circles (4p).
Social Presence (Gefen and Straub, 2003) In general, during this whole experiment (all the task rounds) ... 1. ... there was a sense of human contact during the group work. 2. ... there was a sense of personalness during the group work. 3. ... there was a sense of sociability during the group work. 4. ... there was a sense of human warmth during the group work.	Strongly disagree - Strongly agree (7p)

Item Wording(s)	Answer Options
5. ... there was a sense of human sensitivity during the group work.	
Group Engagement (Wageman, Hackman, and Lehman, 2005) In general, during this whole experiment (all the task rounds) ... 1. ... I felt a real sense of personal satisfaction when our group did well. 2. ... I felt bad and unhappy when our group had performed poorly. 3. ... my own feelings were not affected one way or the other by how well our group performed. 4. ... when our group had done well, I felt that have done well.	Strongly disagree - Strongly agree (7p)
General Group Satisfaction (Wageman, Hackman, and Lehman, 2005) In general, during this whole experiment (all the task rounds) ... 1. ... I enjoyed the kind of work we did in our group. 2. ... working with this group was an exercise in frustration. 3. ... generally speaking, I am very satisfied with this group.	Strongly disagree - Strongly agree (7p)
Personal Growth Within the Group (Wageman, Hackman, and Lehman, 2005) In general, during this whole experiment (all the task rounds) ... 1. ... I learned a great deal from my work with this group. 2. ... my own creativity and initiative were suppressed by this group. 3. ... working with this group stretched my personal knowledge and skills.	Strongly disagree - Strongly agree (7p)

TABLE A.7: Task Survey Items Used in Experiment 2 Only (Only for the MP Condition).

Item Wording(s)	Answer Options
Flow Proneness - Domain General (Ullén et al., 2012) When you do something during your (university) work, how often does it happen that ... 1. ... you feel bored? 2. ... it feels like your ability to perform what you do completely matches how difficult it is? 3. ... you have a clear picture of what you want to achieve, and what you need to do to get there? 4. ... you are conscious of how well or poorly you are performing at what you are doing? 5. ... you feel completely concentrated? 6. ... you have a sense of complete control? 7. ... what you do feels extremely enjoyable? When you are doing household work or other routine chores (e.g. cooking, cleaning, ironing) how often does it happen that ... 1. ... you feel bored? 2. ... it feels like your ability to perform what you do completely matches how difficult it is? 3. ... you have a clear picture of what you want to achieve, and what you need to do to get there?	Never - Rarely - Sometimes - Often - Everyday, or almost everyday (5p)

Item Wording(s)	Answer Options
4. ... you are conscious of how well or poorly you are performing at what you are doing?	
5. ... you feel completely concentrated?	
6. ... you have a sense of complete control?	
7. ... what you do feels extremely enjoyable?	
When you do something in your leisure time, how often does it happen that ...	
1. ... you feel bored?	
2. ... it feels like your ability to perform what you do completely matches how difficult it is?	
3. ... you have a clear picture of what you want to achieve, and what you need to do to get there?	
4. ... you are conscious of how well or poorly you are performing at what you are doing?	
5. ... you feel completely concentrated?	
6. ... you have a sense of complete control?	
7. ... what you do feels extremely enjoyable?	
Math Task Preference (Ulrich et al., 2014)	
How much do you ...	
1. ... like performing mental arithmetic?	Very little - Very much (7p)
2. ... like to write something (a report, short story, etc.)?	
3. ... prefer mental arithmetic over writing?	
Open Remarks (No Reference)	
Is there anything you noticed during the experiment that you would like to comment on?	Open Text Field

TABLE A.8: End Survey Items Used in Experiment 1 and Experiment 2 After Completion of All Tasks.

A.4 Neurophysiological Signal Processing Pipelines

Adhering to recent recommendations for reproducible neurophysiological research (Bigdely-Shamlo et al., 2015), the complete signal processing pipelines for ECG and EEG data are outlined in the following tables.

Steps & Parameters	Reference
1. Data Extraction	
Eyes Open Baseline & Task Phases (5m Windows)	-
2. RR-Interval Extraction	
Signal Filtering (FIR 3-45 Bandpass)	-
R-Peak Segmentation (Hamilton Segmenter)	Hamilton (2002)
Correct R-Peak Locations to the Maximum Within a Tolerance (Defined as the Time Interval)	-
3. HRV Feature Extraction - Initial Processing for All Features	
Artefact Detection (Statistical & Physiological)	Makowski (2016)
Time Domain Feature Extraction: RMSSD, SDNN, Percent of Adjacent NN Intervals not Differing More than 50ms (PNN50)	Malik et al. (1996)
4. HRV Feature Extraction - Add. Pre-processing for Freq. Domain Features	
RR Interval Interpolation (Third Order Splines)	-
Butterworth Bandpass Filter (Range of Target Freq. Band)	-
Hilbert Transformation	-
Freq. Power Extraction (Multitaper Method): HF-HRV (.15 to .40 Hz); LF-HRV (.04 to .15 Hz); Total Power	Malik et al. (1996)
5. Change Score Computation	
ΔHRV = Task - Eyes Open Baseline	Harmat et al. (2015) and Tozman et al. (2015)

TABLE A.9: ECG Processing Pipeline for Experiment 1.

Steps & Parameters	Reference
1. Data Extraction	
Eyes Open Baseline & Task Phases (5m Windows)	-
2. RR-Interval Extraction	
Signal Filtering (FIR 3-45 Bandpass)	-
R-Peak Segmentation (Hamilton Segmenter)	Hamilton (2002)
Correct R-Peak Locations to the Maximum Within a Tolerance (Defined as the Time Interval)	-
3. HRV Feature Extraction - Initial Processing for All Features	
RR Outlier Removal ($\geq$ 2 SD from Mean)	-

Steps & Parameters	Reference
Time Domain Feature Extraction: Mean Average Heart Rate (mHR), RMSSD	Malik et al. (1996)
4. HRV Feature Extraction - Additional Pre-processing for Freq. Domain Features	
RR Interval Interpolation (Cubic)	Morelli et al. (2019)
Freq. Power Extraction: (Welch Windows With Size 120; Linear Detrending): HF-HRV (.15 to .40 Hz)	Malik et al. (1996)
Natural Logarithm Transformation of Freq. Powers	Berntson, Quigley, and Lozano (2007)
5. Erroneous Data Inspection	
Delete Participants Where mHR is $\geq 1.5 *$ Interquartile Range (IQR) From Mean	-
6. Change Score Computation	
ΔHRV = Task - Eyes Open Baseline	Harmat et al. (2015) and Tozman et al. (2015)

TABLE A.10: ECG Processing Pipeline for Experiment 2.

Steps & Parameters	Reference
1. Data Extraction	
Eyes Closed, Eyes Open Baseline & Task Phases	-
Channel Centering: Subtraction of Channel Mean	-
2. Data Cleaning	
Signal Drop Artefacts: Deleting Epochs With Amplitudes >5 SD From Mean	-
Detrending: 0,5 - 45 Hz Bandpass Filter	-
Line Noise Removal: 50 Hz Notch Filter	-
Channel & Paroxysmal Artefact Removal: Visual Rejection of Noisy Channels & Epochs	-
Stationary Artefact Removal (Independent Components): AMICA - ICs: [Blinks, General Discontinuities]	-
3. Feature Extraction	
Epoching: 2s Windows; 50% Overlap; Hann Taper	Ewing et al. (2016)
Freq. Power Extraction (Short-Time FFT)	Cohen (2014)
Power Normalisation: Natural Logarithm Transformation	Ewing et al. (2016)
Freq. Band Extraction: Theta, LoAlpha, HiAlpha, LoBeta, MidBeta, & HiBeta Bands	Berta et al. (2013)
4. Additional Feature Aggregation	
Regional Electrode Pooling (Mean of 8 Frontal Electrodes = AF3, F3, F7, FC5, FC6, F8, F4, AF4) & Lateral Difference (RH-LH for Asymmetry Scores)	Smith et al. (2017)
Temporal Feature Pooling: Median of Experiment Phase	-

Steps & Parameters	Reference
Change Score Computation: ΔPower = Power$_{Task}$ - Power$_{EyesOpenBaseline}$	-

TABLE A.11: EEG Processing Pipeline for Experiment 1.

Steps & Parameters	Reference
1. Data Extraction	
Eyes Closed, Eyes Open Baseline & Task Phases	-
Channel Centering: Subtraction of Channel Mean	-
2. Data Cleaning	
Line Noise Removal: 50 Hz & 100 Hz Re-Referencing: Robust Common Average Reference	Bigdely-Shamlo et al. (2015)
Detrending: 1 Hz High-Pass	-
Trim Outliers: 800mV / 250ms	-
Channel & Paroxysmal Artefact Removal: Artefact Subspace Reconstruction - Burst Criterion 10 SD	Mullen et al. (2015)
Stationary Artefact Removal (Independent Components): AMICA - ICs: [Horizontal & Vertical Eye Movement, Blinks, General Discontinuities] via ADJUST	Mognon et al. (2011)
Processing Inspection: Visual Input-Output Comparison	-
3. Feature Extraction	
Freq. Power Extraction: Morlet Wavelets (55 Frequencies, Range [3, 60], Cycle Range [3,10] Log. Spaced With Freq.) Power Normalisation: dB Power = 10 * $\log_{10}(\mu V^2/Hz)$	Cohen (2014)
Frequency Band Extraction: Theta, Lo2Alpha, HiAlpha, LoBeta, MidBeta, & HiBeta Bands From IAF Peak	Klimesch (1999)
4. Additional Feature Aggregation	
Temporal Feature Pooling: Median of Experiment Phase	-
Completeness Check: Remove Subjects With Short Baseline Data (<80% of Expected Samples); Remove Exp. Phases With <80% of Expected Samples	-
Change Score Computation: ΔPower = Power$_{Task}$ - Power$_{EyesOpenBaseline}$	-
Spatial Pooling (ROIs): Remove Participant When <50% of Electrodes Available	-

TABLE A.12: EEG Processing Pipeline for Experiment 2.

