

Decoding the Functional Reorganization of the Aging Brain

By

Christian Johannes Gölz

A thesis presented for the degree of
Doctor rerum naturalium
(Dr. rer. nat)

Paderborn University
Faculty of natural sciences
2023

Declaration of Authorship

Acknowledgement

Danke.

Abstract

Dealing with an aging population is one of our society’s most significant challenges. Targeted interventions, early detection of unfavorable aging trajectories, and assistive technologies could all help to provide a solution but require simple, easy-to-use, and cost-effective methods to quantify and better understand age-related reorganization of the brain. The approach of this work was, therefore, to gain such an understanding by applying machine learning techniques to electroencephalography (EEG).

To this end, we used datasets obtained from studies using paradigms typical in aging research to investigate information processing during the performance of motor, sensory, and cognitive tasks. We analyzed these datasets using dimensionality reduction methods and classification algorithms to investigate the dedifferentiation of functional organization, i.e., the deterioration of selective information processing and the influence of lifestyle factors, as well as to gain exploratory insights into group patterns of age-related brain reorganization and their influencing factors.

In this way, we demonstrated that the dedifferentiation of the brain’s organization is reflected in the performance of classifiers trained to differentiate different task states or characteristics on the basis of EEG measurements, which opens up possibilities to quantify brain reorganization at the individual level. Following this, our results suggest that different brain systems show different patterns of reorganization that differ in their expression and age of onset. More specifically, dedifferentiation of the motor system was already evident in late middle-aged participants, whereas for elements of the attentional system, we found differences in the dynamics of differentiability only in older adults. In this context, the performance of classifiers trained to predict the group membership showed little misclassification between late middle-aged and older adults, which could indicate gross changes after retirement.

Furthermore, we characterized the influence of lifestyle factors, such as cardiorespiratory fitness and occupational expertise, on this reorganization of the aging brain and identified differential effects of the mentioned factors. More specifically, we found that older adults who exhibited higher cardiorespiratory fitness showed fewer indicators of dedifferentiation and less noise in brain activation patterns, whereas professional expertise was more likely to show up in highly individualized brain organization.

Altogether, machine learning methods were used to test hypotheses about age-related changes and generate new hypotheses. The possibility of quantifying age-related reorganization on an indi-

vidual level may have practical relevance for the development of markers for the early detection of unfavorable age trajectories and for the development of rehabilitation or prevention therapies. Furthermore, the demonstrated interactions between age-related brain reorganization and classifier performance could be important for developing assistive technology based on the use of brain signals and for which performance is crucial.

Zusammenfassung

Der Umgang mit einer alternden Bevölkerung ist eine der größten Herausforderungen für unsere Gesellschaft. Gezielte Interventionen, die frühzeitige Erkennung ungünstiger Altersverläufe und unterstützende Technologien könnten zu einer Lösung beitragen, erfordern jedoch einfache, leicht anzuwendende und kosteneffiziente Methoden zur Quantifizierung und zum besseren Verständnis altersbedingter Reorganisation des Gehirns. Der Ansatz dieser Arbeit war daher, ein solches Verständnis durch die Anwendung von Methoden des maschinellen Lernens auf EEG zu gewinnen.

Zu diesem Zweck haben wir Datensätze aus Studien verwendet, die typische Paradigmen der Altersforschung verwenden, um die Informationsverarbeitung während der Ausführung motorischer, sensorischer und kognitiver Aufgaben zu untersuchen. Wir analysierten diese Datensätze mit Methoden der Dimensionsreduktion und Klassifikationsalgorithmen, um die Dedifferenzierung der funktionellen Organisation, d.h. den Abbau der selektiven Informationsverarbeitung und den Einfluss von Lebensstilfaktoren zu untersuchen sowie explorative Einblicke in Gruppenmuster der altersbedingten Hirnreorganisation und deren Einflussfaktoren zu gewinnen.

Auf diese Weise konnten wir zeigen, dass sich die Dedifferenzierung der Gehirnorganisation in der Leistung von Klassifikatoren widerspiegelt, die darauf trainiert wurden, auf der Grundlage von EEG Messungen Aufgabenzustände oder -merkmale zu unterscheiden, was Möglichkeiten eröffnet, dies auf individueller Ebene zu quantifizieren. Unsere Ergebnisse deuten darauf hin, dass verschiedene Gehirnsysteme unterschiedliche Reorganisationsmuster aufweisen, die sich in ihrer Ausprägung und dem Alter ihres Auftretens unterscheiden. Im Einzelnen zeigte sich die Dedifferenzierung des motorischen Systems bereits bei Teilnehmern im späten mittleren Alter, während wir für Elemente des Aufmerksamkeitssystems Unterschiede in der Dynamik der Differenzierbarkeit erst bei älteren Erwachsenen fanden. In diesem Zusammenhang zeigte die Leistung von Klassifikatoren, die zur Vorhersage der Gruppenzugehörigkeit trainiert wurden, nur geringe Fehlklassifikationen zwischen Teilnehmern im späten mittleren Alter und älteren Erwachsenen, was auf grobe Veränderungen nach dem Eintritt in den Ruhestand hindeuten könnte.

Darüber hinaus haben wir den Einfluss von Lebensstilfaktoren, wie kardiorespiratorische Fitness und berufliche Expertise, auf diese Reorganisation des alternden Gehirns charakterisiert und unterschiedliche Auswirkungen der genannten Faktoren festgestellt. Genauer gesagt fanden wir heraus, dass ältere Erwachsene mit höherer kardiorespiratorischer Fitness weniger Indikatoren für Dedifferenzierung und weniger Rauschen in den Aktivierungsmustern des Gehirns aufwiesen, wohingegen

sich berufliche Expertise eher in einer hochgradig individualisierten Gehirnorganisation zeigte. Zusammenfassend wurden Methoden des maschinellen Lernens eingesetzt, um Hypothesen über altersbedingte Veränderungen zu testen und neue Hypothesen zu entwickeln. Die Möglichkeit, altersbedingte Reorganisation auf individueller Ebene zu quantifizieren, könnte praktische Bedeutung für die Entwicklung von Markern zur Früherkennung ungünstiger Altersverläufe und für die Entwicklung von Rehabilitations- oder Präventionstherapien haben. Darüber hinaus könnten die nachgewiesenen Wechselwirkungen zwischen altersbedingter Reorganisation des Gehirns und der Leistung von Klassifikatoren wichtig für die Entwicklung von assistiver Technologien sein, die auf der Nutzung von Hirnsignalen basieren und für die Leistung entscheidend ist.

Contents

Acknowledgement	ii
Abstract	ii
Zusammenfassung	ii
List of Figures	vi
List of Tables	vii
List of Abbreviations	ix
List of Scientific Contributions	x
1 General Introduction	1
1.1 Motivation	1
1.2 Outline	2
1.3 Aging	3
1.3.1 Age-related Reorganization of the Brain	3
1.3.2 Studying Brain Aging by Electroencephalography	6
1.4 Machine learning	9
1.4.1 Forms of machine learning	10
1.4.2 State-of-the-Art Approaches to Electroencephalographic Data	13
1.4.3 Applications in the Context of Aging Research	15
2 Aims and Scope	17
3 General Methodology	20
3.1 Datasets	20
3.1.1 Dataset I	20
3.1.2 Dataset II	22
3.1.3 Dataset III	22
3.1.4 Electroencephalography: Recording and Preprocessing	24
3.2 Machine Learning Procedures	24

3.2.1 Dimensionality Reduction	26
3.2.2 Classification	28
4 Summary of the Main Results	31
4.1 Research Article I	31
4.2 Research Article II	33
4.3 Research Article III	35
4.4 Research Article IV	36
5 General Discussion	39
5.1 On the Detection of Dedifferentiated Brain Organization	40
5.2 The Role of Lifestyle Factors	41
5.3 Exploratory Insights into Age-Related Reorganization	42
5.4 Methodological Considerations	43
5.4.1 Datasets	43
5.4.2 Machine Learning Approaches	44
5.5 Outlook and Practical Implications	46
5.6 Conclusion	48
References	49
Published Research Articles	62
Published Research Article I	63
Published Research Article II	76
Published Research Article III	88
Published Research Article IV	102

List of Figures

1.1	The computational model of dedifferentiation	5
1.2	The three main forms of machine learning.	10
1.3	Exemplary nested cross-validation procedure.	14
3.1	Schematic presentation of the force-tracking task conducted in Dataset I.	21
3.2	Schematic presentation of the flanker task conducted in Dataset II.	23
3.3	Schematic presentation of the motor, sensory and cognitive tasks conducted in Dataset III.	24
3.4	Machine learning approach used in this thesis	26
4.1	Main results of Published Research Article I.	32
4.2	Main results of Published Research Article II.	34
4.3	Main results of Published Research Article III	36
4.4	Main results of Published Research Article IV	37

List of Tables

2.1	Summary of the approaches followed.	19
3.1	Overview of datasets and participants in each research article.	25
3.2	Dimensionality reduction and classification methods utilized in each research article	26
3.3	Confusion matrix	28
3.4	Summary of metrics to evaluate model performance	29

List of Abbreviations

AI artificial intelligence.

BCI brain computer interface.

CRUNCH compensation-related utilization of neural circuits hypothesis.

CSP common spatial patterns.

DMD dynamic mode decomposition.

EEG electroencephalography.

ERM empirical risk minimization.

ERP event related potential.

FBCSP filter based common spatial patterns.

fMRI functional magnetic resonance imaging.

FN false negative.

FP false positive.

HAROLD hemispheric asymmetry reduction in older adults.

ICA independent component analysis.

LDA linear discriminant analysis.

MLE maximum likelihood estimation.

MRI magnet resonance imaging.

MVC maximum voluntary contraction.

MVPA multivariate pattern analysis.

PASA posterior–anterior shift in aging.

PCA principal component analysis.

STAC scaffolding theory of cognitive aging.

SVD singular value decomposition.

SVM support vector machine.

t-SNE t-distributed stochastic neighbor embedding.

TN true negative.

TP true positive.

UMAP uniform manifold approximation and projection for dimension reduction.

WHO World Health Organization.

List of Scientific Contributions

Peer-reviewed journal articles

Gaidai, R., **Goelz, C.**, Mora, K., Rudisch, J., Reuter, E.-M., Godde, B., Reinsberger, C., Voelcker-Rehage, C. & Vieluf, S. Classification characteristics of fine motor experts based on electroencephalographic and force tracking data. *Brain Research* **1792**, 148001 (2022)^{*1}

Goelz, C., Mora, K., Rudisch, J., Gaidai, R., Reuter, E.-M., Godde, B., Reinsberger, C., Voelcker-Rehage, C. & Vieluf, S. Classification of visuomotor tasks based on electroencephalographic data depends on age-related differences in brain activity patterns. *Neural Networks* **142**, 363–374 (2021)¹

Goelz, C., Mora, K., Stroehlein, J., Haase, F., Dellnitz, M., Reinsberger, C. & Vieluf, S. Electrophysiological signatures of dedifferentiation differ between fit and less fit older adults. *Cognitive Neurodynamics* **15**, 1–13 (2021)¹

Goelz, C., Reuter, E.-M., Fröhlich, S., Rudisch, J., Godde, B., Vieluf, S. & Voelcker-Rehage, C. Classification of age groups and task conditions provides additional evidence for differences in electrophysiological correlates of inhibitory control across the lifespan. *Brain Informatics* **10**, 11 (2023)¹

Gowik, J. K., **Goelz, C.**, Vieluf, S., van den Bongard, F. & Reinsberger, C. Source connectivity patterns in the default mode network differ between elderly golf-novices and non-golfers. *Scientific Reports* **13**, 6215 (2023)

^{*}Shared first authorship between R. Gaidai and C. Goelz

¹Publication considered for thesis

Gölz, C., Voelcker-Rehage, C., Mora, K., Reuter, E.-M., Godde, B., Dellnitz, M., Reinsberger, C. & Vieluf, S. Improved Neural Control of Movements Manifests in Expertise-Related Differences in Force Output and Brain Network Dynamics. *Frontiers in Physiology* **9** (2018)

Niess, A., Widmann, M., Gaidai, R., **Gölz, C.**, Schubert, I., Castillo, K., Sachs, J. P., Bizjak, D., Vollrath, S., Wimbauer, F., Vogel, A., Keller, K., Burgstahler, C., Quermann, A., Kerling, A., Schneider, G., Zacher, J., Diebold, K., Grummt, M., Beckendorf, C., Buitenhuis, J., Egger, F., Venhorst, a., Morath, O., Barsch, F., Mellwig, K.-P., Oesterschlink, J., Wüstenfeld, J., Predel, H.-G., Deibert, P., Friedmann-Bette, B., Mayer, F., Hirschmüller, A., Halle, M., Steinacker, J. M., Wolfarth, B., Meyer, T., Böttinger, E., Flechtner-Mors, M., Bloch, W., Haller, B., Roecker, K. & Reinsberger, C. COVID-19 in German Competitive Sports: Protocol for a Prospective Multicenter Cohort Study (CoSmo-S). *International Journal of Public Health* **67** (2022)

Stroehlein, J. K., Vieluf, S., Zimmer, P., Schenk, A., Oberste, M., **Goelz, C.**, van den Bongard, F. & Reinsberger, C. Learning to play golf for elderly people with subjective memory complaints: feasibility of a single-blinded randomized pilot trial. *BMC Neurology* **21**, 200 (2021)

Strote, C., **Gölz, C.**, Stroehlein, J. K., Haase, F. K., Koester, D., Reinsberger, C. & Vieluf, S. Effects of force level and task difficulty on force control performance in elderly people. *Experimental Brain Research* **238**, 2179–2188 (2020)

Ströhlein, J. K., Vieluf, S., van den Bongard, Franziska, **Gölz, C.** & Reinsberger, C. Golf spielen gegen die Vergesslichkeit: Effekte des Erlernens der Sportart auf das Default Mode Netzwerk des Gehirns. *B&G Bewegungstherapie und Gesundheitssport* **36**, 65–72 (2020)

Vieluf, S., Mora, K., **Gölz, C.**, Reuter, E.-M., Godde, B., Dellnitz, M., Reinsberger, C. & Voelcker-Rehage, C. Age-and expertise-related differences of sensorimotor network dynamics during force control. *Neuroscience* **388**, 203–213 (2018)

Contributions to Conferences

Talks²

Gaidai, R., **Gölz, C.**, Widmann, M., Niess, A. M. & Reinsberger, C. Herausforderungen und Lösungen im Aufbau eines Datenbanksystems im Verbundprojekt „Covid-19 im Spitzensport - Eine multizentrische Kohortenstudie“ (CoSmo-S). Sports Medicine and Health Summit. Hamburg, Germany, Sept. 2023

²First author was always the presenting author

Goelz, C., Gaidai, R., Mora, K., Rudisch, J., Reuter, E. M., Godde, B., Reinsberger, C., Voelcker-Rehage, C. & Vieluf, S. *Classification characteristics of fine motor experts based on electroencephalographic and force tracking data.* 54th Annual Conference of the German Society for Sport Psychology. Muenster, Germany, June 2022

Ströhlein, J., **Gölz, C.**, Vieluf, S., van den Bongard, F. & Reinsberger, C. *Funktionelle Netzwerkcharakteristika des Default Mode Netzwerks bei älteren Golf-Novizen.* Sports Medicine and Health Summit. Virtual Conference, Apr. 2021

Vieluf, S., Reinsberger, C., **Gölz, C.** & Voelcker-Rehage, C. *Age- and expertise related differences of sensorimotor network dynamics during force control.* 23rd University Day of the German Society of Sport Science. Munich, Germany, Sept. 2017

Poster²

Gaidai, R., **Gölz, C.**, Widmann, M., Niess, A. M. & Reinsberger, C. *Herausforderungen und Lösungen im Aufbau eines Datenbanksystems im Verbundprojekt „Covid-19 im Spitzensport - Eine multizentrische Kohortenstudie“ (CoSmo-S).* Nachwuchssymposium Deutsche Gesellschaft für Sportmedizin und Prävention. Tuebingen, Germany, Sept. 2022

Goelz, C., Mora, K., Stroehlein, J., Reinsberger, C. & Vieluf, S. *Electrophysiological signatures of brain network dynamics in elderly.* 26th Annual Meeting of the Organization of Human Brain Mapping. Virtual Conference, June 2020

Vieluf, S., Voelcker-Rehage, C., **Goelz, C.**, Mora, K. & Reinsberger, C. *Age- and expertise-related differences in force control manifest in variability.* 27th Annual Meeting of the Society of the Neural Control of Movements. Dublin, Ireland, May 2017

Chapter 1

General Introduction

1.1 Motivation

"Humans now live longer than at any time in history. But adding more years to life can be a mixed blessing if it is not accompanied by adding more life to years."

Dr. Tedros Adhanom Ghebreyesus, WHO Director-General,
2020

One of the major societal challenges facing Western societies is the demographic shift towards an older population, which will place enormous demands on society and raise questions about health care, infrastructure, family policy, and employment [19]. To avoid overburdening social structures, one of the main objectives is to promote healthy and independent aging and to improve the quality of life in old age. As part of efforts to promote these goals, the World Health Organization (WHO) launched the *Decade of Healthy Aging (2021-2030)*, which aims to encourage global action to improve the lives of older adults, their families, and the communities in which they live with the ultimate goal of *adding life to years* [19].

An essential part of promoting healthy aging and enabling participation in society includes the early identification and treatment of pathological conditions, developing and evaluating targeted interventions for prevention and therapy, or designing assistive technologies for older adults. These efforts require a deep understanding of the dynamics of aging in the context of individual trajectories and general patterns. Since many of the mechanisms leading to cognitive and physical decline are related to changes in the brain, it is of great interest to understand and quantify the aging process at this level.

Aging is a highly complex phenomenon, and the brain itself, as a nonlinear, dynamic, and multi-layered system, exhibits complex properties in space and time [20]. Machine learning offers valuable data-driven methods to unravel this complexity and gain insights by uncovering complex relationships and identifying predictive markers related to the aging process and associated health status. In general, progress in science is more and more characterized by applying methods from artificial

intelligence (AI), including machine learning algorithms, which make it possible to systematically analyze large and complex amounts of data [21]. This development has led to proclamations of an "AI revolution in science" [22] or promoting science has entered a new area characterized by *data-intensive computing* [23]. Moreover, these methods serve as the foundations for solving various practical problems, as demonstrated by applications in many socially relevant areas, such as public transport, e.g., autonomous or self-driving vehicles [24], the medical sector, e.g., diagnostic imaging [25], or social interaction, e.g., tools for communicative interaction [26], and are thus one of the basic building blocks for assistive technology facilitating the participation of older people with disabilities in society. AI and machine learning as a key technology have become a hope for solving societal challenges, including the shift towards an older population.

However, the implementation of machine learning approaches in aging research is still at an early stage compared to the rapid development in the commercial sector, and the most effective applications and integration into the traditional scientific system have yet to be evaluated, despite the potential to better understand the aging brain.

This is the starting point of this dissertation which aims to investigate brain aging using machine learning techniques. The focus is on using these methods to better understand the neurophysiological factors contributing to age-related sensory, motor, and cognitive changes. To this end, existing hypotheses about the aging brain will be tested and validated while new hypotheses will be generated. The results inform the development of assistive technologies to facilitate the participation of older adults in society, the early detection of pathological conditions, or the development of targeted interventions to counteract age-related decline.

1.2 Outline

This dissertation is separated into five main chapters. This chapter describes the thesis's theoretical framework. In Chapter 1.3, a description of aging at the level of the brain focuses on the most relevant concepts for the context of this work and forms the starting point for introducing the added value of applying machine learning in the context of studying the reorganization of the aging brain. Next, machine learning is introduced in Chapter 1.4 to provide the methodological framework. The general terminology and a literature-based overview of the use of machine learning methods in neuroscience and especially in the neuroscientific research on aging will form the basis for the deduction of the research aim and scope of this thesis in the following Chapter 2. The following Chapter 3 includes a description of the general methodological approaches of this work. In the subsequent Chapter 4, the main results of the published research articles underlying this thesis will be presented. These include:

- Research Article I:

Goelz, C. *et al.* Classification of visuomotor tasks based on electroencephalographic data depends on age-related differences in brain activity patterns. *Neural Networks* **142**, 363–374 (2021)

- Research Article II:

Goelz, C. *et al.* Classification of age groups and task conditions provides additional evidence for differences in electrophysiological correlates of inhibitory control across the lifespan. *Brain Informatics* **10**, 11 (2023)

- Research Article III:

Goelz, C. *et al.* Electrophysiological signatures of dedifferentiation differ between fit and less fit older adults. *Cognitive Neurodynamics* **15**, 1–13 (2021)

- Research Article IV:

Gaidai, R., Goelz, C., *et al.* Classification characteristics of fine motor experts based on electroencephalographic and force tracking data. *Brain Research* **1792**, 148001 (2022)*

The thesis concludes with an overarching discussion in which the results are evaluated in the light of the current scientific discourse highlighting practical implications and future research topics (chapter 5).

1.3 Aging

Biologically aging is "the time-dependent functional decline that affects most living organisms" [27]. It can be observed in the reorganization of multiple interacting physiological systems operating at different spatial and temporal scales [28]. The underlying patterns of reorganization within and between these systems are highly individual, as they are subject to internal (e.g., genetic, cellular, molecular) as well as external (e.g., environmental, and lifestyle) influences [28–30]. At the same time, however, overarching, generalizable patterns can be identified [31]. The most recognizable consequences of aging are changes in cognitive, sensory, and motor abilities that challenge the daily lives of older adults [32]. However, not all abilities are equally affected by declines, and the changes are highly individual. While sensory, motor, and cognitive abilities, such as memory and processing speed, are generally declining, abilities in the context of acquired knowledge, such as verbal abilities, tend to be stable or even improve with age [33]. One factor that plays a crucial role in these changes is reorganization at the level of the brain [34]. A profound understanding is, therefore, of particular interest to research efforts as this is a prerequisite to identifying unfavorable trajectories and developing prevention and therapy concepts. It is important to note that the reorganization of the brain can be viewed from many perspectives, so in the following, only the aspects and concepts essential for the understanding of this work will be presented.

1.3.1 Age-related Reorganization of the Brain

Reorganization in the brain's structure includes, among others, atrophy of the gray and white matter and enlargement of cerebral ventricles [35]. The efficiency of neuromodulation declines mainly driven by the loss of dopaminergic receptors indicative of a reorganization of neurotransmitter

systems [36]. Besides this, the study of the functional properties of the brain and their relationship to behavioral changes is of great interest. In neuroimaging studies, both under-activation and over-activation of brain areas have been reported in older adults compared to younger adults during the performance in various tasks with sensory, cognitive as well as motor demands [34, 37]. Regarding activation dynamics, brain activity in response to a stimulus is often slower or delayed. Moreover, the frequency distribution of oscillatory neural activity changes to a slowing of the primary rhythms and altered temporal dynamics, which is interpreted as changes in neural communication [38].

By emphasizing neural communication and information flow, rather than viewing the brain as functionally separate, it can be conceptualized as a complex system whose functional units, i.e., neurons, areas, and subsystems, are interconnected structurally and functionally [39, 40]. In this concept, functional connectivity reflects coherent activation patterns within and between these units. Several distinct but interconnected functional networks have been identified [41]. The dynamic interplay between and within these networks is characterized by segregation and integration at different levels, indicating the flow of information in the brain [42]. Older adults' information flow tends to be less efficient and is characterized by lower within-network connectivity and higher between-network connectivity associated with a less segregated, less modular, and more integrated brain network organization [37, 40, 43]. However, studies on sensorimotor and visual networks seem very heterogeneous, which could indicate individual reorganization patterns [40].

Dedifferentiation

The functional reorganization patterns described in the previous section have been attributed to dedifferentiation [44]. Dedifferentiation refers to the loss of neural specialization or reduced distinctiveness of neural responses resulting in diffuse, nonspecific recruitment of brain resources [45]. Historically, the term originates from behavioral research in which an increased correlation of performance between sensory, cognitive, and sensorimotor domains was reported in older adults [32, 46]. To explain this behavioral dedifferentiation, Li *et al.* [36, 47] provided a computational model. According to this model, deficient neurotransmitter modulation observed in older adults may affect the responsiveness of cortical neurons, leading to higher levels of neuronal noise and ultimately to less differentiated, more diffuse neuronal activation patterns in response to different stimuli [36, 47] (see Figure 1.1 for an overview on the computational model). In several computational simulations, the authors demonstrated that the proposed model could explain behavioral co-variation and several other phenomena, such as decreased average behavioral performance or increased behavioral intra- and inter-person variability [32, 47]. Furthermore, decreased stimulus selectivity of cells, for example, in the primary visual, auditory, or somatosensory cortices with increasing age, has been confirmed in animal models at the single-cell level [45, 48–50]. Neuroimaging studies in humans have also demonstrated that the activation patterns to various visual, cognitive, and motor stimuli are less specific in older compared to young adults [45, 51, 52]. Importantly, dedifferentiation was also attributed to functional networks, as considered in the computational model presented in Fig-

ure 1.1. As such, the aforementioned reorganization of functional networks, characterized by less segmented and modular and less specialized network organization in older adults, has also been referred to as dedifferentiation [37, 40, 45]. In addition, the activation of functional networks during task performance was found to be less specific [53–55]. Additionally, research has indicated that the degree of dedifferentiation might predict behavioral performance in cognitive tasks [45]. Based on this, it can be assumed that age-related decline is at least partially due to the dedifferentiation of cortical representations [45].

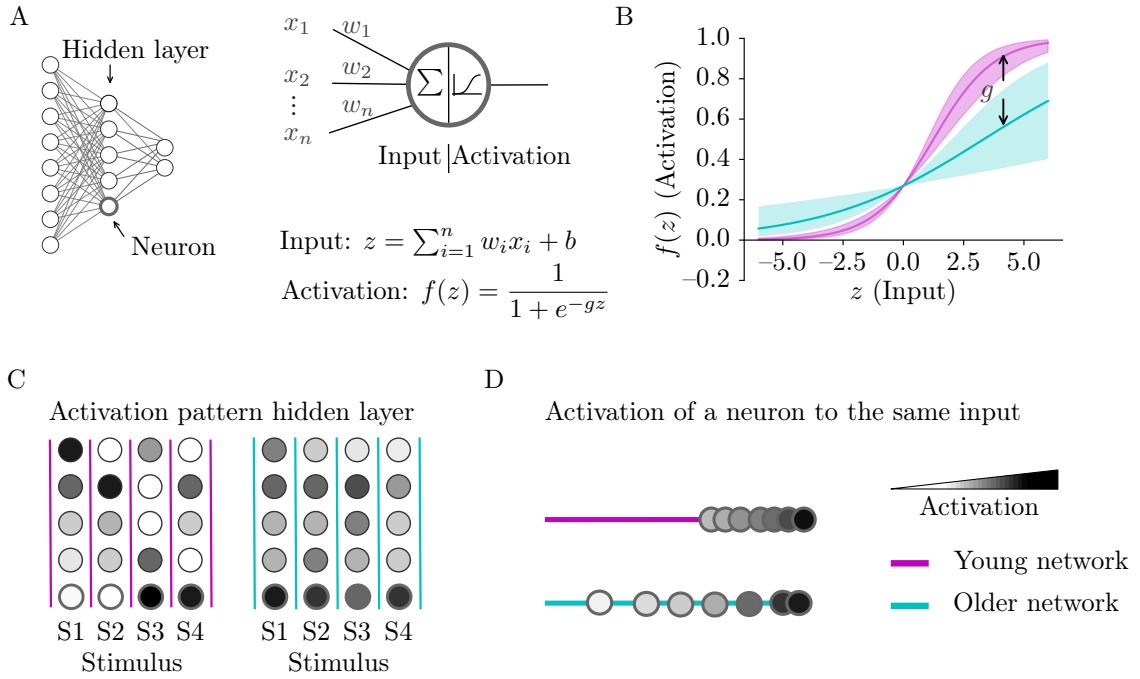


Figure 1.1: The computational model of dedifferentiation. Li and colleagues [32, 36] used a feedforward backpropagation neural network model with logistic activation function $f(z)$ and simulated altered neuromodulation by varying the gain parameter g in $f(z)$ of each neuron (A). Lower g values represent deficient neuromodulation and responsiveness due to aging, resulting in a damped neuron activation (B). Simulations showed that the activation pattern of simulated neurons differs less for different stimuli, i.e., the network’s hidden layer shows a less distinctive representation of the stimulus (C). The activation of a single neuron is more variable in networks with lower g value, i.e., older networks, for multiple stimulations with the same stimulus (D). Adapted from Li *et al.* [36] with permission.