Bibliography

Acharya, U. Rajendra et al. (2018). “Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals”. In: *Computers in Biology and Medicine* 100.August 2017, pp. 270–278.

Admiraal, Wilfried et al. (2011). “The concept of flow in collaborative game-based learning”. In: *Computers in Human Behavior* 27.3, pp. 1185–1194.

Afergan, Daniel et al. (2014). “Dynamic difficulty using brain metrics of workload”. In: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*, pp. 3797–3806.

Agarwal, Ritu and Elena Karahanna (2000). “Time flies when you’re having fun: Cognitive absorption and beliefs about information technology usage”. In: *Management of Information Systems Quarterly* 24.4, pp. 665–694.

Allen, John J.B. and Michael X. Cohen (2010). “Deconstructing the "resting" state: Exploring the temporal dynamics of frontal alpha asymmetry as an endophenotype for depression”. In: *Frontiers in Human Neuroscience* 4.December, pp. 1–14.

Andreassi, John L. (2000). *Psychophysiology: Human Behavior & Physiological Response*. Mahawah: Lawrence Erlbaum Associates, pp. 1–690.

Aparicio, Luana V M et al. (2016). “A systematic review on the acceptability and tolerability of transcranial direct current stimulation treatment in neuropsychiatry trials”. In: *Brain Stimulation* 9.5, pp. 671–681.

Aria, Massimo and Corrado Cuccurullo (2017). “bibliometrix: An R-tool for comprehensive science mapping analysis”. In: *Journal of Informetrics* 11.4, pp. 959–975.

Armstrong, Alayne C. (2008). “The fragility of group flow: The experiences of two small groups in a middle school mathematics classroom”. In: *Journal of Mathematical Behavior* 27.2, pp. 101–115.

Armstrong, Blair C. et al. (2015). “Brainprint: Assessing the uniqueness, collectability, and permanence of a novel method for ERP biometrics”. In: *Neurocomputing* 166, pp. 59–67.

Asakawa, Kiyoshi (2004). “Flow Experience and Autotelic Personality in Japanese College Students: How Do They Experience Challenges in Daily Life?” In: *Journal of Happiness Studies* 5.2, pp. 123–154.

Aubé, Caroline, Eric Brunelle, and Vincent Rousseau (2014). “Flow experience and team performance: The role of team goal commitment and information exchange”. In: *Motivation and Emotion* 38.1, pp. 120–130.

Azevedo, Ruben T et al. (2017). “The calming effect of a new wearable device during the anticipation of public speech”. In: *Scientific Reports* 7.1, pp. 1–7.

Babayan, Anahit et al. (2019). “Data descriptor: A mind-brain-body dataset of MRI, EEG, cognition, emotion, and peripheral physiology in young and old adults”. In: *Scientific Data* 6, pp. 1–21.

Baig, Mirza Mansoor, Hamid Gholamhosseini, and Martin J. Connolly (2013). “A comprehensive survey of wearable and wireless ECG monitoring systems for older adults”. In: *Medical and Biological Engineering and Computing* 51.5, pp. 485–495.

Bakdash, Jonathan Z. and Laura R. Marusich (2017). “Repeated measures correlation”. In: *Frontiers in Psychology* 8.MAR, pp. 1–13.
Bakker, Arnold B. (2008). “The work-related flow inventory: Construction and initial validation of the WOLF”. In: *Journal of Vocational Behavior* 72.3, pp. 400–414.
Bakker, Arnold B. and Marianne van Woerkom (2017). “Flow at Work: a Self-Determination Perspective”. In: *Occupational Health Science* 1.1-2, pp. 47–65.
Bandara, Wasana et al. (2015). “Achieving rigor in literature reviews: Insights from qualitative data analysis and tool-support”. In: *Communications of the Association for Information Systems* 37, pp. 154–204.
Barham, Michael P. et al. (2017). “Acquiring research-grade ERPs on a shoestring budget: A comparison of a modified Emotiv and commercial SynAmps EEG system”. In: *Psychophysiology* 54.9, pp. 1393–1404.
Barros, Marcelo Felipe de Sampaio et al. (2018). “Flow experience and the mobilization of attentional resources”. In: *Cognitive, Affective, & Behavioral Neuroscience* 18.4, pp. 810–823.
Berg-Beckhoff, Gabriele, Grace Nielsen, and Eva Ladekjær Larsen (2017). “Use of Information Communication Technology and Stress, Burnout, and Mental Health in Older, Middle-aged, and Younger workers: Results From a Systematic Review”. In: *International Journal of Occupational and Environmental Health* 23.2, pp. 160–171.
Berka, Chris and Maja Stikic (2017). “On the road to autonomy: Evaluating and optimizing hybrid team dynamics”. In: *Autonomy and Artificial Intelligence: A Threat or Savior?* Springer, pp. 245–262.
Berntson, Gary G, Karen S Quigley, and Dave Lozano (2007). “Cardiovascular Psychophysiology”. In: *Handbook of Psychophysiology*. Ed. by John T. Cacioppo, Louis G. Tassinary, and Gary G. Berntson. 3rd ed. Cambridge: Cambridge University Press, pp. 182–210.
Berta, Riccardo et al. (2013). “Electroencephalogram and Physiological Signal Analysis for Assessing Flow in Games”. In: *IEEE Transactions on Computational Intelligence and AI in Games* 5.2, pp. 164–175.
Beyer, Justus et al. (2015). “Using electroencephalography and subjective self-assessment to measure the influence of quality variations in cloud gaming”. In: *2015 7th International Workshop on Quality of Multimedia Experience, QoMEX 2015*, pp. 1–6.
Bian, Yulong et al. (2016). “A framework for physiological indicators of flow in VR games: construction and preliminary evaluation”. In: *Personal and Ubiquitous Computing* 20.5, pp. 821–832.
Bigdely-Shamlo, Nima et al. (2015). “The PREP pipeline: standardized preprocessing for large-scale EEG analysis”. In: *Frontiers in Neuroinformatics* 9.June, pp. 1–20.
Bikson, Marom et al. (2016). “Safety of transcranial direct current stimulation: evidence based update 2016”. In: *Brain Stimulation* 9.5, pp. 641–661.
Bland, Martin J and Douglas G Altman (1995). “Statistics notes: Calculating correlation coefficients with repeated observations: Part 1—correlation within subjects”. In: *BMJ* 310.6977, p. 446.
Blankertz, Benjamin et al. (2016). “The Berlin Brain-Computer Interface: Progress Beyond Communication and Control”. In: *Frontiers in Neuroscience* 10, pp. 1–24.
Bleichner, Martin G. and Stefan Debener (2017). “Concealed, unobtrusive ear-centered EEG acquisition: Ceegrids for transparent EEG”. In: *Frontiers in Human Neuroscience* 11.April, pp. 1–14.
Bleichner, Martin G., Preben Kidmose, and Jérémie Voix (2020). “Editorial: Ear-Centered Sensing: From Sensing Principles to Research and Clinical Devices”. In: *Frontiers in Neuroscience* 13.January, pp. 2019–2020.

Bless, Herbert and Axel M Burger (2016). "A Closer Look at Social Psychologists ' Silver Bullet : Inevitable and Evitable Side Effects of the Experimental Approach". In: *Perspectives on Psychological Science* 11.2, pp. 296–308.

Bliese, Paul D. (2000). "Within-Group Agreement, Non-Independence, and Reliability: Implications for Data Aggregation and Analysis". In: *Multilevel Theory, Research, and Methods in Organizations*. Ed. by Katherine J Klein and Steve W J Kozlowski. San Francisco: Jossey-Bass, pp. 349–381.

Bobrov, Pavel et al. (2011). "Brain-computer interface based on generation of visual images". In: *PLoS ONE* 6.6, pp. 1–12.

Bombeke, Klaas et al. (2018). "Do Not Disturb: Psychophysiological Correlates of Boredom, Flow and Frustration During VR Gaming". In: *International Conference on Augmented Cognition*, pp. 101–119.

Borghini, Gianluca et al. (2014). "Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness". In: *Neuroscience and Biobehavioral Reviews* 44, pp. 58–75.

Börner, Katy, Chaomei Chen, and Kevin W. Boyack (2005). "Visualizing knowledge domains". In: *Annual Review of Information Science and Technology* 37.1, pp. 179–255.

Boucsein, Wolfram (2012). *Electrodermal Activity*. 2nd. New York: Springer Science + Business Media, pp. 1–635.

Bovend'Eerdt, Thamar J. H., Rachel E. Botell, and Derick T. Wade (2009). "Writing SMART rehabilitation goals and achieving goal attainment scaling: a practical guide". In: *Clinical Rehabilitation* 23, pp. 352–361.

Bowman, Julia et al. (2015). "The development, content validity and inter-rater reliability of the SMART-Goal Evaluation Method: A standardised method for evaluating clinical goals". In: *Australian Occupational Therapy Journal* 62.6, pp. 420–427.

Bradley, Margaret M and Peter J Lang (1994). "Measuring Emotion: The Self-Assessment Manikin and the Semantic Differential". In: *Journal of Behavior Therapy and Experimental Psychiatry* 25.1, pp. 49–59.

Bridwell, David A. et al. (2018). "Moving Beyond ERP Components : A Selective Review of Approaches to Integrate EEG and Behavior". In: *Frontiers in Human Neuroscience* 12.March, pp. 1–17.

Brom, Cyril et al. (2014). "Flow, social interaction anxiety and salivary cortisol responses in serious games: A quasi-experimental study". In: *Computers and Education* 79, pp. 69–100.

Brouwer, Anne Marie et al. (2015). "Using neurophysiological signals that reflect cognitive or affective state: Six recommendations to avoid common pitfalls". In: *Frontiers in Neuroscience* 9.APR, pp. 1–11.