Fornito *et al.* [56] describe dedifferentiation as a fundamental maladaptive mechanism of brain networks that requires compensation. This view is consistent with the argument that dedifferentiation and compensation are complementary mechanisms [34]. However, dedifferentiation could also represent a compensatory response, in that the brain attempts to maintain function in the face of deterioration [57]. By definition, compensation refers to the ability to recruit additional brain resources to compensate for decline and functional loss to maintain cognitive or behavioral functioning [34, 44]. Here, the compensation-related utilization of neural circuits hypothesis (CRUNCH) hypothesizes that compensatory activity changes as a function of task demands [58].

Moreover, compensation often occurs in a specific pattern of under-activation of posterior areas and prefrontal over-activation, known as posterior–anterior shift in aging (PASA) [59]. Another frequently reported pattern is the more bilateral recruitment and loss of hemispheric specialization, known as hemispheric asymmetry reduction in older adults (HAROLD) [60].

Reserve

It is important to note that age-related changes of the brain and behavior are highly individual and dynamic [29, 45, 61]. In this context, the reserve hypothesis defines *reserve* as the accumulated capacity of neural resources over the lifespan that can withstand decline or pathology [57, 62]. Although the concept was initially based on observations that the degree of pathological changes in the brain does not necessarily mean clinical manifestation, it has also been applied to explain the individuality of non-pathological aging [57, 62, 63].

Reserve can be both anatomically quantifiable, referred to as brain reserve, and more functional in nature, referred to as cognitive reserve [57]. At the functional level, compensatory activation, as well as more efficient utilization (less activation of neural resources), and increased capacity (increased availability of neural resources) were described as key mechanisms of cognitive reserve [57, 64]. Brain and cognitive reserve influence each other, and Cabeza *et al.* [62] argue against a strict separation of brain reserve and cognitive reserve. One aspect that explicitly determines the definition of reserve is the lifelong ability of the brain to adapt its structure and function to internal and external requirements. It is known from the animal model that environments rich in cognitive, social, sensory, and motor stimuli contribute to positive plastic changes [65]. As a result, reserve is influenced by an interplay between genetic and environmental factors, including lifestyle factors [62]. Essential elements for increasing reserve have been identified in education, occupation as well as physical activity, with cognitive training, physical fitness, and professional expertise having a considerable impact on the brain’s functional organization [11, 66, 67]. Since reserve is not directly measurable, proxies are often used to compare individuals with high or low scores for a particular proxy value (e.g., physical fitness, education, or professional expertise) [62]. Other complementary concepts, such as the maintenance or the scaffolding theory of cognitive aging (STAC) model, highlight these influencing factors additionally. The concept of maintenance emphasizes the ability of the brain to repair. STAC postulates that lifelong positive and negative plasticity defines a framework that enables compensation and shapes the individual trajectory of aging [68].

1.3.2 Studying Brain Aging by Electroencephalography

The complex interplay of the factors mentioned above leading to the dynamics of age-related reorganization of the brain is highly complex. Understanding these dynamics regarding individual trajectories and overarching patterns is a prerequisite to differentiating healthy from pathological changes and developing and verifying treatments and targeted interventions. This requires uncomplicated, easy-to-use, and cost-effective methods and novel analyses to quantify changes in brain

organization. Several noninvasive methods are available to study the brain's structure and function. Magnet resonance imaging (MRI) is the most widely used method in science to image the structure or, using functional magnetic resonance imaging (fMRI), the function of the brain, which is the dominant method in the study of the functional reorganization described in the previous sections [34]. However, its use in the public health system is mainly limited to cases with a clear indication, making early detection of unfavorable aging trajectories challenging. In addition, limited availability substantially restricts the development of preventive and rehabilitative interventions and therapies and excludes areas and sites with low levels of equipment and expertise. Here, EEG could represent a real added value since it is characterized by simple use, mobility, and relative cost-effectiveness. Although it has a lower spatial resolution than MRI based methods, EEG measures neuronal activity directly with a high temporal resolution which allows for the detection of age-related changes in the temporal dynamics of brain activity and networks, which could be of particular interest to understand age-related changes of the brain and their relation to behavior [38].

Excursus: A Brief Overview on Electroencephalography

EEG measures time-varying electrical fields on the surface of the head by using several sensors placed in a standardized position [69]. The measured signals reflect synchronously active populations of neurons. Electrical activity can only accumulate and be detected on the surface of the head if spatially similar neurons, aligned perpendicular to the surface, are synchronously activated. Based on the conductive properties of the brain, the signal can travel through the different layers to the surface due to volume and capacitive conduction. For this reason, and due to the orientation of neural cell assemblies, the signal in each sensor reflects a summed signal of different neuron patches. The signal expressions are in the range of a few micro-volts and are much lower than other biological and non-biological electrical generators, e.g., muscular activity or line noise, so the EEG signal is often affected by a low signal-to-noise ratio [70].

One of the EEG's most striking signal characteristics is the rhythmic voltage fluctuations that define the signal and are summarized under the term oscillation. Commonly, the EEG signal is analyzed based on the frequency composition of oscillatory activity in loosely defined frequency ranges, i.e., δ (<4 Hz), θ ($\sim 4\text{-}8$ Hz), α ($\sim 8\text{-}12$ Hz), β ($\sim 12\text{-}30$ Hz) and γ (>30 Hz), which have been demonstrated to be related to perceptual, cognitive, motor and emotional processes [70]. Furthermore, the analysis of frequency-dependent synchrony or functional connectivity in terms of statistical dependence of the signals, e.g., by coherence or the phase synchrony of the signal, can provide information about the network characteristics of the brain [71]. Finally, the analysis of event-related activation, so-called event related potentials (ERPs), can provide information on the direct processing of stimuli. The analysis of ERPs involves time-locking the EEG data to the onset of a specific stimulus and averaging

the EEG signal across hundreds of trials to extract a reliable signal related to the processing of the stimulus.

Electroencephalographic Signatures of Age-related Reorganization

Age-related EEG characteristics have been extensively studied. Specifically, it has been reported that aging is associated with changes in the frequency composition of the EEG signal, regardless of any specific task involvement. These changes include a decrease in amplitude within the α frequency band, a shift in the α peak frequency towards lower frequencies, an increase in amplitude within the β frequency band, and varying results regarding changes in the amplitude of the θ and δ bands [38, 72, 73]. Moreover, age-related changes have also been reported in terms of reduced EEG synchrony and a more random, less segregated organization of EEG derived network topology [74, 75]. Altogether, these changes are believed to reflect transformations in brain function and connectivity associated with healthy aging, and variations in these have shown potential utility in diagnosing pathological conditions such as Alzheimer's disease [76]. In contrast, assessing preclinical or mild stages such as mild cognitive impairment poses additional challenges, and research have proposed the potential benefits of incorporating task-related EEG measures for a more effective evaluation [77, 78]. However, this requires a deep understanding of the changes in task-related information processing and reorganization in healthy aging.

Age-related EEG changes in relation to tasks are highly dependent on the task context or domain studied. For example, unilateral motor tasks may display lower frequency specificity and more bilateral spatial expression of α and β frequency power modulations [79]. In contrast, attention tasks may demonstrate enhanced frontal network involvement and power in the θ frequency band [80]. In addition, the neural response to stimuli may exhibit a temporal slowing and altered spatial expression. These changes can be seen, for example, in a delay of early ERP components as well as a more frontal expression of later ERP components in visual attention tasks [81, 82].

Often these characteristics are discussed in terms of dedifferentiation and compensation described above and have been shown to be modulated by lifetime experience such as occupational expertise [11] or physical fitness [61]. However, the relationship between EEG parameters and these mechanisms is often ambiguous. Other EEG findings, for example, point in the opposite direction than described above. Hübner *et al.* [83], for instance, found no age effects in central lateralization in the β frequency band in a complex fine motor control task, which again highlights the dependency on the task context considered. Age-related changes in decreased ERP latency and lower or increased functional connectivity of the examined networks depending on the task context are also reported [38]. Moreover, the interpretation of dedifferentiation is often based on fMRI findings that report over-activation and loss of segregation of brain networks, although the relationship between frequency-specific EEG and fMRI findings acting on different spatial and temporal scales and measurement principles might be unclear. Koen & Rugg [45] further points out that over-activation should be interpreted cautiously and does not necessarily imply loss of

neural specificity, as predicted in the original model of Li *et al.* [36, 47]. The authors, therefore, propose to operationalize dedifferentiation clearly in terms of the selectivity of the neural response between two or more task modulations. While in this operationalization, the evidence regarding dedifferentiation in fMRI studies is quite clear, this has not been explored in EEG studies so far [45].

Altogether the EEG represents an easy-to-use, low-cost method that can provide valuable insights into age-related changes. However, the link to age-related changes reported consistently in the fMRI, such as dedifferentiation, is often challenging and needs to be clarified. EEG signals are temporally and spatially highly dimensional, i.e., large amounts of data points contain intricate patterns of electrical activity. Additionally, the signals often have a low signal-to-noise ratio, making it difficult to detect and visualize age-related brain reorganization and its dynamics. As such, analysis of EEG signals requires advanced methods. In this context, methods from the field of machine learning could be of particular interest. By leveraging machine learning techniques, it is possible to extract meaningful patterns from the high-dimensional EEG data and uncover subtle age-related changes that may not be evident through traditional analysis methods.

1.4 Machine learning

Machine learning emerged in the 1950s to enable computers to learn without being explicitly programmed [84]. It is defined by computational methods combining fundamental concepts from computer science, statistics, probability, and optimization that automatically extract patterns and trends, i.e., *learn* from data [85]. The notion of *learning* therein describes the automated inference of general rules based on the observation of examples using algorithms to solve a specific task or problem [86]. In its basic form, these tasks often involve making predictions based on learned relationships or extracting information based on automatically detected patterns and structures from data. Many real-world problems can be formulated by these tasks, and a rise in machine learning started in the 1990s to 2000s with the availability of computing resources, data, and the development of algorithms, which have found their way into everyday life not only since the current advancements in generative AI systems. Examples can be found in numerous areas, such as predicting stock prices, personalized advertising, or autonomous driving [87].

In science, machine learning is increasingly used as a complementary method to classical statistical analyses because of the ability to make predictions and deal with the multidimensional structure and non-linearity in real-world datasets for drawing inference [88]. Especially in areas where high-dimensional data is prevalent, such as in neuroscience, machine learning methods offer insight by extracting complex patterns purely data-driven [21]. In terms of EEG, machine learning can help identify subtle patterns and nonlinear relationships from the complex multidimensional structure of the data, allowing for more accurate and efficient analysis of brain recordings. Various methods are available for this purpose, which can be roughly characterized based on various properties.

1.4.1 Forms of machine learning

The three main forms of machine learning are supervised, unsupervised, and reinforcement learning. These forms are defined by the type of feedback a machine learning algorithm has access to during learning [89].

Supervised machine learning aims to learn a generalizable relationship between data and associated information, so-called labels or targets. The learned model can then be used to predict the label of new data not used during the learning process. If the labels are categorical, the prediction task is called classification; for continuous labels, the term is regression. Unsupervised machine learning aims to find hidden structures in data without considering associated labels. This could be grouping similar data points, i.e., clustering, or uncovering a meaningful low dimensional representation of high dimensional data, i.e., dimensionality reduction. This type of learning is also referred to as *knowledge discovery* [90]. Reinforcement learning describes the task of learning optimal actions to solve a particular problem by maximizing the reward linked to that action. See Figure 1.2 for an overview.

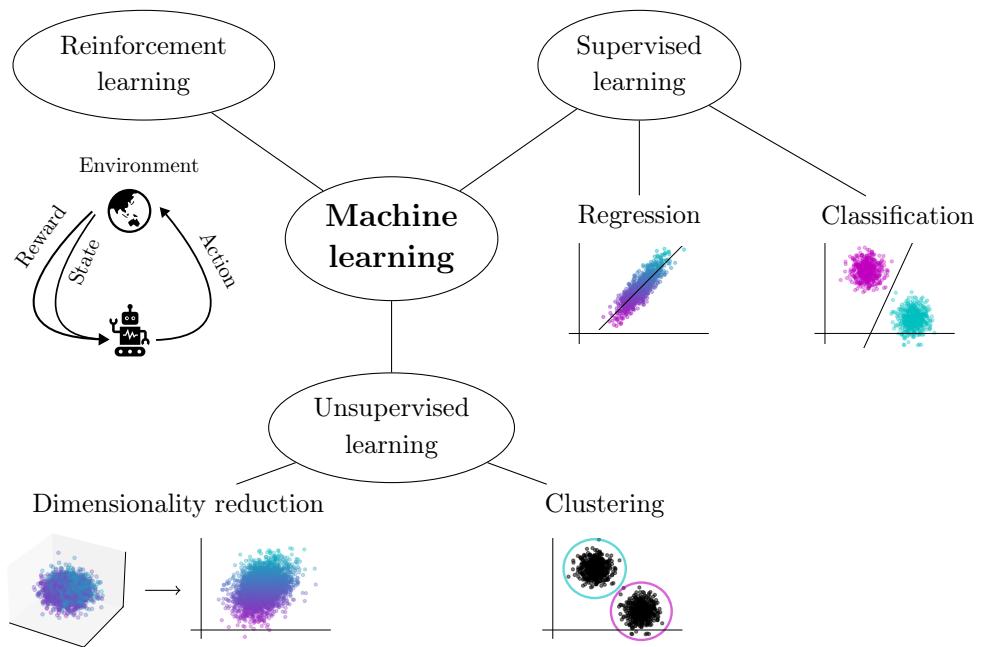


Figure 1.2: The three main forms of machine learning.

In practice, however, a clear separation is often impossible. As such, dimensionality reduction can also be supervised, i.e., labels are provided to learn a new representation of the data [91]. Besides, in semi-supervised learning, the goal is the same as in supervised learning. However, the data set used to learn the relationship contains labeled and unlabeled examples. The hope is to build a stronger representation by providing more information in the form of data [92].

In addition, traditional machine learning is often contrasted with deep learning methods involving

artificial neural networks, which are composed of many layers of interconnected nodes often used in an end-to-end fashion in which the input data is used without any form of preprocessing. Usually, they require a vast amount of data and computational power. In the context of this thesis, the tasks considered involve the processing of EEG from experiments with mid to small sample sizes to learn meaningful patterns and relationships in data. The following sections present state-of-the-art approaches for applications on EEG data.

Excursus: How does a machine learn?

"A computer program is said to learn from experience E with respect to some task T and some performance measure P , if its performance on T , as measured by P , improves with experience E " [93]. In other words, learning in the context of machine learning typically involves solving a specific task by using algorithms that improve their performance by using example data. There are numerous algorithms designed to solve the problems outlined above. Some basic building blocks can be defined, which can be used to describe computational learning formally. In the following description, the view of statistical learning theory is considered, and notation is adapted from Shalev-Shwartz & Ben-David [89] and from von Luxburg & Schölkopf [86].

Learning is always based on data, i.e., measurable information about some phenomenon, consisting of attributes of the phenomenon, so-called features, and an associated label in supervised learning. It is mathematically defined as an open bounded set $\mathcal{Z} \subset \mathbb{R}^n$ of dimension n . Typically there is only a set of examples or training data $S = \{z_i, \dots, z_m\} \subset \mathcal{Z}^m$ available, where $i = 1, \dots, m$, and each z_i is sampled independently from \mathcal{Z} according to an underlying probability distribution \mathcal{D} . Thus the only assumption is that the example data are independent and identically distributed. No assumption on \mathcal{D} is made.

In supervised learning, \mathcal{Z} comprises the space of input data \mathcal{X} and the space of labels or output \mathcal{Y} . The example data S consists of labeled input-output pairs $z_i = x_i, y_i \in (\mathcal{X} \times \mathcal{Y})^m$, where x_i is an input data vector and y_i is its corresponding output label. The pairs are sampled by some unknown joint probability distribution \mathcal{D} on the space $\mathcal{X} \times \mathcal{Y}$.

The space \mathcal{Z} in unsupervised learning comprises the input data space \mathcal{X} only and the example set S consists of unlabelled examples $z_i = x_i \in \mathcal{X}^m$, sampled according to some unknown probability distribution \mathcal{D} on the space \mathcal{X} .

Learning ultimately can be thought of as approximating an underlying ground truth function f , also called model, that represents the relationship between input and output in supervised learning, i.e.,

$$f : \mathcal{X} \rightarrow \mathcal{Y}, \quad (1.1)$$

or the mapping to a space of hidden patterns or structure $\mathcal{W} \subset \mathbb{R}^p$, where p can be equal or smaller than n , i.e.,

$$f : \mathcal{X} \rightarrow \mathcal{W}. \quad (1.2)$$

A learning task can be conceptualized as searching through the space of all possible solution functions. As this is not feasible, a finite class of functions, so-called hypotheses, is typically selected a priori. Thus, learning can be thought of as selecting a hypothesis h from a space of potential solutions \mathcal{H} with $\mathcal{H} = \{h : \mathcal{X} \rightarrow \mathcal{Y}\}$ in supervised learning and $\mathcal{H} = \{h : \mathcal{X} \rightarrow \mathcal{W}\}$ in unsupervised learning.

A learner or learning algorithm is the means of selecting the best element from \mathcal{H} . The cost of a false prediction or an inaccurate representation of the data is quantified using a loss function, $\ell : \mathcal{H} \times \mathcal{Z} \rightarrow \mathbb{R}_+$. In other words, it measures how well a specific hypothesis is doing.

The expected risk is a measure of the average loss of a hypothesis, $h \in \mathcal{H}$ with respect to the probability distribution \mathcal{D} over \mathcal{Z} and can be defined as

$$L_D(h) := \mathbb{E}_{z \sim D}[\ell(h, z)] \quad (1.3)$$

A learner should select a hypothesis with the lowest possible expected risk. However, the underlying probability distribution is unknown. Using S , the expected risk can be estimated using the empirical risk over the training data. This is defined by:

$$L_S(h) := \frac{1}{m} \sum_{i=1}^m \ell(h, z_i). \quad (1.4)$$

Following this, learning can be formalized as solving an optimization problem of the form:

$$\hat{h} = \arg \min_{h \in \mathcal{H}} L_S(h), \quad (1.5)$$

which can be solved computationally. In parameterized models, this often involves the automated selection of those parameters $\theta \in \Theta$ of a chosen class of models that minimize $L_S(h_\theta)$. This optimization problem can then be solved by methods such as gradient descent or, e.g., analytically, using least squares estimation. The solution \hat{h} is the learned model that can be used to solve the task at hand, e.g., predicting the label of new input data or uncovering patterns or structures in data. This is known as empirical risk minimization (ERM).

Upon ERM, more complex learning paradigms can be used to address common problems such as overfitting, in which the learned hypothesis too closely relies on the training data and therefore has low generalization performance, e.g., regularized risk minimization, which introduces regularization to ERM or structural risk minimization that penalizes complex models and encourages simplicity.

Although most machine learning can be conceptualized within the framework of

ERM, there are models that, instead of minimizing risk, assume that the underlying distribution over the data has a specific parametric form, and the goal is to estimate these parameters by using maximum likelihood estimation (MLE) which seeks to find the model parameters that maximize the likelihood of the observed data under the assumed parametric distribution, i.e.,

$$\hat{\theta}_{\text{MLE}} = \arg \max_{\theta \in \Theta} \prod_{i=1}^m p_\theta(z_i), \quad (1.6)$$

where $p_\theta(z)$ is the joint probability function of the assumed parametric distribution and $\hat{\theta}_{\text{MLE}}$ is the estimated value of the parameter vector θ .

1.4.2 State-of-the-Art Approaches to Electroencephalographic Data

Various established supervised and unsupervised algorithms have been utilized in the analysis of EEG data, and the selection is usually based on the goal of the analysis. Unsupervised learning aims to highlight specific information in the data, so the selection is made based on the information one aims to highlight [89]. This is to highlight group structure in EEG data when using clustering or to highlight EEG inherent characteristics in dimensionality reduction. In contrast, selecting a suitable supervised learning algorithm is more guided by its performance, i.e., its ability to derive generalizable rules that allow predictions from the available data. Typically, different classification algorithms and their parameters are selected, trained on one portion of the data, the so-called training data, and then tested for their performance on data not used for training, the so-called testing data [94]. The training data can further be divided into a training and validation portion to compare different model types or user-defined learning algorithm settings, so-called hyperparameters. Finally, the best model configuration is tested for its predictive performance on the test data and reported. However, this three-time division may drastically reduce the data size usable for training and may result in flawed generalization evaluation due to the randomness of the split [95]. Therefore several procedures can be applied. In simple k-fold cross-validation, for example, the training data is divided k-times. Thus each time, a different subset of the data is used for validation while the rest is used for training. Usually, this is repeated for a range of models and subsequent hyperparameters, and the model and hyperparameter performing best on average are selected for final testing. Building on this, nested cross-validation can be used to select the best model and test the generalization performance by adding a second cross-validation loop for the final model evaluation. In this way, an unbiased estimation of the generalization performance of a model can be obtained (see Figure 1.3 for a visual representation of the procedure).

Recent work highlights deep neural networks that can be used for unsupervised and supervised machine learning applications to EEG [96]. However, their advantage comes into play with large data resources, which are often expensive to acquire in the case of EEG [97]. Traditional learning approaches can be more efficient with good performance and promise better interpretability, espe-

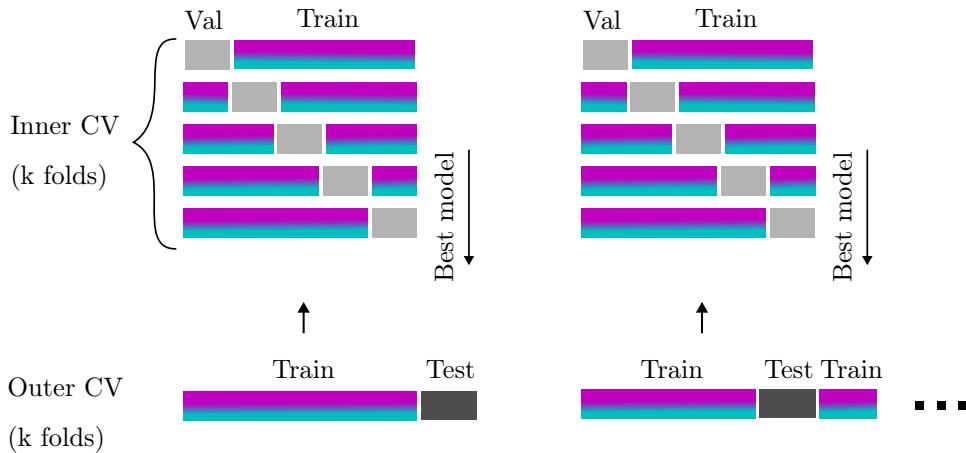


Figure 1.3: Exemplary nested cross-validation procedure. K-fold cross-validation is used in an outer loop for testing the best configuration tuned in an inner cross-validation loop. CV: cross-validation, Val: Validation

cially for comparatively smaller datasets and limited computational resources [98]. Due to the low signal-to-noise ratio and high complexity of EEG data, the inputs in these approaches are often represented by well-known EEG characteristics or features that are believed to be related to the problem being learned. Typical features include time, frequency, time-frequency, connectivity, and theoretical information parameters extracted for each sensor (see Gemein *et al.* [98] for common choices). However, this approach may lead to less flexible and generalizable models with low spatial resolution and vulnerability to low signal-to-noise ratios [99].

Some approaches to address these problems compute the anatomical sources of the EEG signals in the brain using biophysical models as a preprocessing step prior to feature extraction [100, 101]. However, they require a head model based on MRI often unavailable individually or merely estimated based on existing templates. Other approaches use supervised and unsupervised decomposition techniques belonging to the field of dimensionality reduction as a preprocessing step for further prediction tasks or provide information themselves in the sense of knowledge discovery. These methods aim at *unmixing* the highly correlated sensor time series by assumptions about the underlying signal components. For example, independent component analysis (ICA) assumes statistical independence. In contrast, principal component analysis (PCA) assumes that the extracted components are maximally uncorrelated to each other, capturing the largest amount of variance in the data [70]. Dynamic mode decomposition (DMD) is a method that explicitly considers the temporal structure of the signals, which requires that the extracted signal patterns (modes) are dynamically coherent, extracting spatiotemporal coherent structures and thus accounting for the network nature of the brain [102]. Additionally, supervised methods such as common spatial patterns (CSP) [103] or xDAWN [104] extract signal components that correlate with the labels to be predicted.

While the supervised and unsupervised dimensionality reduction methods mentioned so far offer ways of examining the complex EEG signals in terms of components and patterns to generate knowledge, non-linear methods such as t-distributed stochastic neighbor embedding (t-SNE) and uniform manifold approximation and projection for dimension reduction (UMAP) take into account the non-linear relationships between the data points and provide a lower-dimensional representation of the data that is often easier to interpret and visualize [91]. These methods can be beneficial for exploring the relationships between different EEG features or identifying subgroups within a dataset.

It is important to note that these methods can be applied not only to the EEG signals themselves but also to previously extracted EEG parameters or in combination in terms of knowledge discovery. Thus, supervised and unsupervised dimensionality reduction provides data-driven insights into the complex underlying information but also serves as preprocessing for further tasks such as prediction.

1.4.3 Applications in the Context of Aging Research

Traditionally, the previously presented machine learning approaches have been the core building block for developing intelligent systems that can automate tasks or enhance and assist humans in performing their tasks. Such systems are critical in terms of assistive technology, for example, to support older adults with disabilities to live their daily lives, but are also relevant in the medical field. In the latter, the hope is to develop intelligent medical systems to inform clinical theory and support clinical decision-making, i.e., assist in diagnosis and risk management by predicting health status or forecasting treatment responses [105]. In this context, supervised learning is often used to identify markers from EEG by identifying signal features that are predictive of a particular disease or health condition, which is highly important in promoting a healthy aging trajectory [76, 106]. An application, known as brain age estimation, is to estimate biological age based on a regression model trained on neural data, e.g., EEG data, recorded in extensive population studies [107]. The model can then be used to predict the age of an individual. If the brain appears older than it would chronologically, i.e. if the gap between predicted and actual age is large, this may be an early indication of an unfavorable state of health [108].

Another highly relevant application in the context of aging is the development of devices to assist, augment or enhance humans' capabilities, such as brain computer interfaces (BCIs). In BCIs, neural activity is decoded, using classification to generate control commands for various external devices such as computers or prosthetic limbs [109, 110]. Decoding refers to learning a classification or regression model that predicts behavioral outcomes or cognitive states based on neural data. Beyond the application in BCIs, decoding techniques are widely used in neuroscientific research to gain insights into the neural mechanisms underlying perception, cognition, and behavior. This type of analysis is often referred to as multivariate pattern analysis (MVPA) because its goal is to detect multivariate patterns, e.g., a set of voxels in fMRI or an electrical pattern at a given time point in EEG, associated with an experimental condition [111]. While the use has a long history

in the field of fMRI analysis, it has only become more widespread in the field of EEG in recent years. Therefore, decoding approaches to understanding age-related reorganization are mostly limited to fMRI studies. A common approach is to measure dedifferentiation at the individual level, i.e., the loss of neural specificity. Since dedifferentiation, by definition, results in more similar brain activation patterns for different tasks or stimuli, a poorer performance of classifiers trained to discriminate between them based on neural recordings is indicative of a less distinctive neural representation [45, 112]. However, the literature on the application of this approach to EEG data is minimal and restricted to single studies [113].

Classifying group membership or group-level regression can provide additional information about interesting relationships and their generalizability at the group level. Particularly for EEG markers representing functional network characteristics can reveal insightful findings about the relationship to age-related changes [114].

In addition to typical statistically motivated analysis methods that calculate bivariate connectivity between sensors based on the phase difference or coherence of the EEG signals, dimensionality reduction techniques, such as the aforementioned DMD, provide a data-driven way to capture dynamic brain network characteristics. This approach has already been used to map age- or expertise-related changes related to brain networks [11]. Further, unsupervised methods, such as nonlinear dimensionality reduction techniques, were frequently used to describe the structure of data sets with respect to age-related changes. [97, 115].

In summary, machine learning is very diverse and ranges from engineering applications to scientific knowledge discovery. Especially in the latter case, it offers the advantage of automated extraction of patterns from highly complex data that can contribute to studying age-related changes. While decoding approaches are particularly interesting for measuring age-related changes in the organization of neural systems, such as the level of differentiation, group analysis could provide new insights into datasets. Especially classification methods that predict a particular experimental condition or a group membership are particularly suitable. The combination with unsupervised learning algorithms, such as dimensionality reduction methods, could be particularly beneficial and used to visualize high-dimensional data.

Chapter 2

Aims and Scope

The main goal of this dissertation is to study age-related brain reorganization, considering both global patterns and individual trajectories, by applying established methods from supervised and unsupervised machine learning to EEG signals.