Brunner, Sibylle and Julia Schueler (2009). "The rewarding effect of flow experience on performance in a marathon race". In: *Psychology of Sport and Exercise* 10, pp. 168–174.

Bruya, Brian (2010). "Toward a Theory of Attention That Includes Effortless Attention and Action". In: *Effortless Attention: A New Perspective in the Cognitive Science of Attention and Action*. Ed. by Brian Bruya. Cambridge: The MIT Press, pp. 1–28.

Buckner, Randy L., Jessica R. Andrews-Hanna, and Daniel L. Schacter (2008). "The Brain's Default Network: Anatomy, Function, and Relevance to Disease". In: *Annals of the New York Academy of Sciences* 1124.1, pp. 1–38.

Buzsaki, Gyorgy (2006). *Rhythms of the Brain*. Oxford: Oxford University Press, pp. 1–464.

Buzsáki, György and Andreas Draguhn (2004). "Neuronal Oscillations in Cortical Networks". In: *Science* 304.5679, pp. 1926–1929.

Cacioppo, John T, Louis G Tassinary, and Gary G Berntson (2007). "Psychophysiological Science: Interdisciplinary Approaches to Classic Questions About the Mind". In: *The Handbook of Psychophysiology*. Ed. by John T Cacioppo, Louis G Tassinary, and Gary G Berntson. 3rd. Cambridge: Cambridge University Press. Chap. 1, pp. 1–16.

Carlén, Marie (2017). "What constitutes the prefrontal cortex?" In: *Science* 358.6362, pp. 478–482.

Carreiras, Carlos et al. (2015). *BioSPPy - Biosignal Processing in Python*. URL: `https://github.com/PIA-Group/BioSPPy`.

Ceja, Lucia and Jose Navarro (2012). "'Suddenly I get into the zone': Examining discontinuities and nonlinear changes in flow experiences at work". In: *Human Relations* 65.9, pp. 1101–1127.

Chanel, Guillaume and Christian Mühl (2015). "Connecting brains and bodies: Applying physiological computing to support social interaction". In: *Interacting with Computers* 27.5, pp. 534–550.

Chanel, Guillaume et al. (2011). "Emotion Assessment From Physiological Signals for Adaptation of Game Difficulty". In: *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans* 41.6, pp. 1052–1063.

Chatterjee, Debatri, Aniruddha Sinha, and Meghamala Sinha (2016). "A Probabilistic Approach for Detection and Analysis of Cognitive Flow". In: pp. 44–53.

Cheron, Guy (2016). "How to measure the psychological "flow"? A neuroscience perspective". In: *Frontiers in Psychology* 7.DEC, pp. 1–6.

Chin, Michael S. and Stefanos N. Kales (2019). "Is there an optimal autonomic state for enhanced flow and executive task performance?" In: *Frontiers in Psychology* 10.JULY, pp. 1–8.

Cipresso, Pietro et al. (2015). "Psychometric modeling of the pervasive use of Facebook through psychophysiological measures: Stress or optimal experience?" In: *Computers in Human Behavior* 49.June 2013, pp. 576–587.

Clarke, Nicholas (2010). "Emotional intelligence abilities and their relationships with team processes". In: *Team Performance Management* 16.1/2, pp. 6–32.

Cleveland, William S (1979). "Robust locally weighted regression and smoothing scatterplots". In: *Journal of the American Statistical Association* 74.368, pp. 829–836.

Cnaan, Avital, Nan M. Laird, and Peter Slasor (2005). "Mixed Models: Using the General Linear Mixed Model to Analyse Unbalanced Repeated Measures and Longitudinal Data". In: *Tutorials in Biostatistics* 2, pp. 127–158.

Cohen, Michael X. (2017). "Where Does EEG Come From and What Does It Mean?" In: *Trends in Neurosciences* 40.4, pp. 208–218.

Cohen, Mike X (2014). *Analyzing neural time series data: theory and practice*. Cambridge: MIT University Press, pp. 1–578.

Colzato, Lorenza S., Gina Wolters, and Corinna Peifer (2018). "Transcutaneous vagus nerve stimulation (tVNS) modulates flow experience". In: *Experimental Brain Research* 236.1, pp. 253–257.

Courtemanche, François et al. (2020). "Ambient Facial Emotion Recognition: A Pilot Study". In: *Proceedings of the 12th Retreat on NeuroIS*, pp. 313–319.

Csikszentmihalyi, Mihaly (1975). *Beyond Boredom and Anxiety*. San Francisco: Jossey-Bass, pp. 1–272.

— (1990). *Flow: The psychology of optimal experience*. New York: Harper Perennial Modern Classics, pp. 1–336.

— (1996). *Flow and the psychology of discovery and invention*. New York: Harper Collins, pp. 1–480.

— (1997). *Finding flow: The psychology of engagement with everyday life*. New York: Basic Books, pp. 1–166.

Csikszentmihalyi, Mihaly and Jeremy Hunter (2003). “Happiness in everyday life: The uses of experience sampling”. In: *Journal of Happiness Studies* 4, pp. 185–199.

Csikszentmihalyi, Mihaly and Judith LeFevre (1989). “Optimal experience in work and leisure”. In: *Journal of Personality and Social Psychology* 56.5, pp. 815–822.

Culbertson, Satoris S. et al. (2015). *Contagious Flow: Antecedents and Consequences of Optimal Experience in the Classroom*. Vol. 39. 3, pp. 319–349.

Dann, David et al. (2019). “DISKNET–A Platform for the Systematic Accumulation of Knowledge in IS Research”. In: *Proceedings of the 40th International Conference on Information Systems (ICIS)*, pp. 1–9.

David Hairston, W et al. (2014). “Usability of four commercially-oriented EEG systems.” In: *Journal of neural engineering* 11.4, pp. 1–14.

Davis, Fred D (1985). “A technology acceptance model for empirically testing new end-user information systems: Theory and results”. Doctoral Dissertation. Massachusetts Institute of Technology.

Day, Trevor and Paul Tosey (2011). “Beyond SMART? A new framework for goal setting”. In: *Curriculum Journal* 22.4, pp. 515–534.

De Kock, Frederick Gideon (2014). “The neuropsychological measure (EEG) of flow under conditions of peak performance”. Doctoral Dissertation. University of South Africa.

de Moura Jr, Pedro Jacome and Carlo Gabriel Porto Bellini (2019). “The measurement of flow and social flow at work: a 30-year systematic review of the literature”. In: *Personnel Review* 49.2, pp. 537–570.

Debener, Stefan et al. (2015). “Unobtrusive ambulatory EEG using a smartphone and flexible printed electrodes around the ear”. In: *Scientific Reports* 5, pp. 1–11.

Debus, Maike E. et al. (2014). “Making flow happen: The effects of being recovered on work-related flow between and within days”. In: *Journal of Applied Psychology* 99.4, pp. 713–722.

Deci, E L and R M Ryan (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Springer Science + Business Media, pp. 1–375.

Deci, Edward L. and Richard M. Ryan (2000). “The “What” and “Why” of Goal Pursuits: Human Needs and the Self-Determination of Behavior”. In: *Psychological Inquiry* 11.4, pp. 227–268.

Deiber, Marie-Pierre et al. (2007). “Distinction between perceptual and attentional processing in working memory tasks: a study of phase-locked and induced oscillatory brain dynamics”. In: *Journal of Cognitive Neuroscience* 19.1, pp. 158–172.

Delle Fave, Antonella and Fausto Massimini (2005). “The investigation of optimal experience and apathy: Developmental and psychosocial implications”. In: *European Psychologist* 10.4, pp. 264–274.

Delle Fave, Antonella, Fausto Massimini, and Marta Bassi (2011). “Instruments and Methods in Flow Research”. In: *Psychological Selection and Optimal Experience Across Cultures*. Ed. by Antonella Delle Fave, Fausto Massimini, and Marta Bassi. Springer, pp. 59–87.

Delorme, Arnaud and Scott Makeig (2004). “EEGLAB : an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis”. In: 134, pp. 9–21.

Demetrashvili, Nino and Edwin R. Van den Heuvel (2015). “Confidence intervals for intraclass correlation coefficients in a nonlinear dose-response meta-analysis”. In: *Biometrics* 71.2, pp. 548–555.

Derks, Daantje, Agneta H. Fischer, and Arjan E R Bos (2008). “The role of emotion in computer-mediated communication: A review”. In: *Computers in Human Behavior* 24.3, pp. 766–785.

Di Domenico, Stefano I. and Richard M. Ryan (2017). “The Emerging Neuroscience of Intrinsic Motivation: A New Frontier in Self-Determination Research”. In: *Frontiers in Human Neuroscience* 11.March, pp. 1–14.

Dietrich, Arne (2003). “Functional neuroanatomy of altered states of consciousness: The transient hypofrontality hypothesis”. In: *Consciousness and Cognition* 12.2, pp. 231–256.

— (2004). “Neurocognitive mechanisms underlying the experience of flow”. In: *Consciousness and Cognition* 13.4, pp. 746–761.

Dietrich, Arne et al. (2010). “Effortless Attention, Hypofrontality , and Perfectionism”. In: *Effortless Attention: A New Perspective in the Cognitive Science of Attention and Action*. Ed. by Brian Bruya. Cambridge: The MIT Press, pp. 159–178.

Dixon, Mike J. et al. (2019). “Reward reactivity and dark flow in slot-machine gambling: “Light” and “dark” routes to enjoyment”. In: *Journal of Behavioral Addictions* 8.3, pp. 489–498.

Doran, George T (1981). “There’s a S.M.A.R.T. way to write managements’s goals and objectives.” In: *Management Review* 70.11, pp. 35–36.

Drachen, Anders and Anja Lee Pedersen (2009). “Correlation between Heart Rate, Electrodermal Activity and Player Experience in First-Person Shooter Games”. In: *Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games*, pp. 49–54.

Drachen, Anders et al. (2010). “Psychophysiological Correlations with Gameplay Experience Dimensions”. In: *Brain, Body and Bytes, Workshop - CHI 2010*, pp. 1–4.

Ellis, Gary D, Judith E Voelkl, and Catherine Morris (1994). “Measurement and analysis issues with explanation of variance in daily experience using the flow model”. In: *Journal of Leisure Research* 26.4, pp. 337–356.

Engel, David et al. (2014). “Reading the mind in the eyes or reading between the lines? Theory of mind predicts collective intelligence equally well online and face-to-face”. In: *PLoS ONE* 9.12, pp. 1–16.

Engel, David et al. (2015). “Collective Intelligence in Computer-Mediated Collaboration Emerges in Different Contexts and Cultures”. In: *Proceedings of the ACM CHI’15 Conference on Human Factors in Computing Systems* 1, pp. 3769–3778.

Engeser, Stefan and Falko Rheinberg (2008). “Flow, performance and moderators of challenge-skill balance”. In: *Motivation and Emotion* 32.3, pp. 158–172.

Engeser, Stefan and Anja Schiepe-Tiska (2012). “Historical lines and an overview of current research on flow”. In: *Advances in Flow Research* 9781461423.1975, pp. 1–22.

Erhard, K. et al. (2014). “Professional training in creative writing is associated with enhanced fronto-striatal activity in a literary text continuation task”. In: *NeuroImage* 100, pp. 15–23.

Ewing, Kate C., Stephen H. Fairclough, and Kiel Gilleade (2016). “Evaluation of an Adaptive Game that Uses EEG Measures Validated during the Design Process as Inputs to a Biocybernetic Loop”. In: *Frontiers in Human Neuroscience* 10, pp. 1–13.

Fairclough, Stephen H et al. (2013). “Capturing user engagement via psychophysiology: measures and mechanisms for biocybernetic adaptation”. In: *International Journal of Autonomous and Adaptive Communications Systems* 6.1, pp. 63–79.

Ferrez, Pierre W and José del Milan (2007). “Error-Related EEG Potentials in Brain-Computer Interfaces”. In: *Toward Brain-Computer Interfacing*. Ed. by Guido Dornhege et al. Cambridge: MIT Press, pp. 291–301.

Flower, Linda and John R Hayes (1981). “A Cognitive Process Theory of Writing”. In: *College Composition and Communication* 32.4, pp. 365–387.

Fong, Carlton J, Diana J Zaleski, and Jennifer Kay Leach (2015). "The challenge–skill balance and antecedents of flow: A meta-analytic investigation". In: *The Journal of Positive Psychology* 10.5, pp. 425–446.

Fortin-Cote, Alexis et al. (2019). "Affective computing out-of-the-lab: The cost of low cost". In: *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, pp. 4137–4142.

Frey, Carl Benedikt and Michael A. Osborne (2017). "The future of employment: How susceptible are jobs to computerisation?" In: *Technological Forecasting and Social Change* 114, pp. 254–280.

Fullagar, Clive and Antonella Delle Fave (2017). *Flow at work: Measurement and implications*. London: Taylor & Francis, pp. 1–206.

Gaggioli, Andrea et al. (2013). "Psychophysiological correlates of flow during daily activities". In: *Annual Review of Cybertherapy and Telemedicine 2013*. Ed. by B. K. Wiederhold and Giuseppe Riva. Vol. 191. JUNE. IOS Press, pp. 65–69.

Gaggioli, Andrea et al. (2017). "Networked Flow in musical bands". In: *Psychology of Music* 45.2, pp. 283–297.

Galluch, Pamela S, Varun Grover, and Jason Bennett Thatcher (2015). "Interrupting the Workplace: Examining Stressors in an Information Technology Context." In: *Journal of the Association for Information Systems* 16.1, pp. 1–47.

Gallup (2017). *State of the Global Workplace*. New York: Gallup Press, pp. 1–122.

Gefen, David and Detmar Straub (2003). "Managing user trust in B2C e-services". In: *e-Service* 2.2, pp. 7–24.

Gilleade, Kiel Mark, Alan Dix, and Jen Allanson (2005). "Affective Videogames and Modes of Affective Gaming: Assist Me, Challenge Me, Emote Me". In: *roceedings of DIGRA'2005*, pp. 1–7.

Gold, Joshua and Joseph Ciorciari (2019). "A Transcranial Stimulation Intervention to Support Flow State Induction". In: *Frontiers in Human Neuroscience* 13.August, pp. 1–8.

Griethuijsen, Ralf A.L.F. van et al. (2015). "Global patterns in students' views of science and interest in science". In: *Research in Science Education* 45.4, pp. 581–603.

Gruzelier, John et al. (2010). "Acting performance and flow state enhanced with sensory-motor rhythm neurofeedback comparing ecologically valid immersive VR and training screen scenarios". In: *Neuroscience Letters* 480.2, pp. 112–116.

Guger, Christoph et al. (2012). "Comparison of dry and gel based electrodes for P300 brain–computer interfaces". In: *Frontiers in Neuroscience* 6, pp. 1–7.

Hamari, Juho, Jonna Koivisto, and Harri Sarsa (2014). "Does gamification work? - A literature review of empirical studies on gamification". In: *Proceedings of the Annual Hawaii International Conference on System Sciences*, pp. 3025–3034.

Hamilton, Patrick S (2002). "Open source ECG analysis software documentation". In: *Computers in Cardiology* 29, pp. 101–104.

Harari, Yuval Noah (2008). "Combat Flow: Military, Political, and Ethical Dimensions of Subjective Well-Being in War". In: *Review of General Psychology* 12.3, pp. 253–264.

Harmat, L. et al. (2011). "Heart Rate Variability During Piano Playing: A Case Study of Three Professional Solo Pianists Playing a Self-Selected and a Difficult Prima Vista Piece". In: *Music and Medicine* 3.2, pp. 102–107.

Harmat, László et al. (2015). "Physiological correlates of the flow experience during computer game playing". In: *International Journal of Psychophysiology* 97.1, pp. 1–7.

Harmon-Jones, Eddie, Philip A. Gable, and Carly K. Peterson (2010). "The role of asymmetric frontal cortical activity in emotion-related phenomena: A review and update". In: *Biological Psychology* 84.3, pp. 451–462.

Harris, David J., Samuel J. Vine, and Mark R. Wilson (2017a). "Is flow really effortless? The complex role of effortful attention." In: *Sport, Exercise, and Performance Psychology* 6.1, pp. 103–114.

— (2017b). "Neurocognitive mechanisms of the flow state". In: *Progress in Brain Research* 237.1, pp. 221–243.

Hart, Sandra G and Lowell E Staveland (1988). "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research". In: *Advances in Psychology* 52.1, pp. 139–183.

Heerden, Ariana Van (2010). "Creativity, the flow state and brain function". In: *South African Journal of Art History* 25.3, pp. 141–151.

Hesterberg, Tim C. (2015). "What Teachers Should Know About the Bootstrap: Resampling in the Undergraduate Statistics Curriculum". In: *American Statistician* 69.4, pp. 371–386.

Heyne, K., D. Pavlas, and E. Salas (2011). "An Investigation on the Effects of Flow State on Team Process and Outcomes". In: *Proceedings of the Human Factors Society Annual Meeting* 55, pp. 475–479.

Hinterberger, Thilo, Tsutomu Kamei, and Harald Walach (2011). "Psychophysiological classification and staging of mental states during meditative practice". In: *Biomedizinische Technik* 56.6, pp. 341–350.

Hofstede, Geert (1984). "Cultural Dimensions In Management and Planning". In: *Asia Pacific Journal of Management* 1.2, pp. 81–99.

Homan, Richard W., John Herman, and Phillip Purdy (1987). "Cerebral location of international 10-20 system electrode placement". In: *Electroencephalography and Clinical Neurophysiology* 66.4, pp. 376–382.

Hommel, Bernhard (2010). "Grounding attention in action control: The intentional control of selection". In: *Effortless attention: A new perspective in the cognitive science of attention and action*. Ed. by Brian Bruya, pp. 121–140.

Hout, Jef J.J. van den, Orin C. Davis, and Mathieu C.D.P. Weggeman (2018). "The Conceptualization of Team Flow". In: *Journal of Psychology: Interdisciplinary and Applied* 152.6, pp. 388–423.

Huskey, Richard et al. (2018). "Does intrinsic reward motivate cognitive control? A naturalistic-fMRI study based on the Synchronization Theory of Flow". In: *Cognitive, Affective, & Behavioral Neuroscience* 18.5, pp. 902–924.

Ijsselsteijn, Wijnand, Karolien Poels, and Yvonne A W De Kort (2008). "The Game Experience Questionnaire: Development of a self-report measure to assess player experiences of digital games". Doctoral Dissertation. Eindhoven Technical Univeristy.

Jackson, Susan and Herbert Marsh (1996). "Development and validation of a scale to measure optimal experience: The flow state scale". In: *Journal of Sport & Exercise Psychology* 18, pp. 17–35.

Jackson, Susan A., Andrew J. Martin, and Robert C. Eklund (2008). "Long and short measures of flow: The construct validity of the FSS-2, DFS-2, and new brief counterparts". In: *Journal of Sport and Exercise Psychology* 30.5, pp. 561–587.

Jaque, S Victoria, Isabel H Karamanukyan, and Paula Thomson (2015). "A psychophysiological case study of orchestra conductors". In: *Medical Problems of Performing Artists* 30.4, pp. 189–196.

John, E R et al. (1980). "Developmental equations for the electroencephalogram". In: *Science* 210.4475, pp. 1255–1258.

Johnson, Daniel et al. (2015). "Cooperative Game Play with Avatars and Agents: Differences in Brain Activity and the Experience of Play". In: *Proceedings of the*

33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15, pp. 3721–3730.

Kahana, Michael J. et al. (1999). "Human theta oscillations exhibit task dependence during virtual maze navigation". In: *Nature* 399.6738, pp. 781–784.