A prominent aspect of age-related reorganization discussed in the literature at the individual level is the loss of specificity of neural representations or dedifferentiation that accompanies the aging process and plays an essential role in the behavioral decline. This aspect is based on a computational model and has been confirmed in animal and human studies, as presented in Chapter 1.3.1 showing more similar activation patterns of brain areas or networks during task execution [52–55]. According to Koen & Rugg, considering fMRI studies, evidence for that is quite robust when dedifferentiation is operationalized based on the original model of Li *et al.* [32, 36] as loss of neuronal selectivity for different stimuli. In other words, the assumption is that there is less difference in brain states between task conditions. In contrast, results based on EEG studies are often ambiguous as markers, and clear operationalization of dedifferentiation is often missing. Applying machine learning techniques could add value in testing the discriminability of brain states at the individual level. Although this is already applied studies utilizing fMRI, as presented in Chapter 1.4.3, it has only been applied infrequently to EEG data.

Consequently, the first approach of this work is to use machine learning to test the discriminability of task-related EEG signals and draw conclusions about dedifferentiation. Compared with studies using fMRI, this would have the advantage of more directly capturing age-related reorganization and its dynamics at the individual level while offering several advantages in terms of practical availability, low cost, and ease of use.

As postulated by the reserve hypothesis, lifestyle factors contribute significantly to the development of individual age trajectories, and it can be assumed that individuals with a higher degree of reserve exhibit less dedifferentiation based on adaptive and flexible resources. This is based on comparative studies in which individuals with a low expression in a proxy parameter are contrasted with individuals with a high expression (see Chapter 1.3.1). Following this strategy, the second

approach of this work is to compare the differentiability of tasks between individuals with low and high expression in known proxies, such as physical fitness or expertise.

Furthermore, detecting and better understanding global patterns of reorganization is critical to contextualize individual trajectories. As presented in Chapter 1.4.3, group classifiers and unsupervised methods such as dimensionality reduction techniques could be used. The third approach is, therefore, to use these methods exploratively at the group level to extract patterns and relationships from the complex data sets.

Taken together, this work focuses on studying age-related phenomena, such as dedifferentiation, and investigating the replicability of hypotheses, such as reserve. To achieve this, four empirical studies use datasets with participants from different life stages and lifestyles, including work experience and physical fitness. These datasets include experiments covering sensory, motor, and cognitive domains. Results from the analysis are presented in the following research articles that focus on specific sub-questions.

In **Research Article I** we followed the first approach presented above. We investigated the difference in the performance of classifiers trained to discriminate EEG derived brain network activation patterns during visuomotor tracking tasks between younger and older participants. The aim here was to draw conclusions about the reorganization of the motor system and extend a previous publication that found differences between younger and older adults in EEG markers of sensorimotor processing during visuomotor tracking tasks [11].

Following this approach, **Research Article II**, aimed to investigate whether the cortical representation of inhibitory control differs across different age groups ranging from children to older adults. Again, previously published findings, in which distinct mechanisms of selective attention in older adults and children were detected using classical ERP analyses, should be extended [116]. To this end, performance on the classification of two stimulus types of a flanker task, i.e., one with high demands on inhibitory control and one with low demands on inhibitory control, was compared between different age groups. Furthermore, following the third explorative approach, we investigated whether we can train a classifier to determine to which age group a participant belongs based on the EEG data.

Research Article III aimed to examine the potential influence of cardiorespiratory fitness, a lifestyle factor, on brain network patterns of dedifferentiation extracted through dimensionality reduction applied to EEG. This investigation followed the second presented approach and was motivated by the reserve hypothesis, which postulates that cardiorespiratory fitness could impact age-related brain reorganization and the observed patterns of dedifferentiation. While this has already been shown in fMRI studies mainly concerning resting-state brain networks [117], it is not clear whether the differentiability of task-related information processing is affected as well and

whether this is reflected in the EEG.

In addition to cardiorespiratory fitness, another significant lifestyle factor is professional expertise. Therefore, the subsequent **Research Article IV** aimed to characterize middle-aged experts using supervised and unsupervised machine learning techniques. In doing so, machine learning methods were applied as a complement to previous studies in which expertise-related differences were investigated utilizing classical statistical methods [6, 11] in order to detect the influence of professional expertise on the dedifferentiation of brain network activation patterns during fine motor tasks and following the third approach to better understand the phenomenon of expertise employing group classifications.

In summary, the application of machine learning followed three approaches presented in the research articles with the goal of better understanding individual trajectories and overarching patterns of age-related brain reorganization. The first two approaches followed established hypotheses of age-related reorganization (dedifferentiation, and reserve), while the third approach aimed to provide exploratory insights and novel findings. Table 2.1 summarizes the application of these approaches in each research article.

Table 2.1: Summary of the approaches followed.

	Approach 1	Approach 2	Approach 3
	Dedifferentiation	Reserve	Overreaching patterns
Research Article I	X		
Research Article II	X		X
Research Article III	X	X	
Research Article IV	X	X	X

Applying machine learning methods on individual and group levels will allow concluding markers of brain reorganization and help identify the individual status and overarching trajectories. The information gained from these tools could be used to determine and evaluate intervention programs, on-the-job-trainings, and support diagnosis. It may have applications in developing assistive technological systems by providing insights into decoding performance in different age groups and its relation to brain reorganization.

Chapter 3

General Methodology

This chapter introduces the overarching methods of the published research articles on which this thesis is based. The focus is rather on an overall description of the methods, which is crucial for the following summary of the main results than on details in terms of reproducibility. For this, reference is made to the original description as it can be found in the published research articles.

3.1 Datasets

The datasets were selected from experiments in published projects in which different study paradigms were used to investigate age-related differences between age groups and groups with different lifestyle backgrounds. The following is a brief description of the datasets. All included participants gave their written informed consent. For children, guardians gave their written informed consent, and children agreed to participate. Table 3.1 gives an overview of which datasets were used in the respective Research Article I to IV presented in Chapter 4.

3.1.1 Dataset I

Dataset I was collected in one of the experiments conducted in the context of the Bremen-Hand-Study@Jacobs, which investigated the influence of age and expertise on manual dexterity over the working life [118]. This study was in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the German Psychological Society. The dataset was analyzed in Research Article I and IV.

Participants

There are recordings of 59 participants in the dataset. The dataset represents a subset of the Bremen-Hand-Study@Jacobs and considers only participants with valid EEG measurements in the respective experiment. Based on their age and profession, each participant is labeled as a young novice, early middle-aged novice, early middle-aged expert, late middle-aged novice, or late middle-aged expert. The group of novices comprised professional profiles whose daily routine

did not require fine motor control of the hands, such as service workers, insurance agents, office workers, and students. Experts, on the other hand, referred to persons with more than ten years of professional experience in a job with pronounced fine motor requirements for hand control, such as opticians, goldsmiths, dentists, dental technicians, or hearing aid technicians [119]. In Research Article I, only the young ($n = 13$, age: 18 to 25 years) and late middle-aged ($n = 13$, age: 55 to 65 years) novices were considered. In Research Article IV all early and late middle-aged experts ($n = 22$, age: 34 to 65 years), as well as all early and late middle-aged novices ($n = 21$, age: 35 to 64 years), were included in the analyses¹.

Experimental Procedures

The experiment conducted was a force-tracking experiment (see Figure 3.1 and Published Research Article I and II for experimental details). Participants held a force transducer between the thumb and index finger of their respective right and left hands. The task was to apply the correct force to track a target force line displayed on a screen as accurately as possible. A total of 160 trials were conducted. For the first 40 trials, a steady line and the following 40 trials, a sinusoidal line had to be followed both with the right hand. This sequence, i.e., 40 times steady and 40 times sinusoidal force tracking, was then repeated with the left hand.

Grip force and EEG were recorded. Before the experiments resting EEGs with eyes open and eyes closed were recorded for 30 s each while participants sat comfortably on a chair.

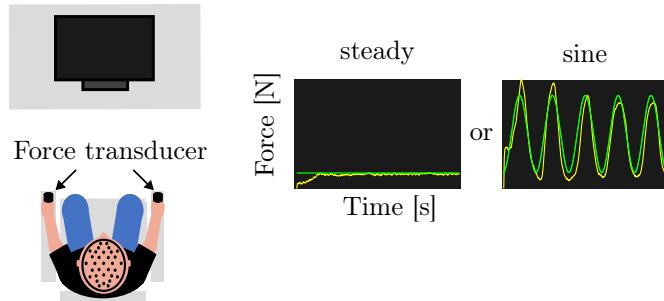


Figure 3.1: Schematic presentation of the force-tracking task conducted in Dataset I. The task was to apply the correct force to a force transducer using a precision grip with the right or left hand to track a target force level (green line) as precisely as possible. Participants received feedback, i.e., they saw their applied force (yellow line).

¹Due to incorrect trigger position and insufficient data quality five participants of the initial sample were excluded

3.1.2 Dataset II

Dataset II contains recordings from three experimental studies, each focusing on a different age group and referred to below as Study 1, Study 2, and Study 3.

Study 1 is the Bremen-Hand-Study@Jacobs presented above. Study 2, the Re-LOAD project, investigated the relationship between motor learning and cognitive function in older adults [83, 120]. Study 3, the CEBRIS project, investigated the influence of physical training on the cognitive functions of children [121]. The German Psychological Society granted ethical approval for Study 1 and 2 and the Ethics Committee of the Faculty of Humanities of the Saarland University, Germany, for Study 3. The dataset is described detailed in [116] and was analyzed in Research Article II.

Participants

The full dataset contains recordings from 222 participants, including 92 participants recorded in Study 1, 81 in Study 2, and 49 in Study 3. Participants are separated into the following age categories [116]: children ($n = 46$, age: 8 to 10 years), young adults ($n = 39$, age: 20 to 29 years), early middle-aged adults ($n = 21$, age: 36 to 48 years), late middle-aged adults ($n = 25$, age: 55 to 64 years), old adults <75 ($n = 40$, age: 66 to 75 years), very old adults >75 ($n = 38$, age: 76 to 83 years)².

Experimental Procedures

All participants performed a modified version of the Flanker task previously reported in Reuter *et al.* [82], Winneke *et al.* [122, 123], and summarized in Reuter *et al.* [116] (see Figure 3.2 and Published Research Article II). The task was to press the correct key corresponding to a central target stimulus surrounded by distracting flanker stimuli as quickly as possible. In Study 1 and Study 3, participants performed 300 trials (approx. 100 trials per stimulus), whereas, in Study 2, they performed 150 trials (approx. 50 trials per stimulus) in randomized order. Other than this, the same experimental procedures were used in all studies, including the EEG measurements system.

3.1.3 Dataset III

Dataset III was collected during an intervention study at Paderborn University in which the feasibility of learning to play golf, as well as its impact on cognitive performance and on (neuro-)biological markers, was investigated [5, 8, 10]. This intervention study was registered at the German Clinical Trials Register (DRKS00014921) and approved by the Ethics Committee of the University of Muenster. The dataset is described in detail in Published Research Article III.

²In the sample used here 13 participants were excluded due to insufficient data quality.

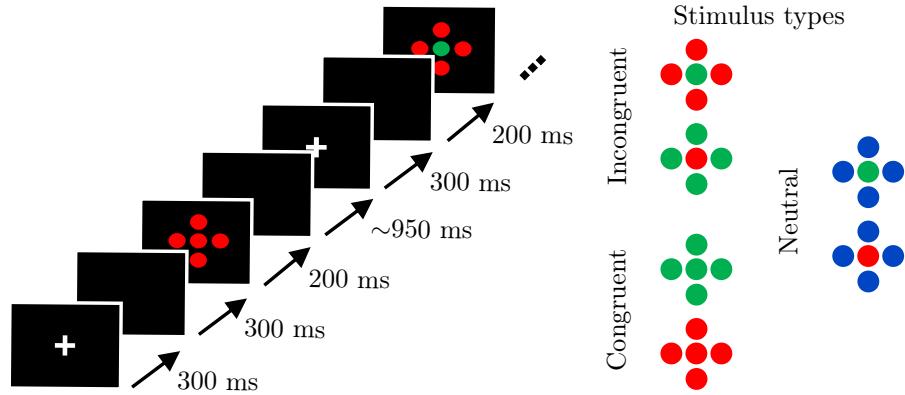


Figure 3.2: Schematic presentation of the flanker task conducted in Dataset II. The task was to press the correct key corresponding to a central target stimulus surrounded by distracting flanker stimuli as quickly as possible. Adapted from Winneke *et al.* [123] with permission.

Participants

The dataset contains recordings from 41 elderly participants. Based on their performance on a 6 min walking test participants' cardiorespiratory fitness was assessed in preceding appointments and two groups were formed, a less fit ($n = 16$, age: 63 to 77 years) and a fit group ($n = 15$, age: 60 to 66 years).

Experimental Procedures

Participants performed sensory, motor, and cognitive tasks each lasting 90 s during which EEG was recorded (see Figure 3.3) Prior to the tasks, EEG was recorded for four minutes in a rest condition in a supine position with eyes closed.

Sensory Task The index fingertips of the participant's left hand were stimulated with the pins of a braille device presenting two pin configurations in random order.

Motor Task The motor task corresponded to the force tracking experiment as described in Chapter 3.1.1. Here, the target was a line that moved from the right to the left on the screen for 90 s and changed level every 3 s in randomized order between heights representing 10 %, 20 % and 30 % of participants' individual maximum voluntary contraction (MVC). The task was to apply the appropriate force to the force transducer to move a cursor so that it followed the target force level presented on a screen as closely as possible.

Cognitive Task The cognitive task was an auditory n-back task. Participants were asked to listen to a sequence of letters presented via two speakers behind them and press the foot switch with the right foot if a letter appeared again two letters later (2-back).

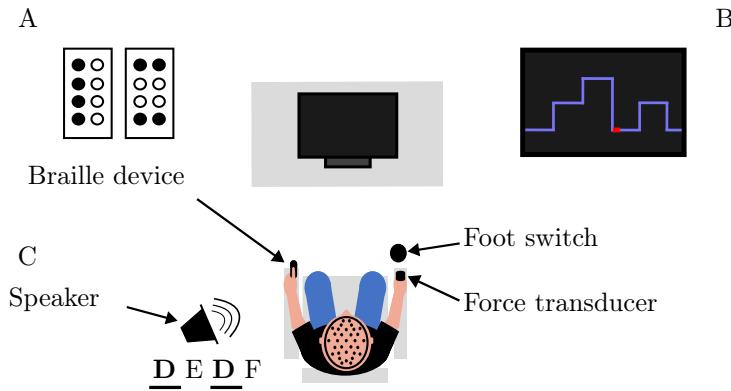


Figure 3.3: Schematic presentation of the motor, sensory and cognitive tasks conducted in Dataset III. The sensory task involved the stimulation of the index finger with two pin patterns of a braille device (A). The motor task was to apply the correct force to a force transducer using a precision grip with the right hand to track a target force level (blue line) as precisely as possible. Participants received feedback, i.e., they saw their applied force (red bar) (B). The cognitive task involved listening to a sequence of letters presented via two speakers behind them (C) and pressing the foot switch with the right foot if a letter appeared again two letters later (2-back).

3.1.4 Electroencephalography: Recording and Preprocessing

The EEGs in Dataset I and Dataset II were recorded with a 32-channel Biosemi active electrode system (ActiveTwo, BioSemi, Amsterdam, Netherlands). In Dataset III EEG measurements were recorded with a 128-channel actiCap active electrode system (BrainProducts, Munich, Germany).

Based on the different analysis methods, objectives, and EEG systems used, the preprocessing differed slightly between the research articles. In each case, the data were down-sampled, re-referenced, and filtered. In Research Article I, III and IV, ICA was used to reduce ocular and muscle artifacts. In Research Article I and IV, trials containing large artifacts were excluded using the *autoreject pipeline* [124]. Due to the high number of electrodes used in Dataset III, bridges were detected and interpolated based on coherence measurements and electrical distance [125]. The exact pipelines are described in the published research articles.

3.2 Machine Learning Procedures

Applying machine learning techniques followed the state-of-the-art practices presented in Chapter 1.4 to enable the three approaches presented in Chapter 2. For this purpose, we used a combination of dimensionality reduction methods and classification.

As a preliminary step, in all approaches taken in this work, the dimensions of the EEG data were reduced to extract specific characteristics, e.g., the network characteristics (see Figure 3.4 1a). Additionally, a second level of dimensionality reduction was optionally applied. This secondary

Table 3.1: Overview of datasets and participants in each research article.

Research Article	Dataset	EEG-System	Paradigm	Participants		
				n	Age range [years]	
I	I	Biosemi: 32-channels	Force control	Late m-a adults	13	55–65
				Young adults	13	18–25
II	II	Biosemi 32-channels	Flanker	Children	46	8–10
				Young adults	39	20–29
				Early m-a adults	21	36–48
				Late m-a adults	25	55–64
				Old adults >75	40	66–75
				Old adults <75	38	76–83
III	III	actiCap: 128-channels	N-back	Low fit	16	63–77
			Tactile oddball	Fit	15	60–66
			Force control			
IV	I	Biosemi: 32-channels	Force control	Experts	22	34–65
				Novices	21	35–64

m-a: middle-aged

reduction served as additional feature extraction for the following analyses or facilitated the exploratory analysis performed in the third approach of this work (see Figure 3.4 1b).

Classification procedures were used at the individual and group levels (see Figure 3.4 2). At the individual level, a classification model was trained for each participant based on the EEG data to predict which experimental condition was present, e.g., which task a participant performed. Following the first approach, this should ultimately allow inferences about dedifferentiation at the individual level and, following the second approach, provide the basis for comparing groups with different lifestyle backgrounds. In addition, following the third approach, classifiers were also trained at the group level to predict group membership and detect general patterns in the dataset. Table 3.2 summarizes the used methods for each research article. In the following, the methods will be briefly described.

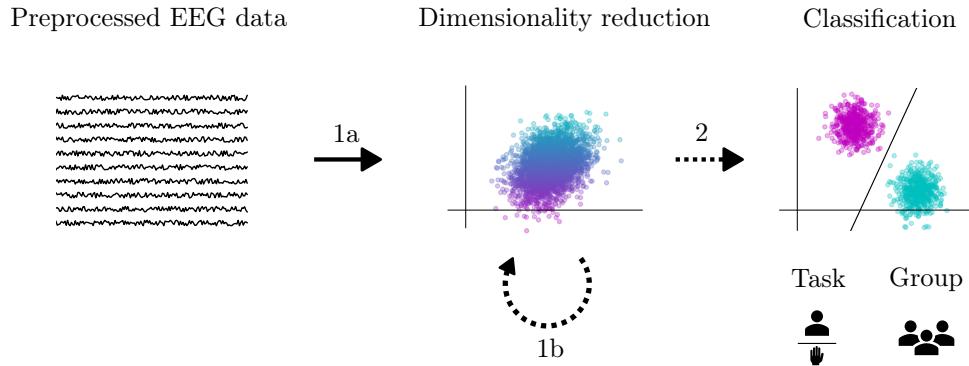


Figure 3.4: Machine learning approach used in this thesis. Dimensionality reduction was used to produce a suitable representation of the EEG data (1a). Optionally a second level dimensionality reduction was applied to extract patterns and for visualization (1b). Finally, Classification was used to classify either the task at the individual level or the group membership (2) using either the results of 1a or 1b.

Table 3.2: Dimensionality reduction and classification methods utilized in each research article

	Dimensionality reduction		Classification	
	Level 1a	Level 1b	Task	Group
Research Article I	DMD	CSP	LDA	SVM
Research Article II	xDAWN	-	SVM	SVM
Research Article III	DMD	PCA	-	-
Research Article IV	DMD	CSP, UMAP	LDA	SVM

DMD: dynamic mode decomposition, CSP: common spatial patterns, LDA: linear discriminant analysis, SVM: support vector machine, PCA: principal component analysis

3.2.1 Dimensionality Reduction

As described above, different dimensionality reduction methods were utilized to extract a suitable representation of the EEG on the first level and further investigate this representation on the second level. The following briefly introduces the methods listed in Table 3.2.

Dynamic Mode Decomposition

To extract dynamic brain network activation patterns from the continuous EEG activity, we chose DMD because it can decompose the signals into spatial activation patterns (modes) that are dynamically coherent, reflecting the network nature of the underlying brain activity [102]. For the computation of DMD, we used the *exact DMD* algorithm introduced by Tu *et al.* [126] and described and validated by Brunton *et al.* [102] for the application to electrophysiological data.

The analysis was based on the preprocessed and windowed EEG data and DMD modes associated with the frequency ranges θ (4 to < 7 Hz), α (7 to < 12 Hz), β_1 (12 to < 16 Hz), and β_2 (16 to < 30 Hz) were considered. We calculated the DMD mode magnitude (absolute value) to obtain the influence of each electrode in a DMD mode representing coherent activation [102].

Common Spatial Patterns

To extract the information from the DMD modes that would allow the best possible differentiation between the tasks, we leveraged supervised dimensionality reduction (see Chapter 1.4.2). The approach followed in this work is based on filter based common spatial patterns (FBCSP), a widely used algorithm for the classification of continuous tasks that extracts a weighting for each EEG channel that maximizes the class discriminative energy for selected frequency bands [127]. By multiplying these weights with the channel values, meaningful features are generated. The weightings were calculated based on DMD magnitudes in each frequency band (see Published Research Article I for details on the implementation).

Principal Component Analysis

PCA aims at extracting the statistically most descriptive components from highly dimensional data [128]. We used the singular value decomposition (SVD) algorithm and reduced the dimension of the DMD modes of all time windows to their main features. Therefore, we calculated a SVD over all frequency-specific modes over all time windows per participant and extracted the singular vectors and values. The singular vectors represent the principal components and capture the most significant or dominant DMD patterns. The singular values capture the proportion of variance accounted for by this pattern. Higher singular values indicate that the associated dominant pattern captures more variation among all DMD modes during task completion and can be considered representative of the stability or prominence of this pattern.

Uniform Manifold Matrix Approximation and Projection

While PCA was used to explain patterns of variation within a participant, we relied on UMAP to capture the structure both on a local level, i.e., within a participant, and on a global level, i.e., between participants. UMAP constructs a low-dimensional representation by modeling the data as a topological manifold, considering both the distances between data points and the local density, and is particularly effective in visualizing and exploring complex data patterns and meaningful relationships in the data [91]. For this, we first calculated the arithmetic mean over the DMD modes of all windows per frequency band, trial, and participant and then applied UMAP to obtain a lower-dimensional representation that captures underlying patterns and meaningful relationships in the data, facilitating visualization and exploration of DMD patterns.

xDAWN

To process the event-related EEG data, we utilized the xDAWN algorithm [104]. By applying the xDAWN algorithm, it is possible to obtain a set of weights that emphasize the relevant EEG activity while suppressing noise and artifacts, leading to improved signal quality for further analysis or classification tasks. In this way, it is possible to induce activation patterns, i.e., neural responses to external stimuli at the level of individual trials. In contrast to the previous methods, we did not use DMD first but applied xDAWN to the preprocessed trials to ensure comparability to previous ERP analyses [116].

3.2.2 Classification

The classification was performed on a trial-by-trial basis, i.e., models were learned that can predict for each participant to which task a given trial belongs or cross-participant to which group the performing participant belongs. For model selection and evaluation, we relied on nested cross-validation (see Chapter 1.4.2). For all datasets, ten splits were used in the inner and outer cross-validation loops, keeping 80 % of the dataset for training and 20 % for testing. In the inner loop, hyperparameters were tuned according to a grid search procedure in which a given parameter space is provided to select the best-fitting parameters [95]. Consequently, one model was tested for each of the ten outer splits. The final metrics are the average over the outer splits and represent the performance of the classification algorithm. For model evaluation, we constructed confusion matrices summarizing all correct and false predicted instances of the test set [129] (see Table 3.3). Additionally, we calculated the common performance metrics which are summarized in Table 3.4.

Table 3.3: Confusion matrix. If an actual value is positive and is classified as positive, it is a true positive (TP) result. If the positive instance is classified as negative, it is a false negative (FN) result. If an instance is actually negative and classified as negative, it is called true negative (TN). Consequently, if an instance is negative and classified as positive, it is a false positive (FP) [129].

		Predicted Class	
		Positive	Negative
Actual Class	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

To estimate the significance of these metrics, we related them to the theoretical chance level, i.e., the metrics that would result if the class labels were randomly assigned. However, because this leads to a biased estimate of the significance of model performance when the distribution is nonuniform and the data sets are small, we additionally determined the random level of the values using permutation procedures. This involved repeatedly shuffling the labels at random to produce a null distribution against which to test the significance of a model. For the larger data sets, however, we relied on an empirical approach and estimated the significance of the metrics by computing a

Table 3.4: Summary of metrics to evaluate model performance

Metric	Formula	Description
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$	Measures the overall correctness of the model's predictions.
Precision	$\frac{TP}{TP + FP}$	Measures the proportion of true positive predictions among all positive predictions.
Recall (Sensitivity)	$\frac{TP}{TP + FN}$	Measures the proportion of true positive predictions among all actual positive samples.
Specificity	$\frac{TN}{TN + FP}$	Measures the proportion of true negative predictions among all actual negative samples.
F1 Score	$\frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$	Harmonic mean of precision and recall, providing a balanced measure between the two.
AUC	-	Area under the receiver operating characteristic curve, which measures the performance of a binary classification model across various threshold settings.

TP: true positive, TN: true negative, FP: false positive, FN: false negative, AUC: area under the receiver operating characteristic curve

binomial distribution. The appropriate method was chosen individually depending on the dataset, model objective, and distribution of classes. The results consequently mark the reference values as theoretical, analytical, or empirical chance level. A detailed justification of the chosen methods is described in the published research articles.

In the research articles, we experimented with multiple algorithms based on recent literature to find the most suitable models for the respective tasks [130]. In doing so, we finally used support vector machine (SVM) and linear discriminant analysis (LDA), which both have a comprehensive application history to neuroscience data and are well suited to our specific problems. SVMs is one of the most commonly used algorithms for neuroscience data [95]. The algorithm generates optimal decision boundaries, called hyperplanes. It can thus handle both linearly separable and non-linearly separable data by using kernel functions to map the input space into a higher-dimensional feature space [130]. LDA is specifically used in the context of BCI and has proven to be successful in task decoding [103]. This algorithm aims to find a linear combination of features that maximizes

the separation between different classes, making it particularly effective when the classes are well-separated [130].

All machine learning pipelines were implemented following the framework of the Python package scikit-learn [131]. If the methods were not implemented, they were implemented using custom scripts in Python or Matlab. A detailed description of the complete analysis pipelines and details of the respective implementation can be found in the published research articles.

Chapter 4

Summary of the Main Results

In the following, the main results of the published research articles will be briefly presented. The style of the illustrations published in the respective articles has been adapted from the original publication. The authors' contributions can be found in the corresponding published research articles at the end of the thesis and have been reviewed and approved by all authors.

4.1 Research Article I

Goelz, C. *et al.* Classification of visuomotor tasks based on electroencephalographic data depends on age-related differences in brain activity patterns. *Neural Networks* **142**, 363–374 (2021)

As described in Chapter 2, dedifferentiation was operationalized in this article as the loss of specificity of brain networks affecting the discriminability of task-related network characteristics as shown in studies utilizing fMRI [45, 52]. We, therefore, investigated differences in classification performance between younger and older adults in visuomotor tracking tasks to infer age-related dedifferentiation of the motor system.

To describe the classifier input, we used classical statistical methods first. Next, we trained a classifier for each participant to output the appropriate task label based on the EEG data recorded during a trial. That is, the classifier should be able to identify, given the EEG data, whether the participant was tracking a sinusoidal target force with the left or right hand or whether the task was to track a steady target force with the right or left hand. To gain deeper insight into the classification performance, we trained classifiers to predict only the hand side or only the task characteristic individually. We compared the DMD derived brain network patterns between the groups and tasks with permutation t-tests and the classification performance between the groups with Man-Whitney-U tests.

We found significant differences in the expression of DMD patterns between the tasks and groups focusing mainly on central and posterior electrodes, most pronounced in the β frequency bands but also in the α and θ frequency ranges (see Figures 2 to 4 in Published Research Article I for the statistical values per electrode). In addition, there were group differences in the spatial distribution concerning a more bilateral and frontal expression of the patterns in the group of late middle-aged adults.