Karran, Alexander et al. (2019). "Towards a hybrid passive BCI for the modulation of sustained attention using EEG and fNIRS". In: *Frontiers in Human Neuroscience* 12.November, pp. 1–20.

Katahira, Kenji et al. (2018). "EEG correlates of the flow state: A combination of increased frontal theta and moderate frontocentral alpha rhythm in the mental arithmetic task". In: *Frontiers in Psychology* 9, pp. 1–11.

Kaye, Linda K. (2016). "Exploring flow experiences in cooperative digital gaming contexts". In: *Computers in Human Behavior* 55, pp. 286–291.

Keeler, Jason R. et al. (2015). "The neurochemistry and social flow of singing: bonding and oxytocin". In: *Frontiers in Human Neuroscience* 9.September, pp. 1–10.

Keil, Andreas et al. (2014). "Committee report: Publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography". In: *Psychophysiology* 51.1, pp. 1–21.

Keith, Mark et al. (2014). "The Effects of Team Flow on Performance: A Video Game Experiment". In: *SIGHCI 2014 Proceedings*, pp. 1–5.

Keith, Mark et al. (2016). "The Effects of Video Gaming on Work Group Performance". In: *Proceedings of the 37th International Conference on Information Systems (ICIS)*, pp. 1–20.

Keller, Johannes (2016). "The Flow Experience Revisited: The Influence of Skills-Demands-Compatibility on Experiential and Physiological Indicators". In: *Flow Experience*. Ed. by László Harmat et al., pp. 351–374.

Keller, Johannes and Herbert Bless (2008). "Flow and regulatory compatibility: an experimental approach to the flow model of intrinsic motivation." In: *Personality and social psychology bulletin* 34.2, pp. 196–209.

Keller, Johannes et al. (2011). "Physiological aspects of flow experiences: Skills-demand-compatibility effects on heart rate variability and salivary cortisol". In: *Journal of Experimental Social Psychology* 47.4, pp. 849–852.

Kendall, Maurice G and B Babington Smith (1939). "The problem of m rankings". In: *The annals of mathematical statistics* 10.3, pp. 275–287.

Kennedy, Patrick, David B. Miele, and Janet Metcalfe (2014). "The cognitive antecedents and motivational consequences of the feeling of being in the zone". In: *Consciousness and Cognition* 30, pp. 48–61.

Kerner, W and J Brückel (2011). "Praxisempfehlung der Deutschen Diabetes Gesellschaft zur Definition, Klassifikation und Diagnostik des Diabetes mellitus". In: *Diabetologie und Stoffwechsel* 6.2, pp. 107–110.

Kerr, Norbert L. and R. Scott Tindale (2004). "Group Performance and Decision Making". In: *Annual Review of Psychology* 55.1, pp. 623–655.

Kitchenham, Barbara et al. (2009). "Systematic literature reviews in software engineering - A systematic literature review". In: *Information and Software Technology* 51.1, pp. 7–15.

Kivikangas, J Matias (2006). "Psychophysiology of Flow Experience: An Explorative Study". Master's Thesis. University of Helsinki, pp. 1–32.

Klarkowski, Madison (2016). "Psychophysiology of Challenge in Play: EDA and Self-Reported Arousal". In: *CHI Extended Abstracts on Human Factors in Computing Systems*, pp. 1930–1936.

— (2017). "The Psychophysiologial Evaluation of the Player Experience". Doctoral Dissertation. Queensland University of Technology, p. 274.

Klasen, Martin et al. (2011). “Neural contributions to flow experience during video game playing”. In: *Social Cognitive and Affective Neuroscience* 7.4, pp. 485–495.
Klimesch, Wolfgang (1999). “EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis”. In: *Brain Research Reviews* 1.29, pp. 169–195.
Kloker, Simon (2020). “Non-addictive information systems”. In: *Information Systems Frontiers* 22.3, pp. 1000–1007.
Knierim, Michael T., Mario Nadj, and Christof Weinhardt (2019). “Flow and optimal difficulty in the portable EEG: On the potentiality of using personalized frequency ranges for state detection”. In: *Proceedings of the 3rd International Conference on Computer-Human Interaction Research and Applications (CHIRA)*, pp. 183–190.
Knierim, Michael T et al. (2017a). “Designing live biofeedback for groups to support emotion management in digital collaboration”. In: *International Conference on Design Science Research in Information System and Technology*, pp. 479–484.
Knierim, Michael T et al. (2017b). “Emotion Feedback in Small Group Collaboration: A Research Agenda for Group Emotion Management Support Systems”. In: *Proceedings of the 17th International Conference on Group Decision and Negotiation*, pp. 1–12.
Knierim, Michael T. et al. (2017c). “The psychophysiology of flow: A systematic review of peripheral nervous system features”. In: *Proceedings of the 9th Retreat on NeuroIS*, pp. 109–120.
Knierim, Michael T et al. (2018a). “Exploring Flow Psychophysiology in Knowledge Work”. In: *Proceedings of the 10th Retreat on NeuroIS*, pp. 239–249.
Knierim, Michael T et al. (2018b). “Flow Neurophysiology in Knowledge Work: Electroencephalographic Observations from Two Cognitive Tasks”. In: *Proceedings of the 5th International Conference on Physiological Computing Systems (PhyCS)*, pp. 42–53.
Knierim, Michael Thomas et al. (2019). “Flow in knowledge work groups - Autonomy as a driver or digitally mediated communication as a limiting factor?” In: *Proceedings of the 40th International Conference on Information Systems (ICIS)*, pp. 1–17.
Knippenberg, Daan van and Michaéla C. Schippers (2007). “Work Group Diversity”. In: *Annual Review of Psychology* 58.1, pp. 515–541.
Kosti, Makrina Viola et al. (2018). “Towards an affordable brain computer interface for the assessment of programmers’ mental workload”. In: *International Journal of Human Computer Studies* 115.1, pp. 52–66.
Kramer, D (2007). “Predictions of performance by EEG and skin conductance”. In: *Indiana Undergraduate Journal of Cognitive Science* 2, pp. 3–13.
Krigolson, Olave E. et al. (2017). “Choosing MUSE: Validation of a low-cost, portable EEG system for ERP research”. In: *Frontiers in Neuroscience* 11.MAR, pp. 1–10.
Krol, Laurens R, Pim Haselager, and Thorsten O Zander (2019). “Cognitive and affective probing: A tutorial and review of active learning for neuroadaptive technology”. In: *Journal of Neural Engineering*, pp. 1–22.
Labonté-Lemoyne, Élise et al. (2016). “Are We in Flow? Neurophysiological Correlates of Flow States in a Collaborative Game”. In: *Proceedings of the 2016 CHI Conference*, pp. 1980–1988.
Lance, By Brent J et al. (2012). “Brain-Computer Interface Technologies in the Coming Decades”. In: *Proceedings of the IEEE* 100 (Special Centennial Issue), pp. 1585–1599.
Latham, Gary P. and Edwin A. Locke (2006). “Enhancing the Benefits and Overcoming the Pitfalls of Goal Setting”. In: *Organizational Dynamics* 35.4, pp. 332–340.

Lawlor, K. Blaine (2012). “Smart Goals: How the Application of Smart Goals can Contribute to Achievement of Student Learning Outcomes”. In: *Journal of Developments in Business Simulation and Experiential Learning* 39, pp. 259–267.

Lee, Chiawen, Kirk Damon Aiken, and Huang Chia Hung (2012). “Effects of college students’ video gaming behavior on self-concept clarity and flow”. In: *Social Behavior and Personality: An International Journal* 40.4, pp. 673–679.

Léger, Pierre Majorique et al. (2014). “Neurophysiological correlates of cognitive absorption in an enactive training context”. In: *Computers in Human Behavior* 34, pp. 273–283.

Li, Mengxiang et al. (2014). “Enhancing User-Game Engagement Through Software Gaming Elements”. In: *Journal of Management Information Systems* 30.4, pp. 115–150.

Loudon, Gareth and Dimitrios Zampelis (2017). “Technology Demo of using Real-time Biofeedback of Heart Rate Variability Measures to Track and Help Improve Levels of Attention and Relaxation”. In: *Proceedings of the 2017 ACM SIGCHI Conference on Creativity and Cognition - C&C ’17*, pp. 257–259.

Løvoll, Helga S. and Joar Vittersø (2014). “Can Balance be Boring? A Critique of the "Challenges Should Match Skills" Hypotheses in Flow Theory”. In: *Social Indicators Research* 115.1, pp. 117–136.

Lux, Ewa et al. (2018). “Live Biofeedback as a User Interface Design Element: A Review of the Literature”. In: *Communications of the Association for Information Systems* 43.1, pp. 257–296.

Lyubomirsky, Sonja, Laura King, and Ed Diener (2005). “The benefits of frequent positive affect: Does happiness lead to success?” In: *Psychological Bulletin* 131.6, pp. 803–855.

Magyaródi, Tímea and Attila Oláh (2017). “The Effect of Social Interaction on Flow Experience”. In: *International Journal of Psychology & Behavior Analysis* 3.126, pp. 1–5.

Maier, Marco et al. (2019). “DeepFlow: Detecting optimal user experience from physiological data using deep neural networks”. In: *IJCAI International Joint Conference on Artificial Intelligence*, pp. 1415–1421.

Makowski, Dominique (2016). *NeuroKit*. URL: `https://github.com/neuropsychology/NeuroKit.py`.

Malik, M et al. (1996). “Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Task Force of the European Society of Cardiology.” In: *European Heart Journal* 17, pp. 354–381.

Mannell, Roger C and William Bradley (1986). “Does greater freedom always lead to greater leisure? Testing a person x environment model of freedom and leisure”. In: *Journal of Leisure Research* 18.4, pp. 215–230.

Manzano, Örjan de et al. (2010). “The psychophysiology of flow during piano playing”. In: *Emotion* 10.3, pp. 301–311.