Overall the classifiers' performance was above chance level within all participants (accuracy: $M = 0.66$, $SD = 0.11$; theoretical chance level: 0.25) and yielded distinct misclassification patterns between younger and late middle-aged adults (see Figure 4.1). The classification of the hand side in the late middle-aged participants performed significantly worse compared to young adults, i.e., the classifier misclassified trials performed with the right hand as left hand trials and vice versa (accuracy late middle-aged adults: $M = 0.70$, $SD = 0.08$; accuracy young adults: $M = 0.82$, $SD = 0.09$; $U = 39.5$, $p = 0.02$, $r = 0.45$). On the other hand, the classification of which target force was followed, i.e., steady vs. sinusoidal, worked significantly better in the late middle-aged adults (accuracy late middle-aged adults: $M = 0.86$, $SD = 0.09$; accuracy young adults: $M = 0.75$, $SD = 0.09$; $U = 40.00$, $p = 0.02$, $r = 0.45$). Consequently, fewer steady trials were classified as sinusoidal trials in the late middle-aged compared to the younger participants or vice versa.

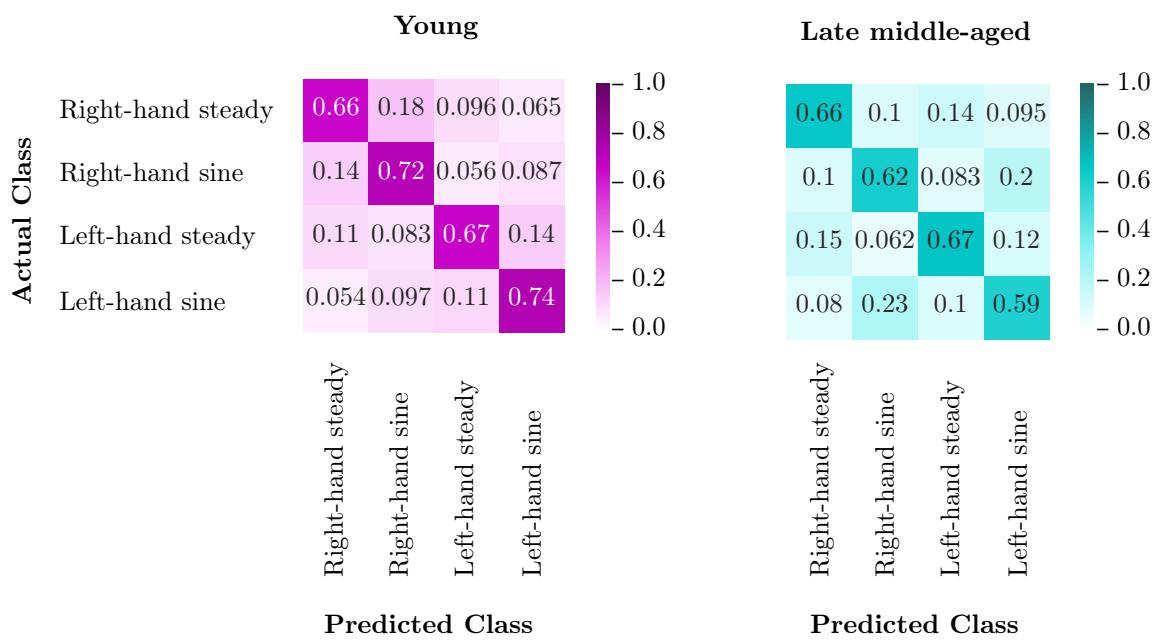


Figure 4.1: Main results of Published Research Article I. The misclassification rate of the hand side (left vs. right) was higher, and the misclassification rate of task type (sinusoidal vs. steady) was lower in late middle-aged compared to young adults.

The lower classification performance in the classification of the body side, i.e., left vs. right hand task execution in the group of late middle-aged adults, points to a less specific and less segregated brain network activation of the motor system. In contrast, the higher classification performance

with respect to the task characteristic sinusoidal vs. steady force tracking might indicate a higher level of compensatory involvement when the task gets more demanding. These findings demonstrate that the reorganization of brain networks is reflected in the classification performance.

4.2 Research Article II

Goelz, C. *et al.* Classification of age groups and task conditions provides additional evidence for differences in electrophysiological correlates of inhibitory control across the lifespan. *Brain Informatics* **10**, 11 (2023)

Following the previous approach, the discriminability of inhibitory and non-inhibitory stimuli within a participant should be investigated using classification techniques. The intention was to study differences in the neural representation of inhibitory control across age groups and extend previous findings on the same data [116].

In this study, we trained a classifier for each participant that could predict which stimulus, i.e., congruent (requires no inhibitory control) or incongruent (requires inhibitory control), was presented based on the EEG data (see Chapter 3.1.2). We also examined how well classification worked over time after stimulus presentation, extracting the time points at which the most accurate classification was possible and the level of performance at that time. We compared the classification trajectories recorded in this manner between the different age groups using one-way ANOVAs, or Kruskal-Wallis tests, followed by t-tests or Dunn's tests for post-hoc comparisons.

To further investigate the group structure, we also trained a group-level classifier to predict which age group the performing participant belongs to based on the EEG recording of a trial.

The maximum classification performance of the model trained to predict which stimulus was presented within a participant was above the chance level in over 95 % of the participants (AUC: M = 0.72, SD = 0.06, analytical chance level: 0.61). The classification performance was further dependent on the group [$F(5,206) = 4.805$, $p < 0.001$], with classification performance lower in the children's group compared to the other groups ($p < 0.05$). When comparing the trajectories, we found that this also differed between the groups ($H(5) = 35.575$, $p < 0.001$) with later performance peaks in the children and the two oldest age groups ($p < 0.05$) (see Figure 4.2 A).

The classification of group membership was overall above chance level (accuracy: 0.55, empirical chance level: 0.17), and a characteristic pattern of misclassifications emerged (see Figure 4.2 B).

The classification of children was the most accurate. An increasing number of misclassifications were observed in the other age groups, where the classifier incorrectly assigned trials to a neighboring age group. This resulted in clusters of adjacent age groups within which this misclassification

accumulated. The first cluster ranged from young adults to late middle-aged adults, and the second cluster included the two oldest age groups. There were higher misclassification rates within these clusters but fewer misclassifications between clusters, especially between the two oldest groups and the adjacent group of late middle-aged adults.

We also studied the time points for which the classification performance of the group model was highest and found a 10 % performance increase after stimulus onset compared to before, with a peak at 100 ms to 200 ms.

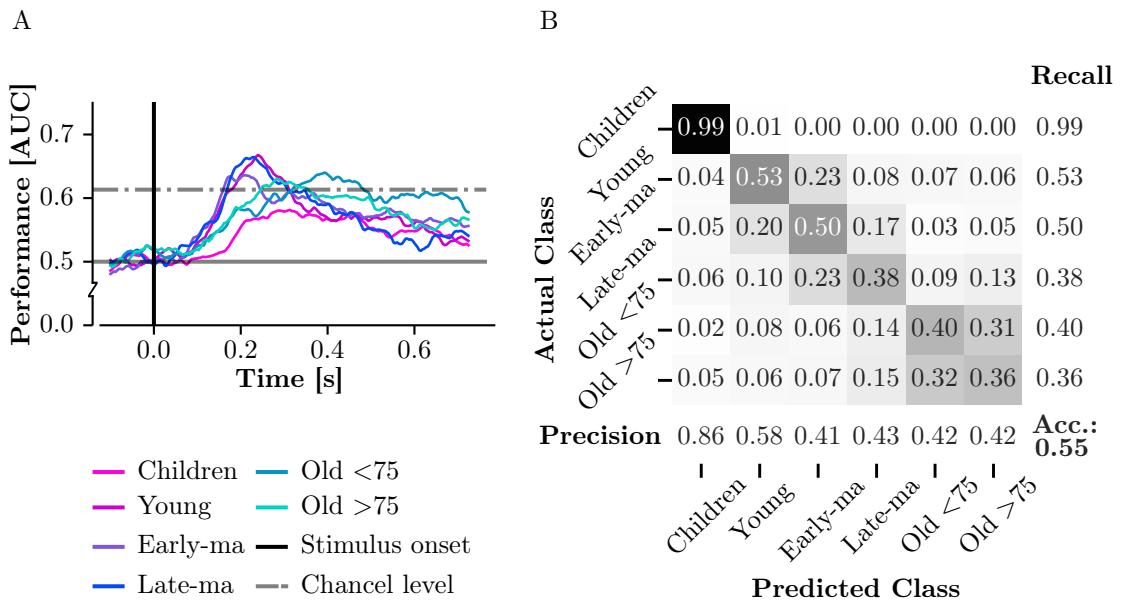


Figure 4.2: Main results of Published Research Article II. The evolution of classification performance of models trained to discriminate between congruent and incongruent trials differed between age groups. Mean trajectories per age group are shown here (A). Classifying between these age groups revealed clusters of groups in which misclassifications happened predominately (B). Acc.: accuracy, AUC: area under the receiver operating characteristic curve, ma: middle-aged.

The results of the task classification suggest that the distinctiveness of the cortical representation of inhibitory control does not differ with older age but that different time windows and, therefore, different processes are important for selective attention at different ages. The higher classification performance during the task than before stimulus onset underscores the added value of task-related EEG. The grouped structure of misclassifications, especially the comparable fewer misclassifications between the oldest group and the late middle-aged groups, could reflect gross changes, e.g., after retirement.

4.3 Research Article III

Goelz, C. *et al.* Electrophysiological signatures of dedifferentiation differ between fit and less fit older adults. *Cognitive Neurodynamics* **15**, 1–13 (2021)

This study followed the assumption from the reserve hypothesis that life factors, such as physical fitness, influence age-related reorganization of the brain. Thus, the aim was to investigate the influence of cardiorespiratory fitness on the dedifferentiation of task-related brain network patterns at rest and during tasks representing the sensory, motor, and cognitive domains, respectively.

We compared the dominant DMD patterns derived by SVD per frequency band between the tasks using permutation t-tests to infer the differentiability of task-related DMD modes in fit and less fit participants. For a statistical evaluation of group differences, we compared the multivariate distribution of obtained t-values with Cramér tests between the groups. We further compared the singular values associated with the dominant pattern to infer the prominence of that pattern throughout task execution between the groups using repeated measures ANOVAs.

The comparison of t-distributions with Cramér tests between the groups revealed higher differentiability of dominant DMD modes in the fit compared to the less fit participants in all frequency bands (all $p < 0.05$, see Figure 3 and Table 2 in Published Research Article III for exact statistical values). However, the difference in the differentiability was most pronounced in the θ and β_2 frequency bands (see Figure 4.3). Furthermore, a significantly lower proportion of total variance could be explained by the dominant pattern in the β_2 frequency range for the less fit compared to the fit group [$F(1,29) = 12.572$, $p = 0.001$, partial $\eta^2 = 0.300$] in the motor (fit: $M = 80.5\%$, $SD = 0.60\%$, less fit: $M = 79.95\%$, $SD = 0.52\%$), the sensory (fit: $M = 80.8\%$, $SD = 0.76\%$, less fit: $M = 80.05\%$, $SD = 0.53\%$) and the cognitive task (fit: $M = 81.07\%$, $SD = 0.78\%$, less fit: $M = 80.21\%$, $SD = 0.54\%$).

The higher degree of task differentiability in the fit group compared to the unfit group supports the idea that physical fitness manifests in task-related brain activation patterns consistent with lower dedifferentiation in older adults. The higher proportion of explained variance in the fit participants can be interpreted as higher prominence of the patterns in this group due to lower noise levels, which is consistent with the predictions by the computational model of dedifferentiation. These findings are consistent with the predictions of the reserve hypothesis and support assumptions from the computational model of dedifferentiation.

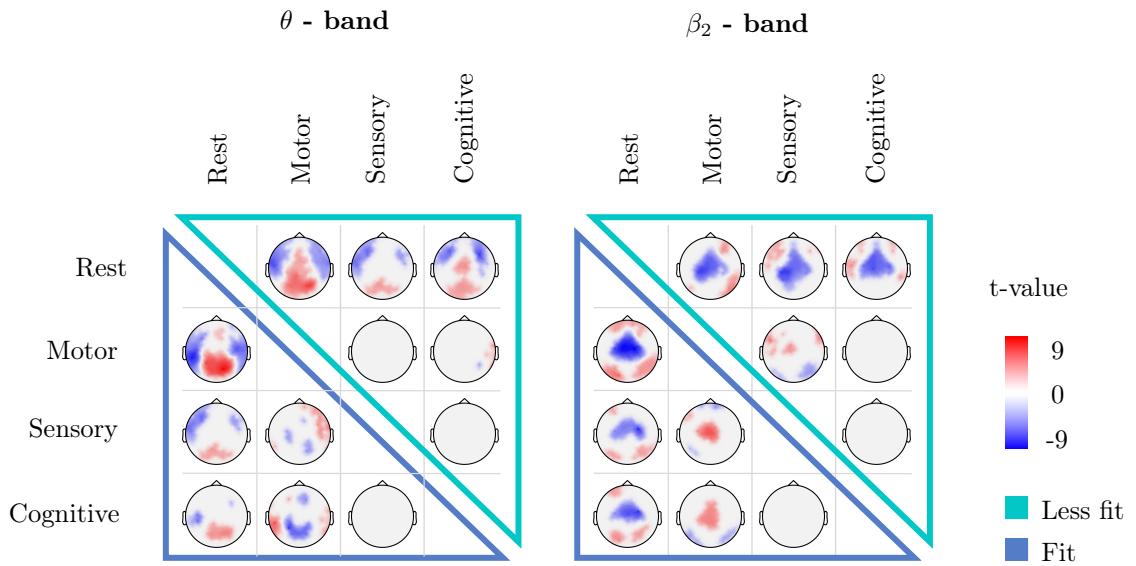


Figure 4.3: Main results of Published Research Article III. Statistical t-maps of significant differences of dominant DMD mode patterns in the θ and β_2 frequency bands between the tasks. Fit participants (blue) showed higher specificity of task-related patterns than less fit participants (turquoise), indicated by higher pronounced differences between the tasks.

4.4 Research Article IV

Gaidai, R., Goelz, C., et al. Classification characteristics of fine motor experts based on electroencephalographic and force tracking data. *Brain Research* **1792**, 148001 (2022)

This study examined professional expertise, another important factor believed to influence age-related reorganization of the brain. The goal was to characterize middle-aged fine motor experts and novices performing a visuomotor tracking task using classification procedures at task and group levels.

As described previously, we used classical statistical methods first to compare the inputs to the classification algorithms. Analogous to Chapter 4.1, we trained a classifier for each participant that outputs the type of task, i.e., steady or sinusoidal force tracking with the left and right hands. We compared the classification performance, i.e., the differentiability of the tasks, between the groups. In addition, we trained a group-level classifier to predict whether the participant performing the force tracking is a fine motor expert or a novice. We used EEG and the force tracking data as input to the classification pipelines.

To further examine the EEG patterns at the trial level, we transformed the DMD modes of all participants via UMAP. We quantified the clustering structure of the UMAP embedding by calculating the Euclidean distance of each trial within each participant. We also computed the centroids

of all trials per participant and determined the Euclidean distance between these centroids.

Although the groups differed at the behavioral level in force tracking with the left hand, we found no statistical differences in DMD modes. Classification of task levels for all subjects performed better than chance (accuracy: $M = 0.68$, $SD = 0.13$, theoretical chance level: 0.25). There were no differences in classification performance between experts (accuracy: $M = 0.66$, $SD = 0.16$) and novices (accuracy: $M = 0.70$, $SD = 0.10$) [$t(41) = 0.96$, $p = 0.35$].

Classification of group membership was not possibly better than chance (theoretical chance level: 0.5), neither based on the EEG ($M = 0.53$, $SD = 0.07$) nor based on the force tracking data ($M = 0.43$, $SD = 0.16$).

Visualization of the EEG feature space revealed patterns of individuality in both groups (see Figure 4.4). This is expressed in a structure of small clusters, with each cluster assigned to one participant. Our comparison of the distances showed a more considerable distance between the individual clusters of the experts ($M = 7.26$, $SD = 2.25$) than between those of the novices ($M = 3.92$, $SD = 1.33$) with more compact clusters for the experts ($M = 1.03$, $SD = 0.55$) compared to the novices ($M = 1.34$, $SD = 0.48$).

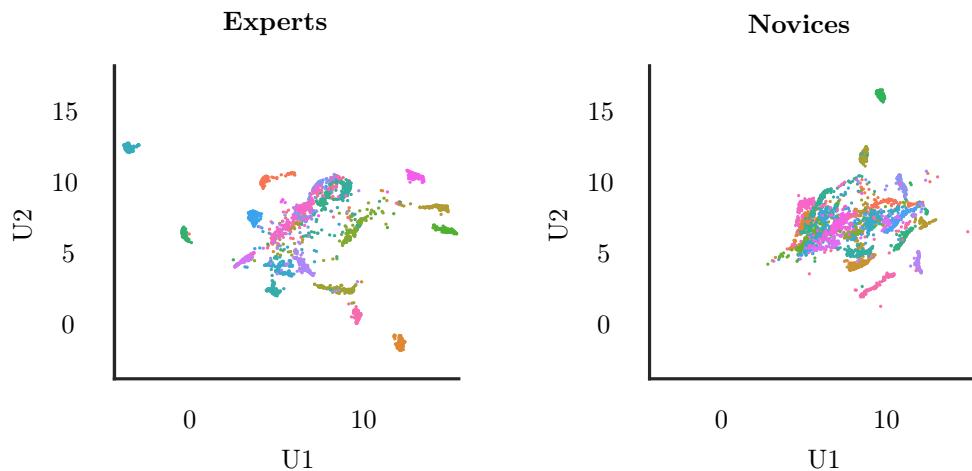


Figure 4.4: Main results of Published Research Article IV. UMAP embedding of EEG feature space of fine motor experts and novices. Each color corresponds to one participant; each dot corresponds to one trial. The experts' data clustered structure is more dispersed, and clusters are more compact.

The results of task classification suggest that contrary to what was predicted by the reserve hypothesis, expertise does not influence the differentiation of task representations. Also, we could not classify between experts and novices. However, the analysis of the group structure indicates a higher individuality of task-related brain activation patterns in experts compared to novices.

The results suggest that professional expertise leads to the development of a very individual and task-specific pattern of central control of tasks that corresponds to the context of the expertise.

Chapter 5

General Discussion

This work aimed to better understand the reorganization of the aging brain by applying machine learning techniques to EEG data. Three approaches were derived from the literature and used in four publications. The first approach was based on a reorganization of the aging brain often described in the literature, namely dedifferentiation, i.e., less specific activation of brain networks expressed in a lower differentiability of brain states. Based on the computational model by Li *et al.* [36, 47], this is attributed to a loss of selectivity of neuronal responsiveness, which has been confirmed in numerous animal and human studies. Building upon this, we tested the differentiability of brain states, e.g., in the form of activation patterns of brain networks, and thus the selectivity of neuronal representations using machine learning methods. This approach provided insights into age-related dedifferentiation and allowed the formulation of new hypotheses, e.g., regarding this phenomenon's temporal aspect.

The second approach of this work addressed the reserve hypothesis. We examined the possible positive effects of certain lifestyle factors, such as physical fitness or professional expertise, on age-related brain reorganization. Using machine learning techniques, we compared individuals with different lifestyle backgrounds and the effects on age-related brain reorganization. This analysis provided new perspectives and helped to test the assumption that individuals with certain lifestyle factors exhibit less dedifferentiation.

In the third approach, dimensionality reduction and group-level classification algorithms were used to identify overarching patterns in brain reorganization. This approach enabled the generation of new insights and hypotheses.

In the following, the main results of these approaches are discussed. However, for a detailed discussion of each specific result, please refer to the corresponding published journal articles.

5.1 On the Detection of Dedifferentiated Brain Organization

Motivated by applications from research using fMRI, in which dedifferentiation is quantified based on the performance of classification algorithms, we applied this to EEG data to investigate the differentiability of signals in motor, cognitive and sensory tasks in participants of different age groups.

In Research Article I, we found differences in the classification performance of fine motor tasks based on EEG data. In particular, performance in classifying which hand was used to perform the task was reduced in late middle-aged compared to younger participants. Studying the classifier input, i.e., the network activation patterns extracted with DMD, we found signs of a more bilateral and widespread activation pattern in late middle-aged adults which might explain the difference in the classifier performance. These results correspond to a dedifferentiated motor network as reported in fMRI studies and are also consistent with other EEG findings that have compared classification between age groups finding reduced performance in the classification of the body side of task execution in older adults [52, 113, 132, 133]. In contrast, we found an increased classification performance in the late middle-aged participants when predicting whether the participant tracked a steady or sinusoidal target force. When considering the classifier input again, we found a stronger frontal involvement in the network activation patterns of the late middle-aged participants and thus suspected that compensation processes play a role here. Compensation could be present, especially in the more complex sinusoidal force tracking, and thus positively influence the differentiability of sinusoidal and steady force tracking. This result is consistent with the CRUNCH hypothesis, which postulated that during task performance, as task difficulty (or load) increases, more cortical resources will be activated [58]. Combining both results, we could demonstrate that age-related reorganization of the aging brain is reflected in the performance of classifiers based on EEG measurements, thus providing new insights into the aging brain.

In Research Article II, we again utilized classification algorithms to investigate whether the cortical representation of selective attention differed between six different age groups ranging from children to older adults. Here we found that the discriminability between stimuli requiring and stimuli not requiring inhibitory control was reduced only in the children's group. This reduced discriminability in the group of children goes along with the assumption of inhibitory control in children that is not yet differentiated at the neural level [116, 134]. However, the results of similar good performance in the classification of the two stimulus categories contradict the assumption of a general dedifferentiation of all cortical systems. Similar results are also known in fMRI studies, which, for example, found no dedifferentiation in the activation of the visual system in response to certain stimulus categories [135]. By using EEG, however, we were able to investigate the temporal response to the stimulus in addition. For this reason, the classifier input did not represent coherent activation patterns as in the previously described analysis, but rather single-trial EEG traces filtered by xDAWN in response to the stimuli, through which it was possible to examine the performance of the classifiers over time and thus gain insights into the temporal dynamics of

information processing. In this way, we found that the differentiability of the stimuli was delayed in older adults. Thus, dedifferentiation might also have a dynamic component related to the rise of neural noise, as postulated in the computational model of Li *et al.* [36, 47]. However, this only occurred in the two oldest age groups and not, as in the previous study, in the late middle-aged, so this effect might occur later in the aging process.



5.2 The Role of Lifestyle Factors



An important aspect to better understand the individuality of age-related changes is to investigate how influencing factors affect the reorganization of the aging brain. From the literature, physical fitness can be deduced as a significant influencing factor that affects the brain's ability to resist degradation processes and maintain function, thus contributing to reserve. In this context, Stillman *et al.* [117] summarized findings of fMRI studies proving that the network structure of the brain at rest of physically fit people show fewer signs of dedifferentiation. Therefore, in Research Article III we investigated to what extent cardiorespiratory fitness affects the dedifferentiation of task-related network activation patterns based on EEG. To do so, we extracted dominant spatiotemporal network patterns and their prominence over time during the performance of motor, sensory and cognitive tasks. In this way, we were able to estimate task-related brain network activation patterns in the main domains in which age-related changes are reported and compare this between a fit and less fit group [33, 37, 46]. This comparison revealed greater differences between the task-related activation patterns in the fit compared to the less fit group, suggesting lower levels of dedifferentiation in the less fit group. We further found that the extracted dominant patterns of the fit group resolved more variance compared to the less fit group, indicating higher stability or prominence of the patterns in the fit group. These differences were frequency dependent mainly to the θ and β_2 bands which could be carefully attributed to processes reflecting cognitive involvement and information integration [71]. From the perspective of the computational model of dedifferentiation, the higher stability could indicate less neural noise and a lower rate of neural variability, which points to a higher efficiency of information processing. Overall, these results follow the prediction of the reserve hypothesis and provide evidence that supports the assumptions of the computational model of dedifferentiation.

Another frequently studied factor of influence in the literature is professional expertise. Thus, in Research Article IV, we compared fine motor experts and novices with respect to the classification performance of visuomotor tasks. We found no group differences in the performance of a classifier trained to classify the same task characteristics as in Research Article I. Consequently, we could not detect any differences in the differentiation of the neural representation of the task between novices and experts. At first glance, these results contradict previous studies from the Bremen-Hand-Study@Jacobs, which showed statistical differences in the expression of network characteristics of experts compared to novices in terms of neural efficiency [6, 11]. However, our study tested a particular assumption of age-related reorganization, namely dedifferentiation, and

did not test neural efficiency as in the previous work. Also, in contrast to Research Article III, the results seem to contradict the reserve concept, based on which a lower level of dedifferentiation would have been expected. Therefore, taking a more differentiated view of the reserve concept concerning individual factors could be necessary to refine this concept. We have only explored one aspect in considering dedifferentiation in the aging brain. As a result, other characteristics of brain reorganization may differ, such as the efficiency of information processing or compensational network involvement. These factors were not accounted for in our analyses.

5.3 Exploratory Insights into Age-Related Reorganization

While the previous results refer to the application at the participant level, we also used the machine learning methods at the group level to gain explorative insights into overreaching group characteristics.

In Research Article II, we found a grouped structure of misclassifications when using a classifier that was trained to determine the group membership in one of the six age groups based on EEG data. The classifier tended to misclassify only between certain age groups. This was especially noticeable in the two oldest age groups, between which the misclassification rate was highest. Less frequently, however, older participants were classified into the group of late middle-aged adults and vice versa. Considering the age structure of the older adults (minimum age of 65 years) and the middle-aged adults (maximum age of 64 years), as well as a retirement age of around 65, this could indicate major cognitive changes after retirement. Deterioration of cognitive abilities and cerebral blood flow after retirement has been reported in the literature [136–138]. However, we could not further test this hypothesis because of a lack of data on the exact retirement age of the participants. In addition, no misclassification occurred between children and the oldest groups. Although differences based on ERP markers are often described as u-shaped [116, 139], this result suggests different mechanisms, i.e., differentiation in children vs. dedifferentiation in older adults. Overall, we found an increase in classification performance shortly after stimulus onset of about 10 %, highlighting the added value of task-related EEG recordings and suggesting early processing of the stimulus to be affected by age-related reorganization.

Also, in Research Article IV, a classifier was trained based on the EEG data to determine the membership to the group of experts or novices. In contrast to the previous article, however, the performance was below the chance level. This finding contrasts with studies in which the classification of the level of expertise was already successful [140–142]. In contrast to these studies, the professional background of the fine motor experts examined in this study, such as opticians and watchmakers, was very diverse, and the visuomotor tracking tasks selected for analysis offered only a general approximation of the context of their expertise. Based on this, we visualized the underlying brain network characteristics using a nonlinear dimensionality reduction in two-dimensional space. Differences in the data structure between experts and novices emerged. Compared to each other, the brain network patterns of the experts showed a higher degree of a clustered structure

in which each cluster could be clearly assigned to a participant. This finding can be interpreted as a higher individuality of the brain network patterns in the experts. A high degree of individuality is also known in the movement patterns of flute players and the muscle synergies of expert powerlifters [143–145]. Our results suggest that individuality, which was gained in the context of professional expertise, is as well present at the network level of the brain.

Overall, the results presented in the four research articles showed that machine learning techniques, including dimensionality reduction and classification of EEG signals, successfully detected age-related brain reorganization. These results provided new starting points for understanding the reorganization process of the aging brain, particularly in relation to dedifferentiation, and partially supported the predictions of the reserve hypothesis.

5.4 Methodological Considerations

Before discussing specific methodological aspects regarding the characteristics of the datasets as well as the application of machine learning procedures, it should be noted that we have not presented *changes* but *differences* as the study design of all studies from which the datasets were derived were cross-sectional. We also focused only on some well-documented specific aspects of the reorganization of the aging brain and explored overarching patterns in an exploratory manner. In addition, the results should be evaluated in light of the following methodological considerations.

5.4.1 Datasets

Concerning the individual datasets and the underlying study protocols, some considerations should be taken into account when evaluating the results presented.

Different age groups were present in the datasets underlying the individual research articles. The older (late middle-aged) participants in Dataset I examined in Research Article I were comparatively young, so for these results, it is to be expected that possible age effects have been underestimated. In Dataset II, the minimum age difference between the six age groups varied from ten years (between children and young adults) to one year (within the two oldest groups). This age structure may have significantly conditioned the classification results in Research Article II. However, the minimum age difference between the oldest and the next younger groups (late middle-aged adults) is also only two years. Hence, the finding of misclassification cutoff between these groups is less likely to be explained by the age structure of the dataset. In general, regarding the age structure, it would be advantageous to have a continuous age range for a fine-grained analysis of age-related reorganization.