Manzano, Örjan de et al. (2013). “Individual differences in the proneness to have flow experiences are linked to dopamine D2-receptor availability in the dorsal striatum”. In: *NeuroImage* 67, pp. 1–6.

Marr, Arthur J (2001). “In the Zone: A Biobehavioral Theory of the Flow Experience”. In: *Athletic Insight: Online Journal of Sport Psychology* 3.1, pp. 1–8.

Martin, Andrew J and Susan A Jackson (2008). “Brief approaches to assessing task absorption and enhanced subjective experience : Examining ‘short’ and ‘core’ flow in diverse performance domains”. In: *Motivation and Emotion* 32.3, pp. 141–157.

Martin, Kelly D and Patrick E Murphy (2017). “The role of data privacy in marketing”. In: *Journal of the Academy of Marketing Science* 45.2, pp. 135–155.

Martinez, Constantino Antonio Garc'ia et al. (2017). *Heart Rate Variability Analysis with the R Package RHRV*. Cham: Springer International Publishing, pp. 1–173.

Mather, Mara and Julian F Thayer (2018). "How heart rate variability affects emotion regulation brain networks". In: *Current opinion in behavioral sciences* 19, pp. 98–104.

Mauri, Maurizio et al. (2011). "Why Is Facebook So Successful? Psychophysiological Measures Describe a Core Flow State While Using Facebook". In: *Cyberpsychology, Behavior, and Social Networking* 14.12, pp. 723–731.

Mehmood, Abid et al. (2016). "Protection of big data privacy". In: *IEEE Access* 4, pp. 1821–1834.

Michels, Lars et al. (2010). "Simultaneous EEG-fMRI during a working memory task: Modulations in low and high frequency bands". In: *PLoS ONE* 5.4, pp. 1–15.

Milnor, William R (1990). *Cardiovascular Physiology*. New York, Oxford: Oxford University Press, pp. 1–489.

Moeller, Aleidine J, Janine M Theiler, and Chaorong Wu (1988). "Goal Setting and Student Achievement: A Longitudinal Study". In: *The Modern Language Journal* 96.2, pp. 153–169.

Mognon, Andrea et al. (2011). "ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features". In: *Psychophysiology* 48, pp. 229–240.

Moller, Arlen C., Brian P. Meier, and Robert D. Wall (2010). "Developing an Experimental Induction of Flow: Effortless Action in the Lab". In: *Effortless Attention*. 6, pp. 191–204.

Moneta, G B and M Csikszentmihalyi (1996). "The effect of perceived challenges and skills on the quality of subjective experience." In: *Journal of personality* 64.2, pp. 275–310.

Moneta, Giovanni B (2012). "On the Measurement and Conceptualization of Flow". In: *Advances in Flow Research*. Ed. by Stefan Engeser, pp. 23–50.

Moneta, Giovanni B. (2017). "Validation of the Short Flow in Work Scale (SFWS)". In: *Personality and Individual Differences* 109, pp. 83–88.

Morelli, Davide et al. (2019). "Analysis of the Impact of Interpolation Methods of Missing RR-intervals Caused by Motion Artifacts on HRV Features Estimations". In: *Sensors* 19.14, pp. 3163–3177.

Mullen, Tim R. et al. (2015). "Real-time Neuroimaging and Cognitive Monitoring Using Wearable Dry EEG". In: *IEEE Transactions on Biomedical Engineering* 62.11, pp. 2553–2567.

Müller, Sebastian C. and Thomas Fritz (2015). "Stuck and frustrated or in flow and happy: Sensing developers' emotions and progress". In: *International Conference on Software Engineering*, pp. 688–699.

Müller-Putz, Gernot R., René Riedl, and Selina C. Wriessnegger (2015). "Electroencephalography (EEG) as a research tool in the information systems discipline: Foundations, measurement, and applications". In: *Communications of the Association for Information Systems* 37.1, pp. 911–948.

Murch, W. Spencer, Stephanie W. M. Chu, and Luke Clark (2017). "Measuring the Slot Machine Zone With Attentional Dual Tasks and Respiratory Sinus Arrhythmia". In: *Psychology of Addictive Behaviors*, pp. 1–11.

Musca, Serban C. et al. (2011). "Data with hierarchical structure: Impact of intraclass correlation and sample size on Type-I error". In: *Frontiers in Psychology* 2.APR, pp. 1–6.

Muthukumaraswamy, Suresh D. (2013). "High-frequency brain activity and muscle artifacts in MEG/EEG: A review and recommendations". In: *Frontiers in Human Neuroscience* 7.MAR, pp. 1–11.

Nacke, Lennart E., Mark N. Grimshaw, and Craig A. Lindley (2010). “More than a feeling: Measurement of sonic user experience and psychophysiology in a first-person shooter game”. In: *Interacting with Computers* 22.5, pp. 336–343.

Nacke, Lennart E and Craig A. Lindley (2008). “Flow and Immersion in First-person Shooters: Measuring the Player’s Gameplay Experience”. In: *Proceedings of the 2008 Conference on Future Play: Research, Play, Share*. New York: ACM Press, pp. 81–88.

Nacke, Lennart E. and Craig a. Lindley (2010). “Affective Ludology, Flow and Immersion in a First-Person Shooter: Measurement of Player Experience”. In: *Loading...* 3.5, pp. 1–21.

Nakamura, Jeanne and Mihaly Csikszentmihalyi (2009). “Flow theory and research”. In: *Oxford Handbook of Positive Psychology*. Ed. by S. Lopez and C. R. Snyder. New York: Oxford Univ. Press, pp. 195–206.

Naumann, Laura et al. (2016). “Prediction of Difficulty Levels in Video Games from Ongoing EEG”. In: *International Workshop on Symbiotic Interaction*, pp. 125–136.

Nikolin, Stevan et al. (2018). “Safety of repeated sessions of transcranial direct current stimulation: a systematic review”. In: *Brain Stimulation* 11.2, pp. 278–288.

Olsson, Bengt Köping and László Harmat (2018). “State Flow, Creativity and Flow Synchronization During Group Based Problem Solving Task”. In: *5th Participatory Innovation Conference*, pp. 146–153.

Oostenveld, Robert and Peter Praamstra (2001). “The five percent electrode system for high-resolution EEG and ERP measurements”. In: *Clinical neurophysiology* 112.4, pp. 713–719.

Ortiz de Guinea, Ana, Ryad Titah, and Pierre-Majorique Léger (2013). “Measure for Measure: A two study multi-trait multi-method investigation of construct validity in IS research”. In: *Computers in Human Behavior* 29.3, pp. 833–844.

Palumbo, Richard V et al. (2017). “Interpersonal Autonomic Physiology: A Systematic Review of the Literature”. In: *Personality and Social Psychology Review* 21.2, pp. 99–141.

Parasuraman, Raja and Glenn F Wilson (2008). “Putting the brain to work: Neuroergonomics past, present, and future”. In: *Human factors* 50.3, pp. 468–474.

Parent-Thirion, A et al. (2015). *Sixth European working conditions survey (EWCS)*. Tech. rep. Office for Official Publications of the European Communities, Luxembourg, LU.

Partington, Sarah, Elizabeth Partington, and Steve Olivier (2009). “The Dark Side of flow: A qualitative study of dependence in big wave surfing”. In: *Sport Psychologist* 23.2, pp. 170–185.

Peifer, Corinna (2012). “Psychophysiological Correlates of Flow-Experience”. In: *Advances in Flow Research*. Ed. by S. Engeser. New York: Springer Science, pp. 139–164.

Peifer, Corinna and Gina Zipp (2019). “All at once? The effects of multitasking behavior on flow and subjective performance”. In: *European Journal of Work and Organizational Psychology*, pp. 1–9.

Peifer, Corinna et al. (2014). “The relation of flow-experience and physiological arousal under stress - Can u shape it?” In: *Journal of Experimental Social Psychology* 53.1, pp. 62–69.

Peifer, Corinna et al. (2015). “Cortisol effects on flow-experience”. In: *Psychopharmacology* 232.6, pp. 1165–1173.

Peifer, Corinna et al. (2019). “Thieves of Flow: How Unfinished Tasks at Work are Related to Flow Experience and Wellbeing”. In: *Journal of Happiness Studies*, pp. 1–20.

Peters, Dorian, Rafael A. Calvo, and Richard M. Ryan (2018). "Designing for motivation, engagement and wellbeing in digital experience". In: *Frontiers in Psychology* 9.May, pp. 1–15.

Petty, Brent G (2015). *Basic Electrocardiography*. New York: Springer, pp. 1–152.

Picard, Rosalind W (2003). "Affective computing: challenges". In: *International Journal of Human-Computer Studies* 59.1-2, pp. 55–64.

Picton, T W et al. (2000). "Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria". In: *Psychophysiology* 37, pp. 127–152.

Pivik, R. T. et al. (1993). "Guidelines for the recording and quantitative analysis of electroencephalographic activity in research contexts". In: *Psychophysiology* 30.1, pp. 547–558.

Pizzagalli, Diego A (2007). "Electroencephalography and High-Density Electrophysiological Source Localization". In: *The Handbook of Psychophysiology*. Ed. by John T. Cacioppo, Louis G. Tassinary, and Gary G. Berntson, pp. 56–84.

Prinzel, Lawrence J et al. (2000). "A Closed-Loop System for Examining Psychophysiological Measures for Adaptive Task Allocation". In: *The International Journal of Aviation Psychology* 10.4, pp. 393–410.

Quinn, Ryan W (2005). "Flow in Knowledge Performance Experience". In: *Administrative Science Quarterly* 50, pp. 610–641.

Radüntz, Thea (2017). "Dual frequency head maps: A new method for indexing mental workload continuously during execution of cognitive tasks". In: *Frontiers in Physiology* 8.DEC, pp. 1–15.

— (2020). "The Effect of Planning, Strategy Learning, and Working Memory Capacity on Mental Workload". In: *Scientific Reports* 10.1, pp. 1–10.

Ramirez, Rafael et al. (2015). "Musical neurofeedback for treating depression in elderly people". In: *Frontiers in Neuroscience* 9.OCT, pp. 1–10.

Ratti, Elena et al. (2017). "Comparison of medical and consumer wireless EEG systems for use in clinical trials". In: *Frontiers in Human Neuroscience* 11.August, pp. 1–7.

Ren, Shiquan, Shuqin Yang, and Shenghan Lai (2006). "Intraclass correlation coefficients and bootstrap methods of hierarchical binary outcomes". In: *Statistics in Medicine* 25.20, pp. 3576–3588.

Rheinberg, Falko (1991). "Flow-experience when motorcycling: A study of a special human condition". In: *Safety-Environment-Future. Proceedings of the 1991 International Motorcycle Conference (forschungsheft nr 7)*, pp. 349–362.

Rheinberg, Falko and Regina Vollmeyer (2003). "Flow-Erleben in einem Computerspiel unter experimentell variierten Bedingungen". In: *Zeitschrift für Psychologie* 4, pp. 161–170.

Rich, Grant J. (2013). "Finding Flow: The History and Future of a Positive Psychology Concept". In: *Positive Psychology: Advances in Understanding Adult Motivation*. Ed. by Jan D. Sinnott. Springer, pp. 43–60.

Richter, Michael and Kate Slade (2017). "Interpretation of physiological indicators of motivation: Caveats and recommendations". In: *International Journal of Psychophysiology* 119.April, pp. 4–10.

Riedl, René, Fred D. Davis, and Alan R. Hevner (2014). "Towards a NeuroIS Research Methodology: Intensifying the Discussion on Methods, Tools, and Measurement". In: *Journal of the Association for Information Systems* 15.10, pp. i–xxxv.

Riedl, René and Pierre-Majorique Léger (2016). "Tools in NeuroIS Research: An Overview". In: pp. 47–72.

Rissler, Raphael et al. (2017). "Towards an integrative theoretical framework of IT-mediated interruptions". In: *Proceedings of the 25th European Conference on Information Systems (ECIS)*, pp. 1950–1967.

Rissler, Raphael; et al. (2018). "Got Flow? Using Machine Learning on Physiological Data to Classify Flow". In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–6.

Ross, Scott R. and Heidi N. Keiser (2014). "Autotelic personality through a five-factor lens: Individual differences in flow-propensity". In: *Personality and Individual Differences* 59, pp. 3–8.

Rouast, Philipp V., Marc Adam, and Raymond Chiong (2018). "Deep Learning for Human Affect Recognition: Insights and New Developments". In: *IEEE Transactions on Affective Computing* 14.8, pp. 1–20.

Roy, Yannick et al. (2019). "Deep learning-based electroencephalography analysis: A systematic review". In: *Journal of Neural Engineering* 16.5.

Ryan, Richard and Edward Deci (2002). "Self-determination theory: An organismic dialectical perspective". In: *Handbook of Self-Determination Research*. Ed. by Edward L Deci and Richard M Ryan. Rochester: The University of Rochester Press, pp. 3–33.

Sadlo, Gaynor (2016). "Towards a Neurobiological Understanding of Reduced Self--Awareness During Flow: An Occupational Science Perspective". In: *Flow Experience*. Ed. by László Harmat, pp. 375–388.

Salanova, Marisa et al. (2014). "Flowing together: a longitudinal study of collective efficacy and collective flow among workgroups." In: *The Journal of psychology* 148.4, pp. 435–455.

Sato, Ikuya (1988). "Bosozoku: Flow in Japanese motorcycle gangs". In: *Optimal experience - Psychological Studies of Flow in Consciousness*. Cambridge: University Press, pp. 92–117.

Schüler, Julia (2007). "Arousal of flow experience in a learning setting and its effects on exam performance and affect". In: *Zeitschrift fur Padagogische Psychologie* 21.3-4, pp. 217–227.

— (2012). "The Dark Side of the Moon". In: *Advances in Flow Research*. Ed. by Stefan Engeser. Springer, pp. 123–137.

Schüler, Julia and Mirjam Pfenninger (2011). "Flow impairs risk perception in kayakers". In: *Sport Psychology*. Ed. by Bruce D Geranto, pp. 237–246.

Schulz-Hardt, Stefan et al. (2006). "Group decision making in hidden profile situations: dissent as a facilitator for decision quality." In: *Journal of personality and social psychology* 91.6, pp. 1080–1093.

Seligman, Martin E. P. and Mihaly Csikszentmihalyi (2000). "Positive psychology: An introduction." In: *American Psychologist* 55.1, pp. 5–14.

Selye, Hans (1980). *Selye's Guide to Stress Research*. New York: Van Nostrand Reinhold, pp. 1–350.

Seneviratne, Suranga et al. (2017). "A survey of wearable devices and challenges". In: *IEEE Communications Surveys & Tutorials* 19.4, pp. 2573–2620.

Severiano, Ana et al. (2011). "Evaluation of Jackknife and Bootstrap for defining confidence intervals for pairwise agreement measures". In: *PLoS ONE* 6.5, pp. 1–11.

Shaffer, Fred and J. P. Ginsberg (2017). "An Overview of Heart Rate Variability Metrics and Norms". In: *Frontiers in Public Health* 5.September, pp. 1–17.

Shearer, Peter B (2016). "Physiological Detection of Flow". Doctoral Dissertation. The University of South Dakota.

Sheldon, Kennon M. and Jonathan C. Hilpert (2012). "The balanced measure of psychological needs (BMPN) scale: An alternative domain general measure of need satisfaction". In: *Motivation and Emotion* 36.4, pp. 439–451.

Shull, Peter B and Dana D Damian (2015). "Haptic wearables as sensory replacement, sensory augmentation and trainer: A review". In: *Journal of NeuroEngineering and Rehabilitation* 12.59, pp. 1–13.

Siegel, Markus, Tobias H. Donner, and Andreas K. Engel (2012). "Spectral fingerprints of large-scale neuronal interactions". In: *Nature Reviews Neuroscience* 13.2, pp. 121–134.

Silvestrini, Nicolas (2017). "Psychological and neural mechanisms associated with effort-related cardiovascular reactivity and cognitive control: An integrative approach". In: *International Journal of Psychophysiology* 119, pp. 11–18.

Sinha, Aniruddha et al. (2015). "Dynamic assessment of learners' mental state for an improved learning experience". In: *Proceedings - Frontiers in Education Conference, FIE*, pp. 1–9.

Smith, Ezra E et al. (2017). "Assessing and Conceptualizing Frontal EEG Asymmetry: An Updated Primer on Recording, Processing, Analyzing, and Interpreting Frontal Alpha Asymmetry". In: *International Journal of Psychophysiology* 111, pp. 98–114.

Soltész, Péter et al. (2014). "Towards understanding the electrophysiology of flow". In: *Flow, Emotional Intelligence and Psychological Immunity: Empirical Studies in Positive Psychological Perspective*, pp. 85–92.

Sors, Arnaud et al. (2018). "A convolutional neural network for sleep stage scoring from raw single-channel EEG". In: *Biomedical Signal Processing and Control* 42, pp. 107–114.

Spurlin, Sam and Mihaly Csikszentmihalyi (2017). "Will Work Ever be Fun Again?" In: *Flow at Work: Measurement and Implications*. Ed. by Clive J. Fullagar and Antonella Delle Fave, pp. 176–187.

Stavrou, Nektarios A (2008). "Intrinsic motivation, extrinsic motivation and amotivation: Examining self-determination theory from flow theory perspective". In: pp. 1–24.

Stavrou, Nektarios A et al. (2007). "Flow Experience and Athletes' Performance With Reference to the Orthogonal Model of Flow". In: *The Sport Psychologist* 21, pp. 438–457.

Stevens, Ronald, Polemnia Amazeen, and Aaron Likens (2013). "The organizational neurodynamics of teams". In: *Nonlinear Dynamics, Psychology, and Life Sciences* 17.1, pp. 67–86.

Stevens, Ronald H. et al. (2012). "Cognitive neurophysiologic synchronies: What can they contribute to the study of teamwork?" In: *Human Factors* 54.4, pp. 489–502.

Stranger, Mark (1999). "The aesthetics of risk: A study of surfing". In: *International review for the sociology of sport* 34.3, pp. 265–276.

Strube, Michael J and L C Newman (2007). "Psychometrics". In: *Handbook of Psychophysiology* 3. Ed. by John T Cacioppo, Louis G Tassinary, and Gary G Berntson, pp. 789–811.

Swann, Christian et al. (2012). "A systematic review of the experience, occurrence, and controllability of flow states in elite sport". In: *Psychology of Sport & Exercise* 13.6, pp. 807–819.

Swann, William B. et al. (2009). "Identity Fusion: The Interplay of Personal and Social Identities in Extreme Group Behavior". In: *Journal of Personality and Social Psychology* 96.5, pp. 995–1011.

Szucs, Denes and John P.A. Ioannidis (2017). "Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature". In: *PLoS Biology* 15.3, pp. 1–18.

Tams, Stefan et al. (2014). "NeuroIS—Alternative or Complement to Existing Methods? Illustrating the Holistic Effects of Neuroscience and Self-Reported Data in the Context of Technostress Research". In: *Journal of the Association for Information Systems* 15, pp. 723–753.

Tanaka, Hirofumi, Kevin D Monahan, and Douglas R Seals (2001). "Age-Predicted Maximal Heart Rate Revisited". In: *Journal of the American College of Cardiology* 37.1, pp. 153–156.

Teplan, M (2002). "Fundamentals of EEG Measurement". In: *Measurement Science Review* 2.2, pp. 1–11.

Thayer, Julian F. et al. (2009). "Heart rate variability, prefrontal neural function, and cognitive performance: The neurovisceral integration perspective on self-regulation, adaptation, and health". In: *Annals of Behavioral Medicine* 37.2, pp. 141–153.