Concerning the proxies of the reserve construct, it should be noted that the construct of the reserve hypothesis is more diverse than the dimensions studied in Research Articles III and IV. In our considerations, we followed the common practice of comparing individuals with low expression with individuals with higher expression of a given proxy, i.e., cardiorespiratory fitness or expertise

[45]. Cardiorespiratory fitness was indirectly measured by a 6 min walk test, which depends on the participants' motivation. However, a high correlation to the more objective VO₂max is reported in the literature [146]. The selection of the fine motor experts was based on the professional group of precision mechanics, whose daily routines and hand use may differ. Conversely, it cannot be excluded that the novices also had a certain level of fine motor skill, which they acquired in their everyday life, so that expertise effects may have been underestimated. However, by the criterion of a minimum of ten years of professional experience, the deliberate practice approach was followed, which was evaluated as a suitable criterion for professional expertise [118, 119]. Furthermore, previous studies have shown that the selected fine motor task appropriately reflects the expertise context [6, 11, 147, 148]. In general, a comparison to young participants could have tested further assumptions of the reserve hypothesis. For example, whether different networks are activated and have higher task specificity or whether we can assume that networks are maintained. However, this was not possible during participant recruitment due to the selected criterion, i.e., at least ten years of work experience in Dataset I or was not considered in the original intervention study on which Dataset III is based on.

In addition to the age structure and the proxies of the reserve concept, it should be noted that the tasks evaluated comprised particular laboratory tasks, whose transferability to everyday life may be limited. Nevertheless, the tasks should map across relevant domains in which age-related changes are reported, i.e., sensory, cognitive, and motor domains. The task types were all based on well-tested and common paradigms in aging research and had the advantage that the results were comparable to other studies. One aspect we cannot exclude is the possible effect of fatigue and motivation, which could have played a role, especially during longer sessions. Other aspects relate to the specific characteristics of the task designs. For example, the force control task in Dataset I was designed in a block design, and the participants already attended several similar experiments in the context of the Bremen-Hand-Study@Jacobs, so practice effects cannot be excluded. These effects could also have been advantageous since age differences in short-term adaptation could thus be eliminated. The flanker task in Dataset II differed slightly between the individual studies on which this dataset is based. This was present in the length of stimulus presentation, number of trials, and gender. However, because the participants of the individual studies were present in several groups and we did not find increased misclassification rates between these groups, we excluded that this had a significant influence on the results. In addition, the influence of gender and the number of trials in previous or comparable studies was estimated to be low [116, 149]. In Dataset III, the task was quite long, which may have contributed to the fatigue effects mentioned above. Still, it allowed the temporal dynamics to be better mapped using machine learning algorithms.

5.4.2 Machine Learning Approaches

Applying machine learning algorithms depends on the dataset and utilized algorithm. Especially different noise levels in the individual EEG recordings could have influenced the results of the group comparisons. However, studies using comparable methods show that this approach is valid,

provided that the noise level is accounted for [149, 150]. We used dimensionality reduction methods on the first level to improve the signal-to-noise ratio and to extract relevant features from the preprocessed EEG data.

In Research Article I, III, and IV we used DMD, which allows to extract spatiotemporal coherent patterns from the EEG data [102]. This approach allowed us to draw inferences about the dynamic properties of complex networks but differs from analyses of bivariate connectivity of voxel or (source) signals followed by graph analytic approaches to describe the brain networks derived in this way. The decision to use DMD in this study was driven by its ability to capture the complex dynamics of networks that include factors such as frequency, growth, and decay. As a result, DMD provides a physiologically plausible and low-dimensional representation of EEG dynamics. Moreover, the effectiveness of DMD in combination with machine learning has been demonstrated in previous studies [102, 151, 152]. In particular, DMD has been shown to be very successful in classifying fine motor tasks [152]. Additionally, our research has already used DMD to extract task-related sensorimotor network dynamics associated with age and expertise [6, 11]. Also, DMD in this work focused on task-related network characteristics derived from EEG signals and not on resting state networks. Thus, we cannot conclude the relationship to the reorganization of large-scale resting state brain networks known in fMRI research, i.e., less separated and modular resting networks. Although it is generally unclear how task-related dedifferentiation and the reorganization of resting networks are related, a recent study shows that these two levels are related and predict performance [153]. Another aspect to consider here is that our analyses were performed in the signal space and were not projected into the anatomical source space, so the anatomical interpretation was not possible with respect to the known resting state networks. This decision was justified by the small number of electrodes in Dataset I and II. Studies show that a valid source reconstruction depends strongly on the number of electrodes [154, 155]. Only for Dataset III a sufficient number of electrodes was given. We, therefore, visualized the found patterns in the source space (see Figure 2 in Published Research Article III) but did not perform DMD in the source space because validation studies for such an approach have not been performed so far. We only had individual MRIs for some participants, so no individual head models were available.

To align our results with previous ERP analyses and enhance the signal-to-noise ratio, we employed the xDAWN algorithm in Research Article IV. This choice allowed us to explore the temporal dynamics in response to stimuli rather than inferring network characteristics, as done in the other research articles. While calculating dynamic functional connectivity over time would have been an alternative approach, it is less extensively validated at this stage. Hence, we relied on the established xDAWN algorithm for our investigation.

The selection of classification algorithms and secondary dimensionality reduction algorithms may also have affected the results obtained. The process of picking suitable algorithms followed standard procedures in the field of machine learning [89]. That is, the choice of the dimensionality reduction algorithms was based on the respective objectives of the analyses, and the selection of the classifiers was based on the performance of the respective algorithm within the training portions.

In addition, we replicated the results obtained by the classifiers with different algorithms and provided them as supplementary material in the respective published research articles. The results of the classification algorithms were carefully evaluated using various metrics such as accuracy, precision, recall, and F1 scores. We used classical algorithms that fit better to the amount of data available, i.e., tens to hundreds of samples. Additionally, we used cross-validation procedures and hyperparameter tuning, e.g., to tune the regularization parameters, limit overfitting, and estimate the generalization performance. 

At this point, it should be mentioned that in Research Article III, we did not use supervised classification to map dedifferentiation but instead extracted the patterns unsupervised via PCA and tested the discriminability using classical statistical methods. This decision was due to the relative amount of data available and the continuous design of the task. For this reason, it is possible that the effects of dedifferentiation were underestimated in both groups studied.

Although the classification results do not allow conclusions about the direction of the effects, these analyses should be seen as a supplement. In all research articles, we, therefore, related the results to previous research and examined the input to the classification with methods that allow conclusions about directional effects. For this purpose, we partly used inferential statistical methods and dimensionality reduction methods. As such, bridging the gap between traditional statistical and modern machine learning analysis was possible.

5.5 Outlook and Practical Implications

The results of this work have provided valuable insights at both the individual and group levels and have implications for future research and practical applications. The analyses conducted in Research Article I have shown that dedifferentiation of the motor system can be detected using EEG measures even in late middle-aged individuals. Therefore, it would be exciting to further explore dedifferentiation in continuous age groups and to investigate whether different brain systems exhibit specific trajectories, such as those reported in MRI findings [156]. We found initial evidence for this in Research Article II, in which we discovered reorganization of attentional control only in the two oldest groups and not, as with respect to the motor system, already in the late middle-aged participants.

Moreover, the temporal dynamics of dedifferentiation has not been studied so far and provides a new starting point for future studies. Supporting this idea, the results of Research Article III suggest that the stability of cortical representations could be a crucial point. Drawing on the computational model by Li *et al.* [36, 47], we propose neural noise to explain our observation. Although initial findings suggest that neural noise may explain dedifferentiation at the EEG level, this requires further investigation [157].

Regarding the reserve concept, we examined only two dimensions and found varying results, so a systematic investigation of further lifestyle factors, such as education, would be of interest. To understand the concept concerning the reorganization of the brain, the development of a gold

standard tool that captures all or as many as possible of the known domains of reserve would be necessary [158].

Exploratory group-level analysis also generated two hypotheses that could be tested in future research. Concerning expertise, the exploratory analysis of the group structure in Research Article III allowed us to show that the factor of individuality, which was previously known only at the muscular or behavioral level, might also exist at the neuronal level. Furthermore, Research Article II suggests significant changes in the functional organization of the brain following retirement, as indicated by the classification of age groups.

In addition to these points, the findings of this work have several practical implications. First, our analyses have shown that the reorganization of the aging brain, particularly the dedifferentiation, can be mapped at the individual level. Although the clinical relevance of dedifferentiation is not fully understood, Fornito *et al.* [56] have described it as a maladaptive mechanism in brain networks. Therefore, our results could be valuable for the early detection of unfavorable age-related changes and contribute to developing and evaluating therapeutic concepts and interventions.

Additionally, in Research Article II, our results demonstrated that task-related EEG data improved age group prediction performance by up to 10 %. This highlights the potential of using task-based EEG in clinical practice to predict the cognitive status of patients and may be particularly relevant for EEG-based detection of mild cognitive impairment, for which resting state EEG alone was shown to be of limited value [77, 78]. An exciting starting point would be to integrate the methods used in this work into the framework of brain age prediction (see Chapter 1.4.3). Perhaps this would help to overcome some limitations of the current framework, such as the limited ability to predict cognitive decline [159]. Interestingly, studies suggest that a dedifferentiated activation pattern of brain networks may be related to poorer behavioral performance independent of age [45], so this may have applications in other areas independent of aging research.

The results of this work may also be relevant for the development of assistive technology for the rehabilitation and prevention of declining cognitive and motor abilities of older adults. In this context, the results could be particularly important for developing adaptive systems, i.e., systems that automatically adapt to the user. Such technologies could be based on individual EEG patterns and adapt the training content based on the signatures of dedifferentiation.

The development of such systems, especially BCIs, often relies on the differentiability of EEG features, e.g., based on the laterality of certain frequency features. However, our results suggest that this may lead to lower classifier performance, and rather age-specific features, such as higher frontality, could be exploited here. Consequently, our results could help address a possible age bias in developing assistive technology.

5.6 Conclusion

By using machine learning methods, this work has contributed to understanding age-related reorganization of the brain. More specifically, using classification and dimensionality reduction methods, this work quantified dedifferentiation based on EEG measurements at the individual level. The results indicate that the expression and characteristics of dedifferentiation could be very different in different brain systems. For example, we found dedifferentiation of the motor system already in late middle-age. In contrast, differences in elements of the attentional system only became apparent in older age and are related to temporal aspects of information processing. It was also confirmed that reorganization depends on lifestyle factors such as cardiorespiratory fitness or work experience. However, different factors could also lead to different adaptation mechanisms and contribute to individuality in old age, underlining the need for a more differentiated view of the reserve concept. 

In addition, two hypotheses emerged from exploratory machine learning analyses at the group level. The first is that professional expertise leads to an individualized neural representation of domain-specific tasks, and the second is that significant changes in the brain's functional organization occur after retirement.

Taken together, predictions of the computational model of dedifferentiation and the reserve hypothesis could be tested and confirmed. In addition, new hypotheses regarding the dynamics of brain reorganization and lifestyle factors were generated. The findings of this work may contribute to the development of markers of age-related changes and, thus, to the detection of unfavorable trajectories. Furthermore, the findings could contribute to the development of age-appropriate assistive systems. Thus, the findings are relevant to promoting healthy aging and developing technical systems that enable older adults to participate in society.

References

1. Gaidai, R., Goelz, C., Mora, K., Rudisch, J., Reuter, E.-M., Godde, B., Reinsberger, C., Voelcker-Rehage, C. & Vieluf, S. Classification characteristics of fine motor experts based on electroencephalographic and force tracking data. *Brain Research* **1792**, 148001 (2022).
2. Goelz, C., Mora, K., Rudisch, J., Gaidai, R., Reuter, E.-M., Godde, B., Reinsberger, C., Voelcker-Rehage, C. & Vieluf, S. Classification of visuomotor tasks based on electroencephalographic data depends on age-related differences in brain activity patterns. *Neural Networks* **142**, 363–374 (2021).
3. Goelz, C., Mora, K., Stroehlein, J., Haase, F., Dellnitz, M., Reinsberger, C. & Vieluf, S. Electrophysiological signatures of dedifferentiation differ between fit and less fit older adults. *Cognitive Neurodynamics* **15**, 1–13 (2021).
4. Goelz, C., Reuter, E.-M., Fröhlich, S., Rudisch, J., Godde, B., Vieluf, S. & Voelcker-Rehage, C. Classification of age groups and task conditions provides additional evidence for differences in electrophysiological correlates of inhibitory control across the lifespan. *Brain Informatics* **10**, 11 (2023).
5. Gowik, J. K., Goelz, C., Vieluf, S., van den Bongard, F. & Reinsberger, C. Source connectivity patterns in the default mode network differ between elderly golf-novices and non-golfers. *Scientific Reports* **13**, 6215 (2023).
6. Götz, C., Voelcker-Rehage, C., Mora, K., Reuter, E.-M., Godde, B., Dellnitz, M., Reinsberger, C. & Vieluf, S. Improved Neural Control of Movements Manifests in Expertise-Related Differences in Force Output and Brain Network Dynamics. *Frontiers in Physiology* **9** (2018).
7. Niess, A., Widmann, M., Gaidai, R., Götz, C., Schubert, I., Castillo, K., Sachs, J. P., Bizjak, D., Vollrath, S., Wimbauer, F., Vogel, A., Keller, K., Burgstahler, C., Quermann, A., Kerling, A., Schneider, G., Zacher, J., Diebold, K., Grummt, M., Beckendorf, C., Buitenhuis, J., Egger, F., Venhorst, a., Morath, O., Barsch, F., Mellwig, K.-P., Oesterschlink, J., Wüstenfeld, J., Predel, H.-G., Deibert, P., Friedmann-Bette, B., Mayer, F., Hirschmüller, A., Halle, M., Steinacker, J. M., Wolfarth, B., Meyer, T., Böttinger, E., Flechtner-Mors, M., Bloch, W., Haller, B., Roecker, K. & Reinsberger, C. COVID-19 in German Competitive Sports:

- Protocol for a Prospective Multicenter Cohort Study (CoSmo-S). *International Journal of Public Health* **67** (2022).
8. Stroehlein, J. K., Vieluf, S., Zimmer, P., Schenk, A., Oberste, M., Goelz, C., van den Bongard, F. & Reinsberger, C. Learning to play golf for elderly people with subjective memory complaints: feasibility of a single-blinded randomized pilot trial. *BMC Neurology* **21**, 200 (2021).
 9. Strote, C., Gölz, C., Stroehlein, J. K., Haase, F. K., Koester, D., Reinsberger, C. & Vieluf, S. Effects of force level and task difficulty on force control performance in elderly people. *Experimental Brain Research* **238**, 2179–2188 (2020).
 10. Ströhlein, J. K., Vieluf, S., van den Bongard, Franziska, Gölz, C. & Reinsberger, C. Golf spielen gegen die Vergesslichkeit: Effekte des Erlernens der Sportart auf das Default Mode Netzwerk des Gehirns. *B&G Bewegungstherapie und Gesundheitssport* **36**, 65–72 (2020).
 11. Vieluf, S., Mora, K., Gölz, C., Reuter, E.-M., Godde, B., Dellnitz, M., Reinsberger, C. & Voelcker-Rehage, C. Age-and expertise-related differences of sensorimotor network dynamics during force control. *Neuroscience* **388**, 203–213 (2018).
 12. Gaidai, R., Gölz, C., Widmann, M., Niess, A. M. & Reinsberger, C. Herausforderungen und Lösungen im Aufbau eines Datenbanksystems im Verbundprojekt „Covid-19 im Spitzensport - Eine multizentrische Kohortenstudie“ (CoSmo-S). Sports Medicine and Health Summit. Hamburg, Germany, Sept. 2023.
 13. Goelz, C., Gaidai, R., Mora, K., Rudisch, J., Reuter, E. M., Godde, B., Reinsberger, C., Voelcker-Rehage, C. & Vieluf, S. Classification characteristics of fine motor experts based on electroencephalographic and force tracking data. 54th Annual Conference of the German Society for Sport Psychology. Muenster, Germany, June 2022.
 14. Ströhlein, J., Gölz, C., Vieluf, S., van den Bongard, F. & Reinsberger, C. Funktionelle Netzwerkcharakteristika des Default Mode Netzwerks bei älteren Golf-Novizen. Sports Medicine and Health Summit. Virtual Conference, Apr. 2021.
 15. Vieluf, S., Reinsberger, C., Gölz, C. & Voelcker-Rehage, C. Age- and expertise related differences of sensorimotor network dynamics during force control. 23rd University Day of the German Society of Sport Science. Munich, Germany, Sept. 2017.
 16. Gaidai, R., Gölz, C., Widmann, M., Niess, A. M. & Reinsberger, C. Herausforderungen und Lösungen im Aufbau eines Datenbanksystems im Verbundprojekt „Covid-19 im Spitzensport - Eine multizentrische Kohortenstudie“ (CoSmo-S). Nachwuchsssymposium Deutsche Gesellschaft für Sportmedizin und Prävention. Tuebingen, Germany, Sept. 2022.
 17. Goelz, C., Mora, K., Stroehlein, J., Reinsberger, C. & Vieluf, S. Electrophysiological signatures of brain network dynamics in elderly. 26th Annual Meeting of the Organization of Human Brain Mapping. Virtual Conference, June 2020.

18. Vieluf, S., Voelcker-Rehage, C., Goelz, C., Mora, K. & Reinsberger, C. *Age- and expertise-related differences in force control manifest in variability*. 27th Annual Meeting of the Society of the Neural Control of Movements. Dublin, Ireland, May 2017.
19. World Health Organization. *Decade of healthy ageing: baseline report* (World Health Organization, Geneva, 2020).
20. Betzel, R. F. & Bassett, D. S. Multi-scale brain networks. *NeuroImage* **160**, 73–83 (2017).
21. Brunton, B. W. & Beyeler, M. Data-driven models in human neuroscience and neuroengineering. *Current Opinion in Neurobiology* **58**. Computational Neuroscience, 21–29 (2019).
22. Appenzeller, T. The AI revolution in science: How deep learning is helping scientists cope with a data deluge. *Science* (2017).
23. Hey, T., Tansley, S., Tolle, K. & Gray, J. *The Fourth Paradigm: Data-Intensive Scientific Discovery* ISBN: 978-0-9825442-0-4 (Microsoft Research, 2009).
24. Leonard, J. J., Mindell, D. A. & Stayton, E. L. Autonomous vehicles, mobility, and employment policy: the roads ahead. *Massachusetts Inst. Technol., Cambridge, MA, Rep. RB02-2020* (2020).
25. Liu, Y., Jain, A., Eng, C., Way, D. H., Lee, K., Bui, P., Kanada, K., de Oliveira Marinho, G., Gallegos, J., Gabriele, S., Gupta, V., Singh, N., Natarajan, V., Hofmann-Wellenhof, R., Corrado, G. S., Peng, L. H., Webster, D. R., Ai, D., Huang, S. J., Liu, Y., Dunn, R. C. & Coz, D. A deep learning system for differential diagnosis of skin diseases. *Nature Medicine* **26**, 900–908 (2020).
26. Adamopoulou, E. & Moussiades, L. *An Overview of Chatbot Technology in Artificial Intelligence Applications and Innovations* (eds Maglogiannis, I., Iliadis, L. & Pimenidis, E.) (Springer International Publishing, Cham, 2020), 373–383.
27. López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. The Hallmarks of Aging. *Cell* **153**, 1194–1217 (2013).
28. Mooney, K. M., Morgan, A. E. & McAuley, M. T. Aging and computational systems biology. *WIREs Systems Biology and Medicine* **8**, 123–139 (2016).
29. Smith, S. M., Elliott, L. T., Alfaro-Almagro, F., McCarthy, P., Nichols, T. E., Douaud, G. & Miller, K. L. Brain aging comprises many modes of structural and functional change with distinct genetic and biophysical associations. *eLife* **9**, e52677 (2020).
30. Cohen, A. A., Ferrucci, L., Fülop, T., Gravel, D., Hao, N., Kriete, a., Levine, M. E., Lipsitz, L. A., Olde Rikkert, M. G. M., Rutenberg, a., Stroustrup, N. & Varadhan, R. A complex systems approach to aging biology. *Nature Aging* **2**, 580–591 (2022).
31. Salthouse, T. A. Trajectories of normal cognitive aging. *Psychology and Aging* **34**, 17–24 (2019).
32. Li, K. Z. & Lindenberger, U. Relations between aging sensory/sensorimotor and cognitive functions. *Neuroscience & Biobehavioral Reviews* **26**, 777–783 (2002).

33. Park, D. C. & Reuter-Lorenz, P. The Adaptive Brain: Aging and Neurocognitive Scaffolding. *Annual Review of Psychology* **60**, 173–196 (2009).
34. Reuter-Lorenz, P. A. & Park, D. C. Human Neuroscience and the Aging Mind: A New Look at Old Problems. *The journals of Gerontology: Series B* **65B**, 405–415 (2010).
35. Fjell, A. M. & Walhovd, K. B. Structural Brain Changes in Aging: Courses, Causes and Cognitive Consequences. *Reviews in the Neurosciences* **21**, 187–222 (2010).
36. Li, S.-C., Lindenberger, U. & Sikström, S. Aging cognition: from neuromodulation to representation. *Trends in Cognitive Sciences* **5**, 479–486 (2001).
37. Sala-Llonch, R., Bartrés-Faz, D. & Junqué, C. Reorganization of brain networks in aging: a review of functional connectivity studies. *Frontiers in Psychology* **6**, 663 (2015).
38. Courtney, S. & Hinault, T. When the time is right: Temporal dynamics of brain activity in healthy aging and dementia. *Progress in Neurobiology* **203**, 102076 (2021).
39. Friston, K. J. Functional and Effective Connectivity: A Review. *Brain Connectivity* **1**, 13–36 (2011).
40. Deery, H. A., Di Paolo, R., Moran, C., Egan, G. F. & Jamadar, S. D. The older adult brain is less modular, more integrated, and less efficient at rest: A systematic review of large-scale resting-state functional brain networks in aging. *Psychophysiology* **60**, e14159 (2023).
41. Uddin, L. Q., Yeo, B. T. T. & Spreng, R. N. Towards a Universal Taxonomy of Macro-scale Functional Human Brain Networks. *Brain Topography* **32**, 926–942 (2019).
42. Sporns, O. Network attributes for segregation and integration in the human brain. *Current Opinion in Neurobiology* **23**, 162–171 (2013).
43. Betzel, R. F., Byrge, L., He, Y., Goñi, J., Zuo, X.-N. & Sporns, O. Changes in structural and functional connectivity among resting-state networks across the human lifespan. *NeuroImage* **102**, 345–357 (2014).
44. Grady, C. The cognitive neuroscience of ageing. *Nature Reviews Neuroscience* **13**, 491–505 (2012).
45. Koen, J. D. & Rugg, M. D. Neural Dedifferentiation in the Aging Brain. *Trends in Cognitive Sciences* **23**, 547–559 (2019).
46. Baltes, P. B. & Lindenberger, U. Emergence of a powerful connection between sensory and cognitive functions across the adult life span: a new window to the study of cognitive aging? *Psychology and Aging* **12**, 12–21 (1997).
47. Li, S.-C., Lindenberger, U. & Frensch, P. A. Unifying cognitive aging: From neuromodulation to representation to cognition. *Neurocomputing* **32-33**, 879–890 (2000).
48. Schmolesky, M. T., Wang, Y., Pu, M. & Leventhal, A. G. Degradation of stimulus selectivity of visual cortical cells in senescent rhesus monkeys. *Nature Neuroscience* **3**, 384–390 (2000).

49. Costa, M., Lepore, F., Prévost, F. & Guillemot, J.-P. Effects of aging on peripheral and central auditory processing in rats. *European journal of Neuroscience* **44**, 2084–2094 (2016).
50. Spengler, F., Godde, B. & Dinse, H. R. Effects of ageing on topographic organization of somatosensory cortex. *NeuroReport* **6** (1995).
51. Tucker-Drob, E. M., Brandmaier, a. M. & Lindenberger, U. Coupled cognitive changes in adulthood: A meta-analysis. *Psychological Bulletin* **145**, 273–301 (2019).
52. Carp, J., Park, J., Hebrank, a., Park, D. C. & Polk, T. A. Age-Related Neural Dedifferentiation in the Motor System. *PLOS ONE* **6**, 1–6 (Dec. 2011).
53. Rieck, J. R., Baracchini, G., Nichol, D., Abdi, H. & Grady, C. L. Reconfiguration and dedifferentiation of functional networks during cognitive control across the adult lifespan. *Neurobiology of Aging* **106**, 80–94 (2021).
54. Geerligs, L., Maurits, N. M., Renken, R. J. & Lorist, M. M. Reduced specificity of functional connectivity in the aging brain during task performance. *Human Brain Mapping* **35**, 319–330 (2014).
55. Antonenko, D. & Flöel, A. Healthy Aging by Staying Selectively Connected: A Mini-Review. *Gerontology* **60**, 3–9 (2013).
56. Fornito, A., Zalesky, a. & Breakspear, M. The connectomics of brain disorders. *Nature Reviews Neuroscience* **16**, 159–172 (2015).
57. Stern, Y. Cognitive reserve. *Neuropsychologia* **47**, 2015–2028 (2009).
58. Festini, S. B., Zahodne, L. & Reuter-Lorenz, P. A. in *Oxford research encyclopedia of psychology* (2018).
59. Davis, S. W., Dennis, N. A., Daselaar, S. M., Fleck, M. S. & Cabeza, R. Qué PASA? The Posterior–Anterior Shift in Aging. *Cerebral Cortex* **18**, 1201–1209 (2007).
60. Cabeza, R. Hemispheric asymmetry reduction in older adults: The HAROLD model. *Psychology and Aging* **17**, 85–100 (2002).
61. Douw, L., Nieboer, D., van Dijk, B. W., Stam, C. J. & Twisk, J. W. R. A Healthy Brain in a Healthy Body: Brain Network Correlates of Physical and Mental Fitness. *PLOS ONE* **9**, 1–8 (Feb. 2014).
62. Cabeza, R., Albert, M., Belleville, S., Craik, F. I. M., Duarte, A., Grady, C. L., Lindenberger, U., Nyberg, L., Park, D. C., Reuter-Lorenz, P. A., Rugg, M. D., Steffener, J. & Rajah, M. N. Maintenance, reserve and compensation: the cognitive neuroscience of healthy ageing. *Nature Reviews Neuroscience* **19**, 701–710 (2018).
63. Esiri, M., Matthews, F., Brayne, C., Ince, P., Matthews, F., Xuereb, J., Broome, J., McKenzie, J., Rossi, M., McKeith, I., Lowe, J. & Morris, J. Pathological correlates of late-onset dementia in a multicentre, community-based population in England and Wales. *The Lancet* **357**, 169–175 (2001).

64. Stern, Y., Habeck, C., Moeller, J., Scarmeas, N., anderson, K. E., Hilton, H. J., Flynn, J., Sackeim, H. & van Heertum, R. Brain Networks Associated with Cognitive Reserve in Healthy Young and Old Adults. *Cerebral Cortex* **15**, 394–402 (2004).
65. Fabel, K., Wolf, S., Ehninger, D., Babu, H., Galicia, P. & Kempermann, G. Additive effects of physical exercise and environmental enrichment on adult hippocampal neurogenesis in mice. *Frontiers in Neuroscience* **3** (2009).
66. Voss, M. W., Weng, T. B., Burzynska, A. Z., Wong, C. N., Cooke, G. E., Clark, R., Fanning, J., Awick, E., Gothe, N. P., Olson, E. A., McAuley, E. & Kramer, A. F. Fitness, but not physical activity, is related to functional integrity of brain networks associated with aging. *NeuroImage* **131**, 113–125 (2016).
67. Soldan, A., Pettigrew, C., Zhu, Y., Wang, M.-C., Bilgel, M., Hou, X., Lu, H., Miller, M. I., Albert, M. & Team, T. B. R. Association of Lifestyle Activities with Functional Brain Connectivity and Relationship to Cognitive Decline among Older Adults. *Cerebral Cortex* **31**, 5637–5651 (2021).
68. Reuter-Lorenz, P. A. & Park, D. C. How Does it STAC Up? Revisiting the Scaffolding Theory of Aging and Cognition. *Neuropsychology Review* **24**, 355–370 (2014).
69. Jackson, A. F. & Bolger, D. J. The neurophysiological bases of EEG and EEG measurement: A review for the rest of us. *Psychophysiology* **51**, 1061–1071 (2014).
70. Cohen, M. X. Where Does EEG Come From and What Does It Mean? *Trends in Neurosciences* **40**, 208–218 (2017).
71. Siegel, M., Donner, T. H. & Engel, a. K. Spectral fingerprints of large-scale neuronal interactions. *Nature Reviews Neuroscience* **13**, 121–134 (2012).
72. Rossini, P. M., Rossi, S., Babiloni, C. & Polich, J. Clinical neurophysiology of aging brain: From normal aging to neurodegeneration. *Progress in Neurobiology* **83**, 375–400 (2007).
73. Ishii, R., Canuet, L., Aoki, Y., Hata, M., Iwase, M., Ikeda, S., Nishida, K. & Ikeda, M. Healthy and pathological brain aging: from the perspective of oscillations, functional connectivity, and signal complexity. *Neuropsychobiology* **75**, 151–161 (2017).
74. Smit, D. J. A., Boersma, M., Schnack, H. G., Micheloyannis, S., Boomsma, D. I., Hulshoff Pol, H. E., Stam, C. J. & de Geus, E. J. C. The Brain Matures with Stronger Functional Connectivity and Decreased Randomness of Its Network. *PLOS ONE* **7**, 1–11 (2012).
75. Samogin, J., Rueda Delgado, L., Taberna, G. A., Swinnen, S. P. & Mantini, D. Age-Related Differences of Frequency-Dependent Functional Connectivity in Brain Networks and Their Link to Motor Performance. *Brain Connectivity* **12** (2022).