Thayer, Julian F. et al. (2012). "A meta-analysis of heart rate variability and neuroimaging studies: Implications for heart rate variability as a marker of stress and health". In: *Neuroscience and Biobehavioral Reviews* 36.2, pp. 747–756.

Tian, Yu et al. (2017). "Physiological signal analysis for evaluating flow during playing of computer games of varying difficulty". In: *Frontiers in Psychology* 8, pp. 1–10.

Tkaczyk, Dominika et al. (2015). "CERMINE: Automatic extraction of structured metadata from scientific literature". In: *International Journal on Document Analysis and Recognition* 18.4, pp. 317–335.

Tomczak, M and E Tomczak (2014). "The need to report effect size estimates revisited. An overview of some recommended measures of effect size". In: *Trends in Sport Sciences* 21.1, pp. 19–25.

Toppi, Jlenia et al. (2016). "Investigating cooperative behavior in ecological settings: an EEG hyperscanning study". In: *PloS one* 11.4, e0154236.

Tozman, Tahmine and Corinna Peifer (2016). "Experimental Paradigms to Investigate Flow-Experience and Its Psychophysiology: Inspired from Stress Theory and Research". In: *Flow Experience*, pp. 329–350.

Tozman, Tahmine, Yichelle Y. Zhang, and Regina Vollmeyer (2017). "Inverted U-Shaped Function Between Flow and Cortisol Release During Chess Play". In: *Journal of Happiness Studies* 18.1, pp. 247–268.

Tozman, Tahmine et al. (2015). "Understanding the psychophysiology of flow: A driving simulator experiment to investigate the relationship between flow and heart rate variability". In: *Computers in Human Behavior* 52, pp. 408–418.

Trull, Timothy J and Ulrich W Ebner-Priemer (2009). "Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research". In: *Psychological Assessment* 21, pp. 457–462.

Tse, Dwight C K, Jeanne Nakamura, and Mihaly Csikszentmihalyi (2020). "Living well by flowing well: The indirect effect of autotelic personality on well-being through flow experience". In: *The Journal of Positive Psychology*, pp. 1–12.

Tse, Dwight C.K. et al. (2016). "Teamwork and flow proneness mitigate the negative effect of excess challenge on flow state". In: *Journal of Positive Psychology* 13.3, pp. 284–289.

Turner, Benjamin O. et al. (2018). "Small sample sizes reduce the replicability of task-based fMRI studies". In: *Communications Biology* 1.1, pp. 1–10.

Ullén, Fredrik et al. (2010). "The Physiology of Effortless Attention: Correlates of State Flow and Flow Proneness". In: *Effortless Attention: A New Perspective in the*

Cognitive Science of Attention and Action. Ed. by Brian Bruya. Cambridge: MIT Press, pp. 205–217.
Ullén, Fredrik et al. (2012). "Proneness for psychological flow in everyday life: Associations with personality and intelligence". In: *Personality and Individual Differences* 52.2, pp. 167–172.
Ulrich, Martin, Johannes Keller, and Georg Grön (2016a). "Dorsal Raphe Nucleus Down-Regulates Medial Prefrontal Cortex during Experience of Flow". In: *Frontiers in Behavioral Neuroscience* 10, pp. 1–9.
— (2016b). "Neural signatures of experimentally induced flow experiences identified in a typical fMRI block design with BOLD imaging". In: *Social Cognitive and Affective Neuroscience* 11.3, pp. 496–507.
Ulrich, Martin et al. (2014). "Neural correlates of experimentally induced flow experiences". In: *NeuroImage* 86, pp. 194–202.
Ulrich, Martin et al. (2018). "The neural correlates of flow experience explored with transcranial direct current stimulation". In: *Experimental brain research* 236.12, pp. 3223–3237.
Valentini, Mariaconsuelo and Gianfranco Parati (2009). "Variables Influencing Heart Rate". In: *Progress in Cardiovascular Diseases* 52.1, pp. 11–19.
Van den Hout, Jef J. J. (2016). "Team Flow From Concept to Application". Doctoral Dissertation. Eindhoven University of Technology.
Vink, Gerko et al. (2014). "Predictive mean matching imputation of semicontinuous variables". In: *Statistica Neerlandica* 68.1, pp. 61–90.
Wageman, Ruth, J Richard Hackman, and Erin Lehman (2005). "Team Diagnostic Survey: Development of an Instrument". In: *The Journal of Applied Behavioral Science* 41.4, pp. 373–398.
Walker, Charles J. (2010). "Experiencing flow: Is doing it together better than doing it alone?" In: *The Journal of Positive Psychology* 5.1, pp. 3–11.
Wan, Chin-Sheng and Wen-Bin Chiou (2006). "Psychological motives and online games addiction: Atest of flow theory and humanistic needs theory for taiwanese adolescents". In: *CyberPsychology & Behavior* 9.3, pp. 317–324.
Wang, Bei et al. (2019). "Bootstrap ICC estimators in analysis of small clustered binary data". In: *Computational Statistics* 34.4, pp. 1765–1778.
Wang, Richard Y and Diane M Strong (1996). "Beyond Accuracy: What Data Quality Means to Data Consumers". In: *Journal of Management Information Systems* 12.4, pp. 5–34.
Wascher, Edmund et al. (2014). "Frontal theta activity reflects distinct aspects of mental fatigue". In: *Biological Psychology* 96, pp. 57–65.
Wascher, Edmund et al. (2019). "Evaluating Mental Load During Realistic Driving Simulations by Means of Round the Ear Electrodes". In: *Frontiers in Neuroscience* 13.September, pp. 1–11.
Weber, René, Richard Huskey, and Britney Craighead (2016). "Flow experiences and well-being: A media neuroscience perspective". In: *Handbook of media use and well-being: International perspectives on theory and research on positive media effects*, pp. 183 –196.
Weber, René et al. (2009). "Theorizing flow and media enjoyment as cognitive synchronization of attentional and reward networks". In: *Communication Theory* 19.4, pp. 397–422.
Webster, Jane and Richard T Watson (2002). "Analyzing the Past to Prepare for the Future: Writing a Literature Review." In: *Management of Information Systems Quarterly* 26.2, pp. xiii –xxiii.

Weippert, Matthias et al. (2010). "Comparison of three mobile devices for measuring R – R intervals and heart rate variability: Polar S810i , Suunto t6 and an ambulatory ECG system". In: *European Journal of Applied Physiology* 109.4, pp. 779–786.

Westcott-Baker, A and R Weber (2012). "The synchronization theory of flow: Consequences of a limited capacity for attention and consciousness". In: *Video game play and consciousness*, pp. 119–141.

Wildman, Jessica L. et al. (2012). "Task Types and Team-Level Attributes: Synthesis of Team Classification Literature". In: *Human Resource Development Review* 11.1, pp. 97–129.

Wilhelm, Frank H. et al. (2006). "Distinguishing emotional from physical activation in ambulatory psychophysiological monitoring". In: *Biomedical Sciences Instrumentation* 42, pp. 458–463.

Williams, E. J. (1949). "Experimental Designs Balanced for the Estimation of Residual Effects of Treatments". In: *Australian Journal of Chemistry* 2.2, pp. 149–168.

Wilson, Edith E. and Giovanni B. Moneta (2016). "The flow metacognitions questionnaire (FMQ): A two factor model of flow metacognitions". In: *Personality and Individual Differences* 90, pp. 225–230.

Wilson, Edith Evelyn (2016). "Metacognitions of Flow Experience: Towards an Understanding of the Self-Regulation of Optimal Experience". Doctoral Dissertation. London Metropolitan University.

Wolf, Sebastian et al. (2015). "Motor skill failure or flow-experience? Functional brain asymmetry and brain connectivity in elite and amateur table tennis players". In: *Biological Psychology* 105, pp. 95–105.

Wolska, Agnieszka et al. (2019). "Which EEG electrodes should be considered for alertness assessment?" In: *CHIRA 2019 - Proceedings of the 3rd International Conference on Computer-Human Interaction Research and Applications*, pp. 40–49.

Woolley, Anita Williams et al. (2010). "Evidence for a collective intelligence factor in the performance of human groups." In: *Science* 330.6004, pp. 686–8.

Wuchty, Stefan, Benjamin F Jones, and Brian Uzzi (2007). "The increasing dominance of teams in production of knowledge." In: *Science* 316.5827, pp. 1036–9.

Yannakakis, Georgios N., Roddy Cowie, and Carlos Busso (2018). "The Ordinal Nature of Emotions: An Emerging Approach". In: *IEEE Transactions on Affective Computing*, pp. 1–20.

Yoo, Jerald et al. (2009). "A wearable ECG acquisition system with compact planar-fashionable circuit board-based shirt". In: *IEEE Transactions on Information Technology in Biomedicine* 13.6, pp. 897–902.

Yoshida, Kazuki et al. (2014). "Brain activity during the flow experience: A functional near-infrared spectroscopy study". In: *Neuroscience Letters* 573, pp. 30–34.

Yotsidi, Vasiliki et al. (2018). "'Add Flow to the Fire': Flow and Hope as a Shield against Burnout of Fire Service Workers". In: *Psychology* 9, pp. 1291–1305.

Zhang, Pengbo et al. (2018). "Learning spatial-spectral-temporal EEG features with recurrent 3D convolutional neural networks for cross-task mental workload assessment". In: *IEEE Transactions on neural systems and rehabilitation engineering* 27.1, pp. 31–42.

— (2019). "Learning Spatial-Spectral-Temporal EEG Features With Recurrent 3D Convolutional Neural Networks for Cross-Task Mental Workload Assessment". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 27.1, pp. 31–42.

Zumeta, Larraitz N. et al. (2016). "Collective efficacy in sports and physical activities: Perceived emotional synchrony and shared flow". In: *Frontiers in Psychology* 6.JAN, pp. 1–12.

Zupic, Ivan and Tomaž Čater (2015). “Bibliometric Methods in Management and Organization”. In: *Organizational Research Methods* 18.3, pp. 429–472.