76. Babiloni, C., Arakaki, X., Azami, H., Bennys, K., Blinowska, K., Bonanni, L., Bujan, A., Carrillo, M. C., Cichocki, a., de Frutos-Lucas, J., Del Percio, C., Dubois, B., Edelmayer, R., Egan, G., Epelbaum, S., Escudero, J., Evans, A., Farina, F., Fargo, K., Fernández, A., Ferri, R., Frisoni, G., Hampel, H., Harrington, M. G., Jelic, V., Jeong, J., Jiang, Y., Kaminski, M., Kavcic, V., Kilborn, K., Kumar, S., Lam, A., Lim, L., Lizio, R., Lopez, D., Lopez, S., Lucey, B., Maestú, F., McGeown, W. J., McKeith, I., Moretti, D. V., Nobili, F., Noce, G., Olichney, J., Onofrij, M., Osorio, R., Parra-Rodriguez, M., Rajji, T., Ritter, P., Soricelli, a., Stocchi, F., Tarnanas, I., Taylor, J. P., Teipel, S., Tucci, F., Valdes-Sosa, M., Valdes-Sosa, P., Weiergräber, M., Yener, G. & Guntekin, B. Measures of resting state EEG rhythms for clinical trials in Alzheimer's disease: Recommendations of an expert panel. *Alzheimer's & Dementia* **17**, 1528–1553 (2021).
77. Fröhlich, S., Kutz, D. F., Müller, K. & Voelcker-Rehage, C. Characteristics of Resting State EEG Power in 80+-year-Olds of Different Cognitive Status. *Frontiers in Aging Neuroscience* **13** (2021).
78. Farina, F., Emek-Savaş, D., Rueda-Delgado, L., Boyle, R., Kiiski, H., Yener, G. & Whelan, R. A comparison of resting state EEG and structural MRI for classifying Alzheimer's disease and mild cognitive impairment. *NeuroImage* **215**, 116795 (2020).
79. Quandt, F., Bönstrup, M., Schulz, R., Timmermann, J. E., Zimerman, M., Nolte, G. & Hummel, F. C. Spectral Variability in the Aged Brain during Fine Motor Control. *Frontiers in Aging Neuroscience* **8**, 305 (2016).
80. Hong, X., Liu, Y., Sun, J. & Tong, S. Age-Related Differences in the Modulation of Small-World Brain Networks during a Go/NoGo Task. *Frontiers in Aging Neuroscience* **8** (2016).
81. Li, L., Gratton, C., Fabiani, M. & Knight, R. T. Age-related frontoparietal changes during the control of bottom-up and top-down attention: an ERP study. *Neurobiology of Aging* **34**, 477–488 (2013).
82. Reuter, E.-M., Voelcker-Rehage, C., Vieluf, S., Parianen Lesemann, F. & Godde, B. The P3 Parietal-To-Frontal Shift Relates to Age-Related Slowing in a Selective Attention Task. *Journal of Psychophysiology* **31**, 49–66 (2017).
83. Hübner, L., Godde, B. & Voelcker-Rehage, C. Older adults reveal enhanced task-related beta power decreases during a force modulation task. *Behavioural Brain Research* **345**, 104–113 (2018).
84. Samuel, A. L. Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development* **3**, 210–229 (1959).
85. Hastie, T., Tibshirani, R. & Friedman, J. *The Elements of Statistical Learning. Data Mining, Inference, and Prediction, Second Edition* Second. ISBN: 978-0-387-84857-0 (Springer-Verlag, New York, 2009).

86. Von Luxburg, U. & Schölkopf, B. in *Handbook of the History of Logic, Volume 10: Inductive Logic* 651–706 (Elsevier North Holland, Amsterdam, 2011).
87. Rudin, C. & Wagstaff, K. L. Machine learning for science and society. *Machine Learning* **95**, 1–9 (2014).
88. Bzdok, D., Altman, N. & Krzywinski, M. Statistics versus machine learning. *Nature Methods* **15**, 233–234 (2018).
89. Shalev-Shwartz, S. & Ben-David, S. *Understanding Machine Learning: From Theory to Algorithms* ISBN: 9781107057135 (Cambridge University Press, 2014).
90. Murphy, K. P. *Machine learning: a probabilistic perspective* ISBN: 9780262018029 (MIT Press, Cambridge, MA, 2012).
91. McInnes, L., Healy, J. & Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426* (2018).
92. Burkov, a. *The Hundred-Page Machine Learning Book* ISBN: 9781999579517 (andriy Burkov, 2019).
93. Mitchell, T. *Machine Learning* ISBN: 9780071154673 (McGraw-Hill, 1997).
94. Daumé, H. *A Course in Machine Learning* (Hal Daumé III, 2017).
95. Varoquaux, G., Raamana, P. R., Engemann, D. A., andrés Hoyos-Idrobo, Schwartz, Y. & Thirion, B. Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines. *NeuroImage* **145**, 166–179 (2017).
96. Roy, Y., Banville, H., Albuquerque, I., Gramfort, A., Falk, T. H. & Faubert, J. Deep learning-based electroencephalography analysis: a systematic review. *Journal of Neural Engineering* **16**, 051001 (2019).
97. Banville, H., Chehab, O., Hyvärinen, A., Engemann, D.-A. & Gramfort, A. Uncovering the structure of clinical EEG signals with self-supervised learning. *Journal of Neural Engineering* **18**, 046020 (2021).
98. Gemein, L. A., Schirrmeister, R. T., Chrabaszcz, P., Wilson, D., Boedecker, J., andreas Schulze-Bonhage, Hutter, F. & Ball, T. Machine-learning-based diagnostics of EEG pathology. *NeuroImage* **220**, 117021 (2020).
99. Saeidi, M., Karwowski, W., Farahani, F. V., Fiok, K., Taiar, R., Hancock, P. A. & Al-Juaid, A. Neural Decoding of EEG Signals with Machine Learning: A Systematic Review. *Brain Sciences* **11**, 1525 (2021).
100. Khan, S., Hashmi, J. A., Mamashli, F., Michmizos, K., Kitzbichler, M. G., Bharadwaj, H., Bekhti, Y., Ganesan, S., Garel, K.-L. A., Whitfield-Gabrieli, S., Gollub, R. L., Kong, J., Vaina, L. M., Rana, K. D., Stufflebeam, S. M., Härmäläinen, M. S. & Kenet, T. Maturation trajectories of cortical resting-state networks depend on the mediating frequency band. *NeuroImage* **174**, 57–68 (2018).

101. Westner, B. U., Dalal, S. S., Hanslmayr, S. & Staudigl, T. Across-subjects classification of stimulus modality from human MEG high frequency activity. *PLOS Computational Biology* **14**, 1–14 (2018).
102. Brunton, B. W., Johnson, L. A., Ojemann, J. G. & Kutz, J. N. Extracting spatial–temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition. *Journal of Neuroscience Methods* **258**, 1–15 (2016).
103. Blankertz, B., Tomioka, R., Lemm, S., Kawanabe, M. & Muller, K.-r. Optimizing Spatial filters for Robust EEG Single-Trial Analysis. *IEEE Signal Processing Magazine* **25**, 41–56 (2008).
104. Rivet, B., Souloumiac, A., Attina, V. & Gibert, G. xDAWN algorithm to enhance evoked potentials: application to brain–computer interface. *IEEE Transactions on Biomedical Engineering* **56**, 2035–2043 (2009).
105. Woo, C.-W., Chang, L. J., Lindquist, M. A. & Wager, T. D. Building better biomarkers: brain models in translational neuroimaging. *Nature Neuroscience* **20**, 365–377 (2017).
106. Mei, J., Desrosiers, C. & Frasnelli, J. Machine Learning for the Diagnosis of Parkinson’s Disease: A Review of Literature. *Frontiers in Aging Neuroscience* **13** (2021).
107. Engemann, D. A., Mellot, A., Höchenberger, R., Banville, H., Sabbagh, D., Gemein, L., Ball, T. & Gramfort, A. A reusable benchmark of brain-age prediction from M/EEG resting-state signals. *NeuroImage* **262**, 119521 (2022).
108. Gonzeaud, J. *et al.* Accelerated functional brain aging in pre-clinical familial Alzheimer’s disease. *Nature Communications* **12**, 5346 (2021).
109. Saha, S., Mamun, K. A., Ahmed, K., Mostafa, R., Naik, G. R., Darvishi, S., Khandoker, A. H. & Baumert, M. Progress in Brain Computer Interface: Challenges and Opportunities. *Frontiers in Systems Neuroscience* **15**, 578875 (2021).
110. Anumanchipalli, G. K., Chartier, J. & Chang, E. F. Speech synthesis from neural decoding of spoken sentences. *Nature* **568**, 493–498 (2019).
111. Holdgraf, C. R., Rieger, J. W., Micheli, C., Martin, S., Knight, R. T. & Theunissen, F. E. Encoding and Decoding Models in Cognitive Electrophysiology. *Frontiers in Systems Neuroscience* **11** (2017).
112. Park, J., Carp, J., andrew Hebrank, Park, D. C. & Polk, T. A. Neural Specificity Predicts Fluid Processing Ability in Older Adults. *Journal of Neuroscience* **30**, 9253–9259 (2010).
113. Chen, M. L., Fu, D., Boger, J. & Jiang, N. Age-Related Changes in Vibro-Tactile EEG Response and Its Implications in BCI Applications: A Comparison Between Older and Younger Populations. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **27**, 603–610 (2019).

114. Petti, M., Toppi, J., Babiloni, F., Cincotti, F., Mattia, D. & Astolfi, L. EEG Resting-State Brain Topological Reorganization as a Function of Age. *Computational Intelligence and Neuroscience* **2016**, 6243694 (Feb. 2016).
115. Kottlarz, I., Berg, S., Toscano-Tejeida, D., Steinmann, I., Bähr, M., Luther, S., Wilke, M., Parlitz, U. & Schlemmer, A. Extracting Robust Biomarkers From Multichannel EEG Time Series Using Nonlinear Dimensionality Reduction Applied to Ordinal Pattern Statistics and Spectral Quantities. *Frontiers in Physiology* **11**, 614565 (2021).
116. Reuter, E.-M., Vieluf, S., Koutsandreuou, F., Hübner, L., Budde, H., Godde, B. & Voelcker-Rehage, C. A Non-linear Relationship Between Selective Attention and Associated ERP Markers Across the Lifespan. *Frontiers in Psychology* **10**, 30 (2019).
117. Stillman, C., Donofry, S. & Erickson, K. Exercise, Fitness and the Aging Brain: A Review of Functional Connectivity in Aging. *Archives of Psychology* **3** (2019).
118. Voelcker-Rehage, C., Reuter, E.-M., Vieluf, S. & Godde, B. in *Age-Differentiated Work Systems* (eds Schlick, C. M., Frieling, E. & Wegge, J.) 391–415 (Springer, Berlin, Heidelberg, 2013). ISBN: 978-3-642-35057-3.
119. Ericsson, K. & Smith, J. *Toward a General Theory of Expertise: Prospects and Limits* ISBN: 9780521406123 (Cambridge University Press, 1991).
120. Hübner, L., Godde, B. & Voelcker-Rehage, C. Acute Exercise as an Intervention to Trigger Motor Performance and EEG Beta Activity in Older Adults. *Neural Plasticity* **2018**, 4756785 (2018).
121. Koutsandréou, F., Wegner, M., Niemann, C. & Budde, H. Effects of Motor versus Cardiovascular Exercise Training on Children’s Working Memory. *Medicine & Science in Sports & Exercise* **48** (2016).
122. Winneke, A. H., Godde, B., Reuter, E.-M., Vieluf, S. & Voelcker-Rehage, C. The Association Between Physical Activity and Attentional Control in Younger and Older Middle-Aged Adults. *GeroPsych* **25**, 207–221 (2012).
123. Winneke, A. H., Hübner, L., Godde, B. & Voelcker-Rehage, C. Moderate Cardiovascular Exercise Speeds Up Neural Markers of Stimulus Evaluation During Attentional Control Processes. *Journal of Clinical Medicine* **8**, 1348 (2019).
124. Jas, M., Engemann, D. A., Bekhti, Y., Raimondo, F. & Gramfort, A. Autoreject: Automated artifact rejection for MEG and EEG data. *NeuroImage* **159**, 417–429 (2017).
125. Alschuler, D. M., Tenke, C. E., Bruder, G. E. & Kayser, J. Identifying electrode bridging from electrical distance distributions: A survey of publicly-available EEG data using a new method. *Clinical Neurophysiology* **125**, 484–490 (2014).
126. Tu, J. H., Rowley, C. W., Luchtenburg, D. M., Brunton, S. L. & Kutz, J. N. On dynamic mode decomposition: Theory and applications. *Journal of Computational Dynamics* **1**, 391–421 (2014).

127. Ang, K. K., Chin, Z. Y., Wang, C., Guan, C. & Zhang, H. Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b. *Frontiers in Neuroscience* **6**, 39 (2012).
128. Brunton, S. L. & Kutz, J. N. in *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control* 154–194 (Cambridge University Press, 2019).
129. Fawcett, T. An introduction to ROC analysis. *Pattern Recognition Letters* **27**, 861–874 (2006).
130. Shoorangiz, R., Weddell, S. J. & Jones, R. D. in *Handbook of Neuroengineering* (ed Thakor, N. V.) 2463–2501 (Springer Nature Singapore, Singapore, 2023). ISBN: 978-981-16-5540-1.
131. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. & Duchesnay, É. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830 (2011).
132. Zich, C., Debener, S., De Vos, M., Frerichs, S., Maurer, S. & Kranczioch, C. Lateralization patterns of covert but not overt movements change with age: An EEG neurofeedback study. *NeuroImage* **116**, 80–91 (2015).
133. Cassady, K., Ruitenberg, M. F. L., Reuter-Lorenz, P. A., Tommerdahl, M. & Seidler, R. D. Neural Dedifferentiation across the Lifespan in the Motor and Somatosensory Systems. *Cerebral Cortex* **30**, 3704–3716.
134. Waszak, F., Li, S.-C. & Hommel, B. The development of attentional networks: Cross-sectional findings from a life span sample. *Developmental Psychology* **46**, 337–349 (2010).
135. Voss, M. W., Erickson, K. I., Chaddock, L., Prakash, R. S., Colcombe, S. J., Morris, K. S., Doerksen, S., Hu, L., McAuley, E. & Kramer, A. F. Dedifferentiation in the visual cortex: An fMRI investigation of individual differences in older adults. *Brain Research* **1244**, 121–131 (2008).
136. Celidoni, M., Dal Bianco, C. & Weber, G. Retirement and cognitive decline. A longitudinal analysis using SHARE data. *Journal of Health Economics* **56**, 113–125 (2017).
137. Rohwedder, S. & Willis, R. J. Mental Retirement. *Journal of Economic Perspectives* **24**, 119–38 (2010).
138. Rogers, R. L., Meyer, J. S. & Mortel, K. F. After Reaching Retirement Age Physical Activity Sustains Cerebral Perfusion and Cognition. *Journal of the American Geriatrics Society* **38**, 123–128 (1990).
139. Mueller, V., Brehmer, Y., von Oertzen, T., Li, S.-C. & Lindenberger, U. Electrophysiological correlates of selective attention: A lifespan comparison. *BMC Neuroscience* **9**, 18 (2008).
140. Hosp, B. W., Schultz, F., Höner, O. & Kasneci, E. Soccer goalkeeper expertise identification based on eye movements. *PLOS ONE* **16**, 1–22 (2021).

141. Winkler-Schwartz, A., Yilmaz, R., Mirchi, N., Bissonnette, V., Ledwos, N., Siyar, S., Azarnoush, H., Karlik, B. & Del Maestro, R. Machine Learning Identification of Surgical and Operative Factors Associated With Surgical Expertise in Virtual Reality Simulation. *JAMA Network Open* **2**, e198363–e198363 (2019).
142. Shourie, N. Cepstral analysis of EEG during visual perception and mental imagery reveals the influence of artistic expertise. *Journal of Medical Signals & Sensors* **6**, 203–217 (2016).
143. Albrecht, S., Janssen, D., Quarz, E., Newell, K. M. & Schöllhorn, W. I. Individuality of movements in music – Finger and body movements during playing of the flute. *Human Movement Science* **35**, 131–144 (2014).
144. Caramiaux, B., Bevilacqua, F., Wanderley, M. M. & Palmer, C. Dissociable effects of practice variability on learning motor and timing skills. *PLOS ONE* **13**, 1–18 (2018).
145. Kristiansen, M., Madeleine, P., Hansen, E. A. & Samani, A. Inter-subject variability of muscle synergies during bench press in power lifters and untrained individuals. *Scandinavian journal of Medicine & Science in Sports* **25**, 89–97 (2015).
146. Zhang, Q., Lu, H., Pan, S., Lin, Y., Zhou, K. & Wang, L. 6MWT Performance and its Correlations with VO₂ and Handgrip Strength in Home-Dwelling Mid-Aged and Older Chinese. *International Journal of Environmental Research and Public Health* **14**, 473 (2017).
147. Vieluf, S., Mahmoodi, J., Godde, B., Reuter, E.-M. & Voelcker-Rehage, C. The influence of age and work-related expertise on fine motor control. *GeroPsych: The journal of Gerontopsychology and Geriatric Psychiatry* **25**, 199–206 (2012).
148. Vieluf, S., Godde, B., Reuter, E.-M. & Voelcker-Rehage, C. Effects of age and fine motor expertise on the bilateral deficit in force initiation. *Experimental Brain Research* **231**, 107–116 (2013).
149. Vahid, A., Mückschel, M., Stober, S., Stock, A.-K. & Beste, C. Applying deep learning to single-trial EEG data provides evidence for complementary theories on action control. *Communications Biology* **3**, 112 (2020).
150. Bae, G.-Y., Leonard, C. J., Hahn, B., Gold, J. M. & Luck, S. J. Assessing the information content of ERP signals in schizophrenia using multivariate decoding methods. *NeuroImage: Clinical* **25**, 102179 (2020).
151. Kunert-Graf, J. M., Eschenburg, K. M., Galas, D. J., Kutz, J. N., Rane, S. D. & Brunton, B. W. Extracting Reproducible Time-Resolved Resting State Networks Using Dynamic Mode Decomposition. *Frontiers in Computational Neuroscience* **13**, 75 (2019).
152. Shiraishi, Y., Kawahara, Y., Yamashita, O., Fukuma, R., Yamamoto, S., Saitoh, Y., Kishima, H. & Yanagisawa, T. Neural decoding of electrocorticographic signals using dynamic mode decomposition. *Journal of Neural Engineering* **17**, 036009 (2020).

153. Cassady, K., Gagnon, H., Freiburger, E., Lalwani, P., Simmonite, M., Park, D. C., Peltier, S. J., Taylor, S. F., Weissman, D. H., Seidler, R. D. & Polk, T. A. Network segregation varies with neural distinctiveness in sensorimotor cortex. *NeuroImage* **212**, 116663 (2020).
154. Song, J., Davey, C., Poulsen, C., Luu, P., Turovets, S., anderson, E., Li, K. & Tucker, D. EEG source localization: Sensor density and head surface coverage. *Journal of Neuroscience Methods* **256**, 9–21 (2015).
155. Lantz, G., Grave de Peralta, R., Spinelli, L., Seeck, M. & Michel, C. Epileptic source localization with high density EEG: how many electrodes are needed? *Clinical Neurophysiology* **114**, 63–69 (2003).
156. Raz, N. & Rodriguez, K. M. Differential aging of the brain: Patterns, cognitive correlates and modifiers. *Neuroscience & Biobehavioral Reviews* **30**, 730–748 (2006).
157. Pichot, R. E., Henreckson, D. J., Foley, M. & Koen, J. D. Neural noise is associated with age-related neural dedifferentiation. *bioRxiv* (2022).
158. Nogueira, J., Gerardo, B., Santana, I., Simões, M. R. & Freitas, S. The Assessment of Cognitive Reserve: A Systematic Review of the Most Used Quantitative Measurement Methods of Cognitive Reserve for Aging. *Frontiers in Psychology* **13**, 847186 (2022).
159. Tetereva, A. & Pat, N. The (Limited?) Utility of Brain Age as a Biomarker for Capturing Cognitive Decline (2023).

Published Research Articles

Published Research Article I



Classification of visuomotor tasks based on electroencephalographic data depends on age-related differences in brain activity patterns

C. Goelz^a, K. Mora^{b,c}, J. Rudisch^d, R. Gaidai^a, E. Reuter^e, B. Godde^f, C. Reinsberger^a, C. Voelcker-Rehage^{d,g}, S. Vieluf^{a,h,*}

^a Institute of Sports Medicine, Paderborn University, Paderborn, Germany

^b Remote Sensing Centre for Earth System Research, Leipzig University, Leipzig, Germany

^c German Centre for Integrative Biodiversity Research (iDiv), Leipzig, Germany

^d Department of Neuromotor Behavior and Exercise, Institute of Sport and Exercise Sciences, University of Münster, Münster, Germany

^e Department of Sport and Health Sciences, Technical University of Munich, Munich, Germany

^f Department of Psychology & Methods, Jacobs University Bremen, Bremen, Germany

^g Institute of Human Movement Science and Health, Chemnitz University of Technology, Chemnitz, Germany

^h Division of Epilepsy and Clinical Neurophysiology, Department of Neurology, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA



ARTICLE INFO

Article history:

Received 16 October 2020

Received in revised form 12 March 2021

Accepted 22 April 2021

Available online 13 May 2021

Keywords:

Dynamic mode decomposition
Fine motor control
Machine learning
Decoding
Dedifferentiation
Aging

ABSTRACT

Classification of physiological data provides a data driven approach to study central aspects of motor control, which changes with age. To implement such results in real-life applications for elderly it is important to identify age-specific characteristics of movement classification. We compared task-classification based on EEG derived activity patterns related to brain network characteristics between older and younger adults performing force tracking with two task characteristics (sinusoidal; constant) with the right or left hand. We extracted brain network patterns with dynamic mode decomposition (DMD) and classified the tasks on an individual level using linear discriminant analysis (LDA). Next, we compared the models' performance between the groups. Studying brain activity patterns, we identified signatures of altered motor network function reflecting dedifferentiated and compensational brain activation in older adults. We found that the classification performance of the body side was lower in older adults. However, classification performance with respect to task characteristics was better in older adults. This may indicate a higher susceptibility of brain network mechanisms to task difficulty in elderly. Signatures of dedifferentiation and compensation refer to an age-related reorganization of functional brain networks, which suggests that classification of visuomotor tracking tasks is influenced by age-specific characteristics of brain activity patterns. In addition to insights into central aspects of fine motor control, the results presented here are relevant in application-oriented areas such as brain computer interfaces.

© 2021 Elsevier Ltd. All rights reserved.

1. Introduction

Decoding and classification of motor behavior based on EEG are elementary parts of EEG analysis and an essential basis for areas of application including, for example, brain computer interfaces (BCI) for persons with motor disabilities. BCI systems, designed to recognize movement specific brain states, require the identification of segregated functional brain networks and processes (Kasahara, DaSalla, Honda, & Hanakawa, 2015; Saha & Baumert, 2020). Considering the clinical application areas of BCIs, for example, in the field of stroke rehabilitation (Ramos-Murguiad et al., 2013) or in prevention and training, e.g. in

form of neurofeedback (Gomez-Pilar, Corralejo, Luis, Alvarez, & Hornero, 2016), older adults form a main target group. More generally, the development of brain networks is an important factor which influences classification performance (Ahn & Jun, 2015). Although inter- and intrasubject variability are studied in recent research in the field of BCIs (Sannelli, Vidaurre, Müller, & Blankertz, 2019), the influence of age-related changes on classification results has been rarely investigated (Ahn & Jun, 2015). Chen, Fu, Boger, and Jiang (2019) found a 15.9% lower accuracy of the classification of body side when the left- and the right-hand were passively stimulated in older adults compared to younger adults. The authors suspect an association with altered, less lateralized, spatial patterns of EEG signals in the elderly compared to younger participants. Beyond that, the authors highlight the relevance of this finding, as lateralization of brain activity patterns based on the contralateral organization of the motor system

* Correspondence to: Paderborn University, Institute of Sports Medicine, Warburger Str. 100, 33098 Paderborn, Germany.

E-mail address: vieluf@sportmed.upb.de (S. Vieluf).

is often exploited in BCI algorithms. However, less is known about whether the results can be confirmed for active motor tasks that could allow a transfer to stroke rehabilitation for example. Furthermore, it has not yet been investigated whether the described changes also apply to other task dimensions. A thorough understanding of this would be a prerequisite for real-life applications.

In addition to the practical relevance of investigating the influence of age-related changes on the classification performance, the underlying machine learning models offer the possibility of a data-driven insight into the aging brain (Cichy & Pantazis, 2017; Schirmeister et al., 2017). With increasing age, brain activity is less differentiated and noisier (Reuter-Lorenz & Lustig, 2005), i.e., brain networks are less task specific (Carp, Park, Hebrank, Park, & Polk, 2011). This is summarized under the term dedifferentiation (Reuter-Lorenz & Lustig, 2005). Therefore, it is reasonable to assume that decoding of neural signals might be less effective in older adults (Roland et al., 2011). Age-related changes include a reduction of distinctness and segregation in the functional organization of large-scale brain networks (Cassady, Ruitenberg, Reuter-Lorenz, Tommerdahl, & Seidler, 2020; Salallonch, Bartrés-Faz, & Junqué, 2015). Using fMRI Seidler et al. (2010) showed functional brain activation to be more widespread and bilateral during visual and cognitive tasks. Several authors report similar results during motor tasks where older compared to younger participants showed stronger activation of the ipsilateral motor cortex (Carp et al., 2011; Ward & Frackowiak, 2003) as well as higher connectivity within sensorimotor and visual networks during visuomotor tapping (Cassady et al., 2020). Moreover, older adults seem to recruit additional sensory as well as prefrontal areas during motor tasks which has been interpreted as compensational mechanisms and additional recruitment of cognitive resources (Berghuis et al., 2019; Reuter-Lorenz & Cappell, 2008). Compensation hence implies that older adults use different neural circuits than young adults when performing a motor task (Cabeza, Anderson, Locantore, & McIntosh, 2002; Davis, Dennis, Daselaar, Fleck, & Cabeza, 2008; Heuninckx, Wenderoth, Debaere, Peeters, & Swinnen, 2005).

These changes are noticeable in altered spatial patterns and spectral composition of EEG signals (Al Zoubi et al., 2018) and thus influence the requirements for discriminability of task-relevant brain network processes for classification, as previously shown by Chen et al. (2019) for one task dimension, i.e., body side of execution. Expanding on these findings, the differentiability of electrophysiological brain network patterns of different task dimensions could provide insight into age-related alterations described above.

One possibility to study these patterns is offered by dynamic mode decomposition (DMD). This method, which originated in the field of fluid dynamics, is capable of detecting spatio-temporal coherent patterns from high dimensional data and has recently been applied in several neuroscientific settings (Brunton, Johnson, Ojemann, & Kutz, 2016; Gölz et al., 2018; Vieluf et al., 2018). DMD provides a framework to deduce brain activity patterns that relate to brain network characteristics, i.e., coherent dynamics of distinct sources (Fries, 2005). It allows to assess the pattern dynamics in terms of neuronal oscillations as markers of neuronal network dynamics by constructing a best fit linear dynamical system directly from the measurement data (Brunton et al., 2016). In a previous study, we were able to show age-related differences in EEG derived central control processes during a force maintenance task utilizing DMD (Vieluf et al., 2018). More specifically, we found that in older adults compared to younger adults' brain activity was less differentiated between left- and right-hand tasks. Furthermore, we could detect signs of a less segregated sensorimotor network, higher internal network communication

as well as additional cognitive recruitment during force control in older adults.

To better understand these alterations and draw conclusions about their relevance in practical settings, we aimed to study the classification of fine motor movements in four visuomotor tracking tasks based on EEG recordings recorded in a follow-up experiment of Vieluf et al. (2018). First, by comparing electro-physiological differences between younger and older adults in relation to different tasks, we aimed to confirm previously reported age-related differences in EEG activity associated with visuomotor tasks. We expected that older adults have less differentiated and segregated networks, similar to our previous findings (Vieluf et al., 2018). Second, we investigated classification differences between age-groups. We expected to gain insight into the age-related differences in functional brain activity, such as higher bilateral brain activation, compensatory involvement and additional attentional resources by studying the discriminative performance in relation to the different task dimensions. Our findings provide further insights into age-related alterations of movement-related brain networks and show the practical relevance these alterations might have.

2. Methods

2.1. Participants

Twenty-six participants were included in this study. All participants enrolled voluntarily and gave their informed consent after recruitment through flyers, telephone calls, and newspaper announcements within the Bremen-Hand-Study@Jacobs (Voelcker-Rehage, Reuter, Vieluf, & Godde, 2013). A questionnaire was used to collect information about the demographic background and health status of all participants. All subjects described themselves as healthy and free of neurological diseases and had normal or corrected to normal hearing and vision. The Edinburgh Handedness Inventory (Oldfield, 1971) further identified all participants as right-handed. Participants received a compensation of 8 Euros per hour. The study was in accordance with the Declaration of Helsinki and approved by the ethics committee of the German Psychological Society. Based on their age, participants were assigned to a young adults (YA: N = 13, age = 18–25 years, 8 females, 5 males) or an older adults (OA: N = 13, age = 55–65 years, 8 females, 5 males) group.

2.2. Experimental procedure

Within the framework of the Bremen-Hand-Study@Jacobs (Voelcker-Rehage et al., 2013), participants took part in several experiments. The visuomotor force-control experiment analyzed here was designed as a force tracking experiment conducted blockwise. Participants sat about 80 cm in front of a computer screen (19", frame rate 60 Hz). Their arms rested on arm pads. Thumb and index finger grasped a force transducer (model Mini-40, ATI Industrial Automation, Garner, NC, United States) in a pinch grip. Using the right- or left-hand, the task was to follow a target force level for five seconds by applying the required amount of force to the force transducer. The target line was either presented as a constant level (steady), i.e., a straight line, corresponding to 2 N or as a sinusoidal curve (sine) ranging from 2 N to 12 N with a frequency of 1 Hz on the y-axis whereas time was presented on the x-axis. The time axis (x-axis) covered 5 s, allowing participants to see one second of the upcoming target line and 4 s of the preceding target line and the applied force. A total of 160 trials were carried out, each lasting 5 s with an individual break of 5 s to 7 s during which participants were instructed to focus on a fixation cross on the screen in front of

them. Initially, 80 trials were performed with the right hand. The first 40 trials involved the steady force level and the following 40 trials the sine force level. The sequence was then repeated with the left hand. The participants had an individual break between each task block. Prior to the experiment, the maximum voluntary contraction (MVC) was recorded with three maximum precision grip trials. Each grip lasted 5 s with approximately 2 min break in between.

The experiment and grip force acquisition were realized using customized LabVIEW (National Instruments, Austin) software. Grip force was recorded with 120 Hz sampling rate and amplitude resolution of 0.06 N via the force transducer.

EEG was recorded with 32 active electrodes (ActiveTwo, BioSemi, Amsterdam, Netherlands) placed on the scalp according to the international 10–20 System. Ocular artifacts as well as mastoid potential were recorded with additionally placed electrodes. Common Mode Sense (CMS) and Driven Right Leg (DRL) electrodes were placed next to Cz. All EEG signals were recorded with a sampling rate of 2048 Hz applying an online filter between 0.16 and 100 Hz. Prior to the experiments resting EEGs with eyes open and eyes closed were recorded for 30 s each while participants sat comfortably on a chair.

2.3. Data analysis

Data analysis was conducted using Python (version 3.7.6). For EEG data analysis and classification the additional MNE package (version 0.2.3 Gramfort et al., 2013) and scikit-learn (version 0.22.1, Pedregosa et al., 2011) were used. All data was analyzed on an individual level and the classifiers were trained individually for each participant. The procedure for individual EEG classification is outlined in Fig. 1 and described in detail in the following subsections.

2.3.1. Force data

Force data was analyzed to identify incorrect task execution using the central 3 s of each trial. The force signal was filtered using a 4th order Butterworth filter at 30 Hz. To evaluate the performance of each participant and test for correct task execution, we calculated the mean absolute deviation from the target force as an indicator of accuracy. Trials were excluded if their z-score was above three. Results are available in the supplementary material (Figure S1).

2.3.2. EEG data

EEG data was resampled to 200 Hz and re-referenced to the linked mastoids. Signals were bandpass filtered between 4 and 30 Hz using a FIR zero phase filter created with MNE python. To correct ocular artifacts caused by blinking and eye movements we applied ICA with the FastICA algorithm (Hyvärinen & Oja, 1997). When the source signals of a component correlated with the ocular signals, this component was automatically marked and set to zero before back projection. Next data was cut between 1 s and 4 s so that the middle 3 s of each trial were used for the analysis. Trials containing artifacts such as electrical or myogenic artifacts were removed based on the Autoreject pipeline (Jas, Engemann, Bekhti, Raimondo, & Gramfort, 2017).

2.3.2.1. Dynamic mode decomposition. Each artifact free trial was mean centered and windows of 0.5 s (100 datapoints) were decomposed with DMD. Consecutive windows have an overlap of 0.25 s (50 datapoints) to account for data non-stationarities. The exact DMD algorithm proposed by Tu, Rowley, Luchtenburg, Brunton, and Kutz (2014) and used by Brunton et al. (2016) was applied in this work.

Basically, DMD is an algorithm that allows to approximate the relation of all signals in pairs of consecutive time instances

of a measurement $X \in \mathbb{R}^{n \times m}$, where n denotes sensors and m denotes measurement points at time t . With DMD it is possible to decompose X to extract spatio-temporal patterns of coherent brain activation (Brunton et al., 2016). To increase the number of calculated DMD modes and thus improve the approximation accuracy we used the delay embedding technique as applied by Brunton et al. (2016); stacking depth $h = 4$ was selected. Since the DMD algorithm takes advantage of the Singular Value Decomposition (SVD), the optimal rank r of the input matrix and therefore the number of SVD modes was determined using the method of Donoho and Gavish (2014). The decomposed spatial patterns are organized in a matrix $\Phi \in \mathbb{C}^{hn \times r}$, containing all DMD modes (i.e., spatial patterns) and a diagonal matrix $\Lambda \in \mathbb{C}^{r \times r}$, containing the corresponding DMD eigenvalues. From this, the original measurement X can be approximated by the model:

$$\hat{X} = \Phi \exp(\Omega t) z,$$

where Ω contains the continuous eigenvalues transformed with $\Omega = \log(\Lambda)/\Delta t$.

The variable t denotes time whereby $\Delta t = 0.005$ s and z is computed from the first data point x_1 of X with $x_1 = \Phi z$. From Ω we calculated the frequencies in cycles per second (Hz)

$$f = |\text{imag}(\Omega)|/2\pi$$

for all DMD modes.

The optimal stacking depth h as well as window size and overlap were determined by an error analysis of the data of five randomly selected subjects. As unique spatial information is present in the first n rows of a mode, we used $\Phi \in \mathbb{C}^{n \times r}$.

2.3.2.2. Feature extraction. For feature extraction we modified and applied an approach based on Filter Bank Common Spatial Pattern (FBCSP; Ang, Chin, Wang, Guan, & Zhang, 2012). FBCSP is a widely used and very successful algorithm for classifying motion-related EEG signals. In essence, linear combinations of EEG channels are used to calculate weightings (spatial filters) that maximize the class discriminative energy for selected frequency bands. The spatial filters in each frequency band of interest are usually calculated based on the Common Spatial Pattern (CSP) algorithm. The rationale behind this was to extract the information from the DMD modes of all windows that would allow the best possible differentiation between the tasks. In this way it is possible to relate the results of the classification to differences in brain network activity.

To estimate spatial filters, we first extracted the absolute DMD values associated with the θ - (4 to < 7 Hz), α - (7 to < 12 Hz), β_1 - (12 to < 16 Hz) and β_2 - (16 to < 30 Hz) frequencies creating a mode matrix Φ_b for each frequency band b . For each frequency band we extracted the spatial filters from the training data with the CSP algorithm. More precisely, we calculated the transformation matrix $W_b \in \mathbb{R}^{n \times n}$ by solving the generalized eigenvalue problem:

$$S_{b,1} W_b = (S_{b,1} + S_{b,2}) W_b D_b.$$

$S_{b,1}$ and $S_{b,2}$ represent estimates of the covariance matrices of the b -th mode matrix Φ_b of the respective task and D_b contains the eigenvalues of $S_{b,1}$ on its diagonal (Ang et al., 2012). Each column vector w_{bj} is called spatial filter. For each frequency band we extracted the two most informative spatial filters w_{bj} , i.e., columns of W_b forming the transformation matrix \bar{W}_b as described in Barachant, Bonnet, Congedo, and Jutten (2010). This number was chosen to avoid overfitting and is based on preliminary tests that showed no significant improvement in classifier performance when more filters were selected. The MNE Python implementation of the CSP algorithm modified for our purposes was used here. Multiclass CSP is implemented from Grosse-Wentrup and Buss (2008).

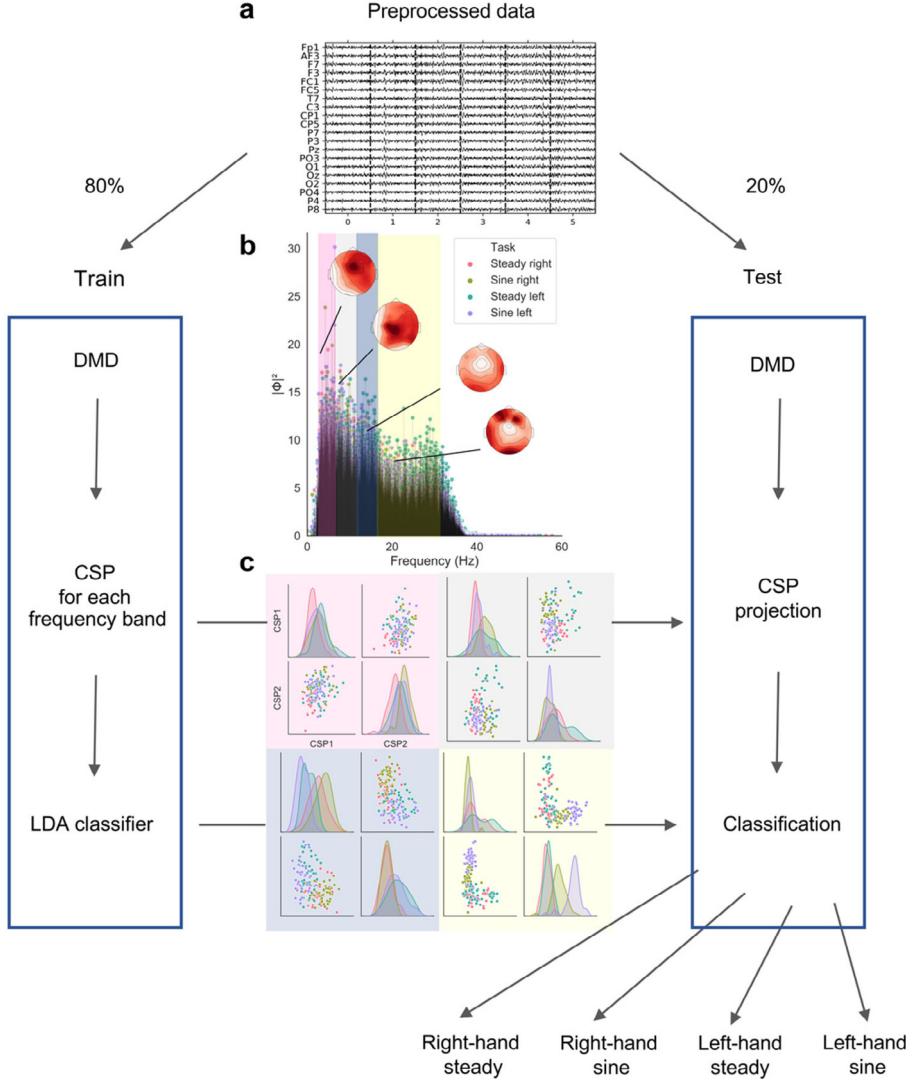


Fig. 1. Individual classification approach. For each participant preprocessed data epochs (a) are divided in training and test data set. Each set was next decomposed by DMD (b: DMD spectrum and corresponding spatial pattern). DMD mode magnitudes corresponding to the θ , α , β_1 or β_2 frequency bands were spatially filtered and projected into CSP Space (c). A LDA classifier was trained on the training data and applied to the test data set.

2.3.2.3. Classification. After preprocessing, all trials were assigned to either a training or test data set. The ratio of this allocation comprised 80% training (i.e., 32 trials per class) and 20% test (i.e., 8 trials per class). Subsequently, all trials were segmented in overlapping windows and decomposed using DMD as described above. For task classification a Linear Discriminative Analysis (LDA) classifier with shrinkage was trained and tested for each participant using the scikit-learn toolbox implemented in Python (Pedregosa et al., 2011). As features the DMD mode magnitudes in each frequency band b and trial i , $\Phi_{b,i}$, were projected into CSP space with the transformation matrix \bar{W}_b , which was extracted from the training data. Next, the logarithmic variance of the projected DMD mode magnitudes of all time windows was calculated and used as features as described in Barachant et al. (2010). Thus, for each trial and frequency band the CSP feature $v_{b,i}$ was calculated with

$$v_{b,i} = \log(\text{Var}(\bar{W}_b^T \Phi_{b,i})),$$

using training and test data, respectively. The CSP features $v_{b,i}$ are then collected in the matrix $V \in \mathbb{R}^{t \times 2f}$ where t denotes the total number of trials and f the total number of frequency bands used (see Fig. 1 for a visual presentation of the procedure). Note that

$2f$ is required to reflect the frequency bands from both training and test data. To (1) chose the best hyperparameter of the LDA classifier (solver, number of components) and to (2) determine the metrics of the classifier, 10×10 fold cross-validation was applied. More precisely the procedure was run in two loops. In the outer loop, the data set was split into training and test data. In a second inner loop, the best hyperparameters of the classification algorithm were determined using 10-fold cross-validation. This procedure was repeated 10 times. In each fold, the classification performance was evaluated based on accuracy, F1 value, and receiver operating characteristics curve (ROC). Accuracy indicates the absolute percentage of all correctly classified trials, whereas the F1 value is the weighted average of the precision and recall. To evaluate the classifier performance in terms of discriminative ability we further studied the ROC (see Fig. 6) and used the area under the curve as metric (ROC AUC). As ROC is restricted to binary classification problems, we relied on the one vs. all procedure implemented in scikit-learn to calculate the ROC AUC for the 4-class classification, thus restricted Fig. 6 to the two-class cases.

Classification was carried out at different levels. First, a four-class classification was carried out. All trials of the four different tasks were included. Then, two-class classifications were

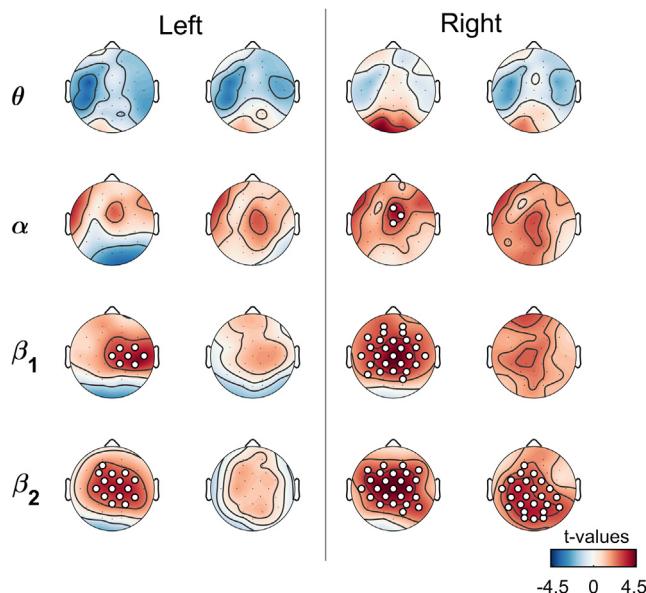


Fig. 2. Comparison of DMD mean mode magnitude between groups in left- and right-hand sine and steady tasks displayed as statistical t-maps (red: OA > YA, blue: OA < YA). Significant differences are marked with white circles (FDR corrected p -value < 0.05). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

performed to compare the classifier performance for each task dimension separately. With this approach we intended to study all levels of classification and intend to increase comparability to two-class approaches reported in the literature (Chen et al., 2019). First all left- and right-handed trials were classified as one class regardless of whether it was a sine or steady tracking task. Next, all sine or steady-tracking trials, regardless of whether the trial was performed with the left- or right-hand, were used as one class each. In order to ensure the comparability of this classification step, we selected 20 trials of each class randomly in each fold. Thus, 40 trials each were the basis for the left- and right-hand as well as sine and steady task classifications. Lastly, the classification was performed for the following combinations: left-hand sine vs. left-hand steady, right-hand sine vs. right-hand steady, left-hand sine vs. right-hand sine, left-hand steady vs. right-hand steady.

2.4. Statistics

Statistical Analysis was conducted using SPSS for MAC version 24.0 (IBM Corp., Armonk, NY, USA) and MNE Python (version 0.2.3 Gramfort et al., 2013). Following Brunton et al. (2016) and Vieluf et al. (2018) the DMD mean mode magnitudes were calculated per frequency band and task for statistical comparison. For this, DMD modes of all windows and trials associated with the θ - (4 to < 7 Hz), α - (7 to < 12 Hz), β_1 - (12 to < 16 Hz) and β_2 - (16 to < 30 Hz) frequencies per task and per participant obtaining DMD mean mode magnitudes per task per frequency band for each participant. To describe the input feature basis prior to CSP feature extraction and LDA classification we compared the DMD mean mode magnitudes between groups and between tasks. We used permutation t-tests to account for the high dimensionality of the EEG data and therefore the inaccuracies in test assumption requirements and reduction of type I and type II errors (Maris & Oostenveld, 2007). P-values were corrected with false discovery rate (Benjamini & Hochberg, 1995). To test for group differences in laterality we calculated the difference between C3 and C4. Moreover, to test for additional frontal activation we calculated

the difference between Pz and Fz. Values below 0 indicate higher DMD mean mode magnitudes on the central right or frontal electrode, respectively. Values above 0 indicate higher DMD mean mode magnitudes on the central left or parietal side. We analyzed these indices each with repeated measures analysis of variance (ANOVA) with the between factor group (2; YA and OA) and the within factor task (4; right-hand steady, right-hand sine, left-hand steady, left-hand sine) for each frequency band (θ , α , β_1 , β_2). Significant interactions and main effects were followed by Bonferroni corrected pairwise comparisons. In case of violation of the assumption of sphericity we use the Greenhouse-Geisser correction and report corrected degrees of freedom and p-values. The effect sizes are reported as η_p^2 .

We compared the classifier performance measures (accuracy, ROC AUC and F1) between groups with Mann-Whitney U test due to violations of normality assumption. Effect sizes are reported as r-values. All statistical tests were conducted on significance level set to $\alpha = 0.05$.

3. Results

3.1. Group and task comparison of DMD mean modes

The group comparisons of DMD mean mode magnitudes are illustrated in Fig. 2 as topographic t-maps. Significant differences, obtained with permutation t-tests, are marked as white circles ($p < 0.05$). For a better understanding of the classifier input and to maintain a direct link to the classification, here we report absolute DMD patterns (note that this procedure is different than in Vieluf et al., 2018 where task related DMD values are reported). For the left-hand steady tasks, we found significant differences between the groups on central electrodes in the β_1 and the β_2 frequency band, with higher DMD mean mode magnitudes for the OA group. Group comparison of the right-hand tasks revealed significant higher DMD mean mode magnitudes in the α , β_1 and β_2 band over central, frontal and parietal electrodes for the steady task. Significant group differences for the right-hand sine task were observed for central, parietal, and occipital electrodes in the β_2 range.

Comparisons of left- and right-hand sine as well as left- and right-hand steady tasks are illustrated in Fig. 2 as topographic t-maps for the OA and YA groups separately. Comparing left- and right-hand tasks, we found significant task differences over occipital electrodes in the θ -band for the steady task in the OA group. In the YA group, significant differences between the left and right-hand sine tasks were evident over central, parietal and occipital electrodes in the β bands localized on the left side. Regarding the comparison of sine and steady tasks we found differences in occipital electrodes which were present for the right-hand task in the YA group and in the left-hand task in the OA group (all $p < 0.05$).

Comparing C3C4 indices with repeated measures ANOVA revealed a significant interaction between task and group in the β_1 band [$F(1.96, 47.193) = 3.214, p = 0.05, \eta_p^2 = 0.118$] and the β_2 band [$F(2.42, 58.166) = 3.22, p = 0.038, \eta_p^2 = 0.118$]. Post-hoc comparison revealed significant differences between YA and OA in the left-hand steady task for the β_1 frequency band with lower values for OA. In the β_2 range in the right-hand sine task, the OA showed higher values (all $p < 0.05$; see Fig. 4a).

For PzFz indices the task-by-group interaction was significant for the α band [$F(1.75, 42.11) = 8.34, p = 0.001, \eta_p^2 = 0.258$]. Post-hoc tests revealed a significant difference between the groups in the left-hand steady task ($p = 0.005$) with higher values for the YA.

In the θ and β_1 bands there was a significant main effect of task for the PzFz indices [θ : $F(3, 72) = 6.06, p = 0.001, \eta_p^2 =$

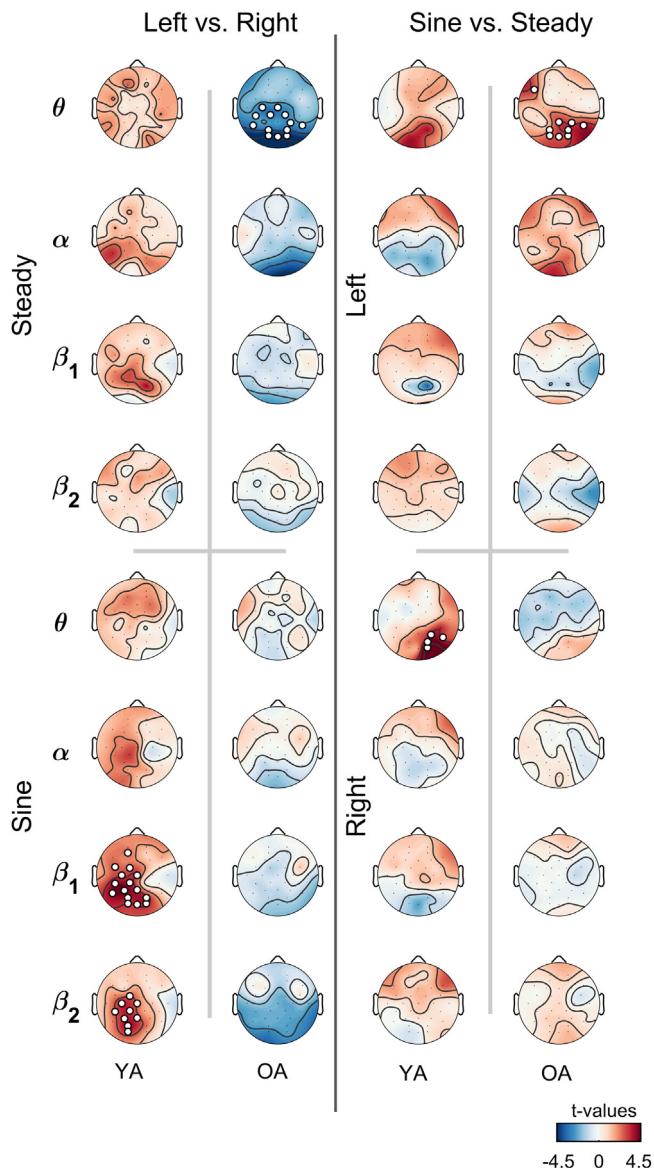


Fig. 3. Comparisons of DMD mean mode magnitudes between left- and right-hand as well as sine and steady tasks within groups displayed as statistical t-maps. Significant differences are marked with white circles (FDR corrected p -value < 0.05). Red: Right > Left/Steady > Sine, Blue: Right < Left/Steady < Sine. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

0.394 ; β_1 : $F(3, 72) = 9.156$, $p = 0.006$, $\eta_p^2 = 0.447$], but no main effect of group [θ : $F(1, 24) = 0.55$, $p = 0.47$, $\eta_p^2 = 0.018$; β_1 : $F(1, 24) = 0.159$, $p = 0.694$, $\eta_p^2 = 0.007$]. Pairwise comparisons revealed significant differences between the right-hand steady, right-hand sine and left-hand sine tasks in the θ band with lower values in the left-hand steady task. In the β_1 band participants had higher values in the left-hand steady task compared to both hands sine tasks and lower values in the left-hand sine compared to the right-hand sine task (all $p < 0.05$; see Fig. 4b).

In summary, we found group differences of spatio-temporal coherent patterns in all tasks involving central electrodes. By comparing the tasks, these patterns revealed higher task specificity. This was indicated by fewer significant differences, especially between the sine tasks which were more pronounced in the YA. In addition, we found selective differences between OA

and YA of the coherent patterns with regard to laterality and frontality.

3.2. Evaluation of classifier performance

All classifiers performed significantly above chance level for both groups ($p < 0.05$). Regarding the group comparison, the four-class classification revealed no significant difference between groups in all performance metrics (accuracy: $U = 54.5$, $p = 0.125$, $r = 0.30$; F1: $U = 57.00$, $p = 0.169$, $r = 0.28$; ROC AUC: $U: 57.00$, $p = 0.169$, $r = 0.28$). However, visualizing the classifiers performance in an error or a confusion matrix (Fig. 5), on a descriptive level, classification accuracy was higher in the group of YA for the dimension body side (9.75 % wrongly classified compared to 18 % in the OA group). Conversely, accuracy seemed to be reduced in the group of YA for the dimension task characteristic (14.25 % wrongly classified compared to 10.5 % in the OA group, see confusion matrix Fig. 5).

Following this observation, we compared the classifier performance of left-hand vs. right-hand and sine vs. steady task classification separately. Furthermore, to check whether the differences are differently expressed between the body side and task characteristics levels we considered all the binary classifications as a follow up analysis. The performance of all two-class classifiers at both levels is shown as ROC curves in Fig. 6. All classification results are shown in Tables 1 and 2 for each participant. Mann–Whitney U test confirmed significantly higher accuracy, F1 and ROC AUC scores for the left-hand vs. right-hand classification in the YA compared to the OA group (accuracy: $U = 39.5$, $p = 0.019$, $r = 0.45$; F1: $U = 40.00$, $p = 0.022$, $r = 0.45$; ROC AUC: $U = 43.00$, $p = 0.034$, $r = 0.42$). For the sine vs. steady classification Mann–Whitney U test confirmed significantly lower accuracy, F1 and ROC AUC scores for the YA compared to the OA group (accuracy: $U = 40.00$, $p = 0.022$, $r = 0.45$; F1: $U = 41.00$, $p = 0.026$, $r = 0.44$; ROC AUC: $U = 37.00$, $p = 0.014$, $r = 0.48$) (see Fig. 6a and b top panel ROC curves).

Next, we compared the classification of each task. There were significant differences between the groups only for the left-hand sine vs. right-hand sine task with higher scores in the YA group (left-hand vs. right-hand sine: accuracy: $U = 29.00$, $p = 0.003$, $r = 0.56$; F1: $U = 29.00$, $p = 0.003$, $r = 0.56$; ROC AUC: $U = 41.00$, $p = 0.026$, $r = 0.44$; left-hand vs. right-hand steady: accuracy: $U = 69.00$, $p = 0.448$, $r = 0.15$; F1: $U = 71.00$, $p = 0.511$, $r = 0.14$; ROC AUC: $U = 71.00$, $p = 0.511$, $r = 0.14$). The follow up analysis of the individual comparisons sine vs. steady revealed no significant difference between the groups for the right- and left-hand sine vs. steady classification separately (left-hand sine vs steady: accuracy: $U = 59.5$, $p = 0.204$, $r = 0.25$; F1: $U = 59.00$, $p = 0.204$, $r = 0.26$; ROC AUC: $U = 53.00$, $p = 0.113$, $r = 0.32$; right-hand sine vs steady: accuracy: $U = 57.00$, $p = 0.169$, $r = 0.28$; F1: $U = 57.00$, $p = 0.169$, $r = 0.28$; ROC AUC: $U = 52.00$, $p = 0.101$, $r = 0.33$) (see Fig. 6c).

In essence, the four-class classification showed comparable performance in both groups. However, the investigation of false positive and negative rates revealed differences between age groups. These could be confirmed by examining the binary classification. While the classification of the body side performed better in YA, the classification of task characteristics was better in OA.

4. Discussion

In the current study, we aimed to explore differences in spatio-temporal coherent patterns related to brain network characteristics relevant for the classification of visuomotor tracking tasks based on EEG recordings between YA and OA. Following up on

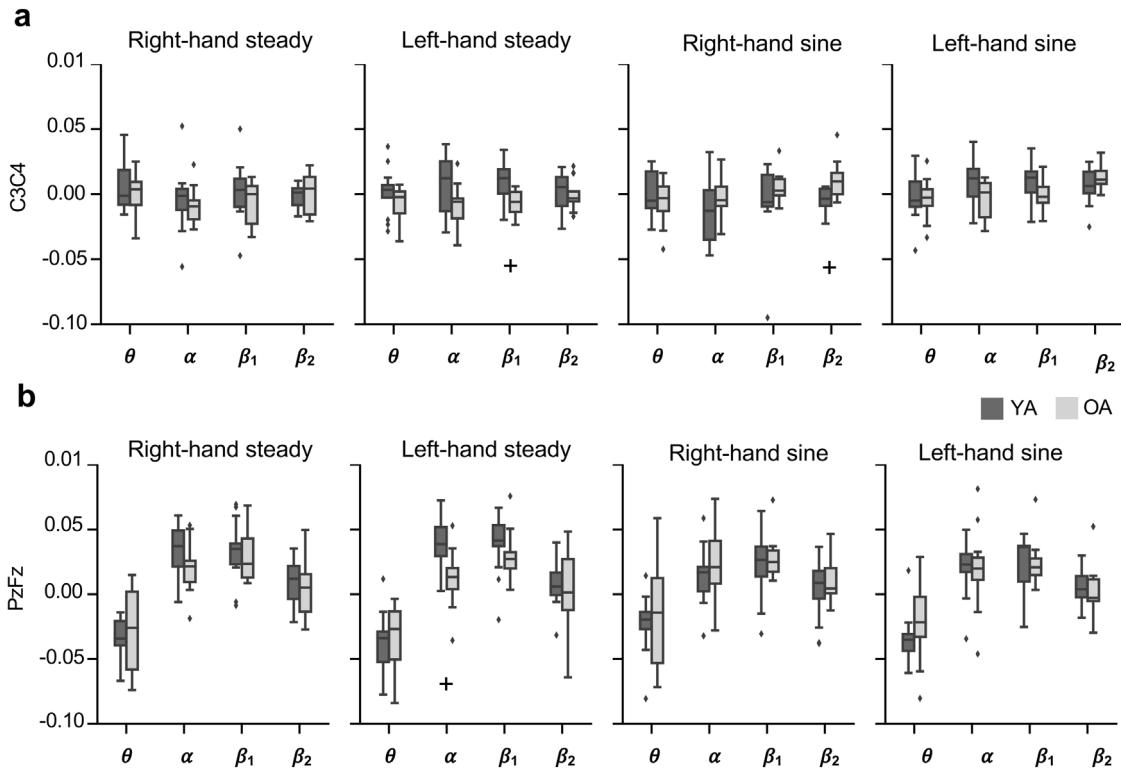


Fig. 4. Laterality and frontality indices calculated based on the DMD mean mode magnitudes. C3C4 indices (a) and PzFz (b) indices for each frequency band and task. Significant differences between the groups are marked with +.

Table 1
First level: Classification performance of 10×10 fold cross validation.

	4 class			sine vs. steady			left vs. right		
	Accuracy	ROC AUC	F1	Accuracy	ROC AUC	F1	Accuracy	ROC AUC	F1
OA1	0.85	0.97	0.84	1.00	1.00	1.00	0.80	0.88	0.80
OA2	0.59	0.85	0.58	0.87	0.95	0.87	0.69	0.80	0.68
OA3	0.65	0.85	0.65	0.83	0.90	0.83	0.68	0.80	0.66
OA4	0.55	0.78	0.55	0.71	0.78	0.70	0.75	0.83	0.75
OA5	0.71	0.92	0.71	0.89	0.96	0.89	0.74	0.85	0.73
OA6	0.47	0.74	0.46	0.71	0.76	0.70	0.74	0.81	0.73
OA7	0.72	0.92	0.72	0.88	0.96	0.87	0.73	0.79	0.72
OA8	0.63	0.86	0.63	0.79	0.86	0.79	0.72	0.77	0.71
OA9	0.53	0.76	0.52	0.79	0.87	0.78	0.63	0.67	0.63
OA10	0.65	0.86	0.64	0.91	0.97	0.91	0.63	0.72	0.61
OA11	0.76	0.92	0.75	0.88	0.93	0.87	0.75	0.84	0.73
OA12	0.48	0.74	0.47	0.89	0.96	0.89	0.57	0.58	0.56
OA13	0.64	0.87	0.63	0.97	1.00	0.97	0.69	0.77	0.68
YA1	0.57	0.84	0.56	0.69	0.78	0.68	0.90	0.95	0.90
YA2	0.73	0.90	0.72	0.66	0.73	0.66	0.91	0.96	0.91
YA3	0.71	0.89	0.70	0.71	0.78	0.71	0.68	0.75	0.67
YA4	0.67	0.89	0.66	0.75	0.82	0.75	0.91	0.98	0.91
YA5	0.78	0.94	0.77	0.66	0.75	0.66	0.95	0.98	0.95
YA6	0.85	0.97	0.84	0.87	0.94	0.87	0.99	1.00	0.99
YA7	0.76	0.93	0.76	0.83	0.93	0.83	0.77	0.88	0.77
YA8	0.54	0.79	0.53	0.61	0.69	0.61	0.70	0.76	0.69
YA9	0.56	0.78	0.54	0.79	0.82	0.79	0.59	0.58	0.58
YA10	0.63	0.83	0.62	0.60	0.69	0.59	0.64	0.70	0.62
YA11	0.73	0.94	0.73	0.77	0.84	0.76	0.93	0.97	0.92
YA12	0.69	0.90	0.68	0.91	0.94	0.91	0.87	0.92	0.87
YA13	0.82	0.94	0.82	0.94	0.98	0.94	0.83	0.90	0.83

Vieluf et al. (2018), we studied differences in the extracted patterns between groups in four visuomotor tracking task (sine and steady force tracking executed with the left and right hand) to get insights into which brain activity patterns might influence the classification.

Group differences in all tasks were most pronounced for central electrodes, especially in the β_2 band. Between groups, we found task differences when comparing left- and right-hand sine

tasks in frontal and central electrodes in the θ and α bands. Assuming additional bilateral and frontal recruitment for older adults we calculated central laterality (C3C4) and fronto-parietal indices (PzFz) and found significant group differences. Especially the central laterality index as well as the gradient between frontal and parietal of the left-hand steady tasks in the α band was lower in OA. Moreover, we found differences in the central laterality index of the right-hand sine task in β_2 . Taken together, these

Table 2
Second level: Classification performance of 10×10 fold cross validation.

	sine: left vs. right			steady: left vs. right			left: sine vs. steady			right: sine vs. steady		
	Accuracy	ROC	AUC	F1	Accuracy	ROC	AUC	F1	Accuracy	ROC	AUC	F1
OA1	0.85	0.93	0.85	0.96	1.00	0.96	1.00	1.00	1.00	1.00	1.00	1.00
OA2	0.81	0.90	0.81	0.96	0.99	0.96	0.93	0.99	0.93	0.99	1.00	0.99
OA3	0.74	0.86	0.73	0.81	0.89	0.81	0.93	0.97	0.92	0.78	0.87	0.77
OA4	0.76	0.85	0.76	0.81	0.89	0.81	0.81	0.93	0.80	0.79	0.88	0.79
OA5	0.63	0.63	0.62	0.94	0.99	0.93	0.91	0.97	0.91	0.91	0.99	0.91
OA6	0.76	0.85	0.76	0.74	0.81	0.73	0.70	0.79	0.69	0.81	0.91	0.81
OA7	0.85	0.94	0.85	0.83	0.92	0.82	0.94	0.99	0.94	1.00	1.00	1.00
OA8	0.91	0.99	0.91	0.91	0.98	0.91	0.95	0.99	0.95	0.93	0.97	0.93
OA9	0.63	0.67	0.63	0.89	0.95	0.89	0.88	0.95	0.87	0.83	0.92	0.82
OA10	0.74	0.85	0.73	0.76	0.81	0.75	0.99	1.00	0.99	0.84	0.91	0.84
OA11	0.85	0.89	0.84	0.92	0.97	0.92	0.90	0.95	0.90	0.89	0.97	0.88
OA12	0.79	0.90	0.78	0.68	0.78	0.67	0.93	0.98	0.92	0.89	0.97	0.89
OA13	0.79	0.87	0.79	0.89	0.96	0.89	0.94	0.99	0.94	0.99	1.00	0.99
YA1	0.92	0.96	0.92	0.87	0.96	0.87	0.84	0.92	0.84	0.82	0.93	0.82
YA2	0.96	0.99	0.96	0.93	0.96	0.92	0.68	0.79	0.67	0.83	0.93	0.82
YA3	0.90	0.95	0.90	0.96	1.00	0.96	0.71	0.77	0.71	0.90	0.97	0.90
YA4	0.98	0.99	0.97	0.92	0.96	0.92	0.91	0.96	0.91	0.81	0.87	0.81
YA5	0.97	1.00	0.97	0.93	0.98	0.93	0.86	0.94	0.85	0.72	0.85	0.72
YA6	0.99	1.00	0.99	1.00	1.00	1.00	0.86	0.95	0.86	0.98	1.00	0.97
YA7	0.91	0.98	0.91	0.88	0.96	0.87	0.98	1.00	0.98	0.96	0.98	0.96
YA8	0.77	0.82	0.76	0.82	0.91	0.82	0.71	0.77	0.71	0.74	0.78	0.74
YA9	0.75	0.83	0.75	0.76	0.80	0.75	0.85	0.93	0.85	0.70	0.71	0.69
YA10	0.68	0.73	0.67	0.79	0.89	0.79	0.96	1.00	0.96	0.71	0.77	0.70
YA11	0.99	1.00	0.99	0.94	0.98	0.94	0.77	0.85	0.77	0.90	0.97	0.90
YA12	0.86	0.90	0.85	0.88	0.91	0.87	0.97	0.98	0.97	0.89	0.95	0.89
YA13	0.97	0.99	0.97	0.86	0.93	0.85	0.96	1.00	0.96	0.96	1.00	0.96

differences could be indicative of altered motor network function reflecting dedifferentiated and compensational brain activation. This could be a decrease in the segregation of the network as well as the need for integration of different central control processes related to attentional processes.

Classification accuracies were above chance level for each of the four visuomotor tracking tasks. Focusing on the two dimensions, task characteristics (sine vs. steady) and body side (left-hand vs right-hand), showed that the performance of the left vs. right-hand task classification was better in YA, and the classification performance of the sine vs. steady task was better in the OA independent of the body side of execution. This reduced body side specificity with age might be a result of a reduction in segregation of the motor network seen e.g., in lateralization of brain activation in OA as described above. Furthermore, our findings indicate that classification of tasks based on their characteristics is more accurate for OA, which might relate to a reorganization of central processes with age.

4.1. Group and task comparison of DMD mean modes

In the following section we first discuss DMD-derived brain activity patterns (related to analysis depicted in Figs. 2–4), which formed the basis of the task classification. For a better understanding of the classification results and their relation to neurophysiological changes, it is helpful to consider the input characteristics. We did not correct for a baseline to ensure a stronger relationship between DMD results and the classifier performance. Therefore, these results must be considered in light of age-related differences that are already evident in resting networks (Sala-Llonch et al., 2015).

In our previous paper, we reported less segregated sensorimotor network activation and higher need of information integration in OA reflected in higher β_2 activation (Vieluf et al., 2018). Extending those findings, we can confirm higher β DMD mean mode magnitudes, i.e., higher energy of spatially correlated structures in the group of OA consistently in all tasks at (fronto- and parieto-) central electrodes as shown in Fig. 2. Differences in the more complex sine task additionally included occipital electrodes in the

β_2 band that were restricted to the right-hand task. Such changed functional connectivity of central (motor) areas with frontal, parietal, and occipital areas might point to altered motor network functioning in OA (Hong, Liu, Sun, & Tong, 2016; Seidler et al., 2010). The β band has been discussed in terms of the integration and exchange of information in different networks (Hipp, Engel, & Siegel, 2011). Thus, the β band specific differences could indicate a greater need of the integration of attention- and sensory-related information processing in the group of OA as proposed by Vieluf et al. (2018). Albeit speculatively, these differences may reflect higher demands in maintaining status quo of the sensorimotor system due to higher level of noise and dedifferentiated activation of the motor system as reported with aging (Reuter-Lorenz & Lustig, 2005). Besides, it should be noted that for the least complex steady right-task, we found higher α DMD modes in OA over frontal electrodes. The α band has been associated with attentional processes (Klimesch, 2012). More specifically, increased α power in cognitive as well as visuomotor tasks have been suggested to reflect an increased engagement of visual spatial attention (Baravalle, Guisande, Granado, Rosso, & Montani, 2019; Van Diepen, Foxe, & Mazaheri, 2019). Thereby α power gates the attentional focus, by suppressing the processing of irrelevant stimuli, locations, or features (Jensen & Mazaheri, 2010). Thus, increased frontal α DMD mode magnitudes even in the relatively easiest task might indicate an increased need to engage attentional control processes for OA.

To investigate the hypothesis of less differentiation in brain activity in OA, we compared groupwise DMD mean mode magnitudes between the tasks (see Fig. 3). DMD mean mode magnitudes differed most prominently between the right- and left-hand tasks. This was stronger in the group of YA and most prominent in the sine tracking task which could point to different motor control strategies (Siegel, Donner, & Engel, 2012; Yordanova, Falkenstein, & Kolev, 2020) as well as different motor representation of the left- and right-hand task as found in fMRI (Carp et al., 2011) and EEG (Pfurtscheller, 2001) investigations.

Based on reports on the activation of bilateral motor areas in unimanual tasks with age (Carp et al., 2011; Ward & Frackowiak, 2003) we calculated laterality indices of sensorimotor relevant

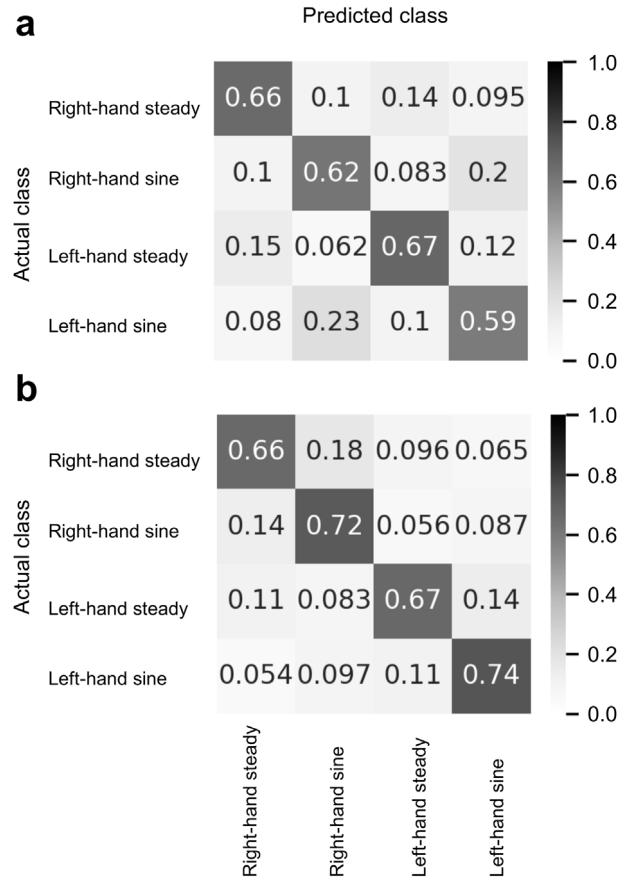


Fig. 5. Confusion Matrix of the 4-class classification. OA (a) and the YA group (b).

electrodes (C3 and C4; see Fig. 4a). We found higher, i.e., more positive lateralization indices, in the β_1 frequency band for the left-hand steady task and lower, i.e., more negative lateralization indices, for the right-hand sine task comparing YA to OA. Such differences in lateralization could point to higher ipsilateral DMD mean mode magnitudes and might correspond to the higher relevance of bilateral recruitment with age, which has been previously interpreted as compensatory involvement and dedifferentiated activation (i.e., less contralateral organization) of neural resources of the motor system (Carp et al., 2011; Sailer, Dichgans, & Gerloff, 2000; Zich, Debener, De Vos, Frerichs, Maurer, & Kranczioch, 2015). With the current results, we cannot clearly differentiate between both effects due to the high complexity and interrelation of these effects as well as the high heterogeneity of the behavioral results in the OA (see Figure S1, supplementary material). This is also reflected in the lack of clear correlations between the behavioral level and the laterality and frontality indices. Thus, both effects could be present in the group. Future work designed to dissociate between these alternatives might, however, test this in different performance groups and with more high-resolution methods. Nevertheless, we found no correlations between the laterality indices and behavioral performance, which could indicate a rather independent non-compensatory effect (see supplementary material Figure S3).

We further expected to find signs of recruitment of brain areas related to attentional processes as described in Vieluf et al. (2018) and calculated frontality indices between Fz and Pz (see Fig. 4b). We assumed this to be reflected in higher frontal activation in older adults especially present in the θ frequency band (Reuter-Lorenz & Cappell, 2008). However, frontality indices in the θ

range did not differ between age groups. Nevertheless, there is a group difference in the α band in the left-hand steady task with values closer to zero indicative of a lower gradient between frontal and parietal α activity in OA. As described above, this could be an expression of altered attentional control (Klimesch, 2012). We assume that changes in attentional and additional cognitive recruitment could be shown more clearly on the basis of task-related analyses.

4.2. Classification performance

In the following, we focus on the results of the classification (summarized in Figs. 5 and 6). We will discuss these findings in relation to above-described age-related brain networks alterations. The classification of the tasks achieved accuracies above chance level in both groups. Nevertheless, we found differences in the classification of the task dimensions (body side: left-hand vs. right-hand and task characteristic: sine vs. steady) between the groups as shown in Fig. 6. Regarding the classification of the body side, the classifier achieved lower performance scores in the OA group, only in the more complex task (sine tracking task). This is in line with Chen et al. (2019) as well as Zich et al. (2015). Chen et al. (2019) found reduced accuracies in the classification of body side of a passive stimulation task in older adults. Zich et al. (2015) compared overt and covert movements and found lower body side classification accuracies for older adults. Both author groups relate their findings to a reduced lateralization of sensorimotor rhythms in older adults. Similarly, we found age-related changes in lateralization, but only in the left-hand steady task as well as the right-hand sine task. Besides, DMD mean mode magnitudes of left- and right-hand sine tasks differed significantly only in the group of YA. This body side specific differences confirms task specificity, which could be due to higher lateralization of brain activation as indication of classification performance. We assume lower distinctiveness of the motor system in older adults (Carp et al., 2011; Cassady et al., 2020; Ward & Frackowiak, 2003). Moreover, the poorer classification performance of the left-hand and right-hand tasks could be explained by this finding. Interestingly, we did not find a lower classifier performance in OA compared to YA in the left-hand vs. right-hand steady tasks. This might suggest that the age-related loss of segregation is more pronounced in more complex tasks. This finding could further indicate a dependence of task complexity for dedifferentiated brain activation as reported for the motor system in older adults (Carp et al., 2011) and extends these previous findings. Taken together the classification results could be an indicator of general reorganization processes and related to aforementioned DMD differences found in this study as well as in Vieluf et al. (2018).

Whereas Chen et al. (2019) and Zich et al. (2015) only report differences in classification performance by body side, we extend these findings by including differences in the classification of task characteristics (sine vs. steady). Similar to above mentioned lower distinctiveness of the motor system in OA, we expected a lower classifier performance in the OA. However, we found a higher classification performance in the group of OA independent of the side of task execution. This finding may indicate that different task demands may have caused rather more diverse brain activity patterns in OA than in YA. Looking at the group differences of the DMD mean modes, it seems that the age-related differences in the less complex tasks are more pronounced, while the groups in the more complex tasks show more similar brain activity patterns. As discussed above, this may be due to a greater need for integration of attentional and sensory information for OA particularly evident in the less complex task, which may be reflected in the differences in the frontality index. However, in OA such integration could be highly depended on the task characteristic as indicated by the classifier results.

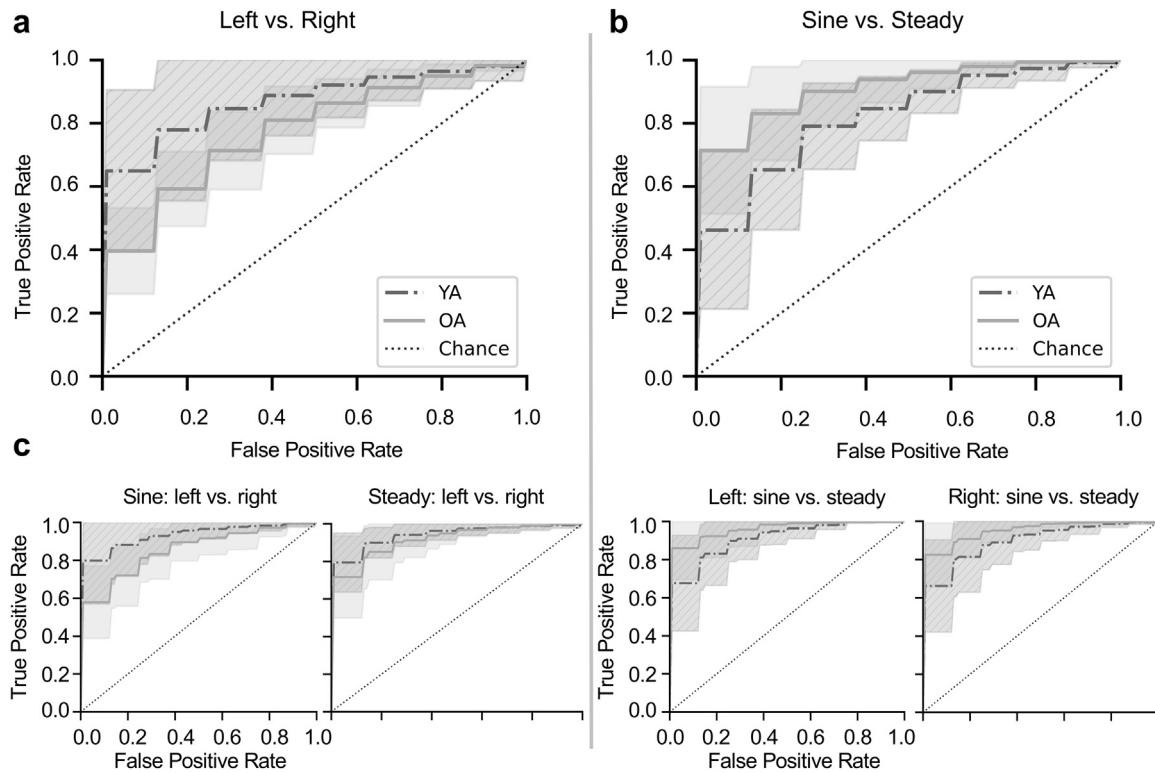


Fig. 6. ROC curve of both two-class classifiers as stepwise classification approach. ROC curves represent the group mean with associated standard deviation as gray (OA) and gray striped (YA) areas. a: left vs. right; b: sine vs. steady; c: sine: left vs. right, steady: left vs. right, left: sine vs. steady, and right: sine vs. steady (from left to right).

In sum, we confirmed previously reported differences in classification performance between tasks (Chen et al., 2019) now for fine motor control tasks and show for the first time the relationship between classification and task characteristics in older adults. We extended the existing knowledge and detect classification results, which depend on processes relevant to attention. Not only the laterality but also the activation of attention-related processes, could be an important factor for classification performance. This observation is supported by BCI systems relying on attention relevant processes like systems based on the P300 shape and dynamics as well as systems based on steady state evoked potentials/response (SSVEP/R) (Abiri, Borhani, Sellers, Jiang, & Zhao, 2019; Middendorf, McMillan, Calhoun, & Jones, 2000). Although they are not rooted in the exploitation of sensorimotor signals used here, they indicate the relevance of attentional processes.

Considering age-related changes in electrophysiological brain activity patterns detected with DMD as well as the classification results, we were able to show indications of the reorganization of central mechanisms in fine-motor force control tasks. Besides the practical relevance, it was also possible to capture age-related changes on the basis of the classifier performance. Machine learning algorithms and multivariate methods such as DMD can thus contribute to data-driven research on age-related changes and extend classical EEG analyses without making assumptions of the underlying data.

Our findings suggest that the influence of individual neurophysiological properties should be integrated into the development of BCI systems, as suggested by Blankertz et al. (2010). In addition, target group-specific characteristics, such as age-related differences, should be incorporated into the design process of BCI systems. This could include the use of age-specific electrode placement and the selection of suitable features and algorithms.

4.3. Methodological consideration

The data collection of the current study was embedded in a larger study context in which participants performed active fine motor control and sensory perception tasks. As a consequence, task classification was done based on an active task rather than on covert movements or passive stimulation as reported in the literature (Chen et al., 2019; Zich et al., 2015). This bears several advantages, first of all, by using an active task we could demonstrate the feasibility of classification in young and older adults. Second, by monitoring the motor outcome, we could account for noncompliance. Non-compliance is seen as one of the reasons for low classification performance in motor imaginary BCI systems (Blankertz et al., 2010). Third, having active tasks facilitated the classification of task characteristic and thus extending previous findings of age-related differences on classification performance (Chen et al., 2019; Zich et al., 2015). Therefore, however, the transferability of our results to BCI settings based on motor imaginary might be limited.

Furthermore, the block design of the trials as well as the broader context of the Bremen-Hand-Study@Jacobs might have impacted the results. In the context of the overall study frame, similar fine motor tracking tasks were repeatedly performed. On one hand, this had the advantage that we can assume that the participants were already familiar with the experimental procedure. On the other hand, plastic changes or training effects that have already taken place could mask age-related effects and might have influenced the classification performance. In contrast, we believe that intensive practice could have been beneficial for the interpretability of the data, as it eliminated age-specific differences in short-term adjustment effects. Finally, regarding the age of the participants, it is important to note that the group of OA was comparatively young (55–65 years old), so that even older participants might show more pronounced changes.

Another aspect to consider is using an alternative machine learning approach. Of course, other approaches might result in different classification performances and could account at least partly for neurophysiological changes as described in Blankertz et al. (2010) for example, with adaptive classification and neural networks. In particular, deep learning methods can automatically learn unexplored features, which could also yield high accuracy for both YA and OA. However, these methods are more suitable for large amounts of data which we do not currently have and often act as a “black box”. This study uses an LDA classifier. The choice of this method, as well as the selection of the chosen parameters could have had an influence on the results. To determine the optimal parameters individually, we used an approach called hyperparameter tuning with Grid Search implemented in Scikit-learn (Pedregosa et al., 2011). Furthermore, we compared the results of the LDA classification with other common algorithms (Random Forest, Support Vector Machine) and obtained comparable results (see Supplementary Material Tables S1 and S2). Moreover, we chose the FBCSP algorithm, on which we based our approach, in combination with an LDA classifier, a very commonly used and in several BCI competitions benchmarked algorithm (Schirrmeister et al., 2017). To extract the features, we have used DMD to create the filter banks preliminary to the CSP. The patterns extracted via DMD based on the interrelation of the EEG signals can be related to the characteristic coherent behavior of physiological networks. Thus, this approach differs from the graph analytical approach of networks constructed by applying bivariate connectivity of EEG signals. DMD was recently found to be highly successful in classifying fine motor tasks (Shiraishi et al., 2020). Moreover, we successfully applied this method in previous work to extract age- and expertise-related sensorimotor network dynamics (Vieluf et al., 2018).

4.4. Conclusion

Based on age-related changes of brain networks, such as additional recruitment of bilateral motor areas (Carp et al., 2011; Ward & Frackowiak, 2003) and attentional resources (Berghuis et al., 2019), we aimed to study differences in the classification of EEG data recorded in active visuomotor tracking tasks.

In summary we found electrophysiological patterns associated with an altered sensorimotor network in OA. Lower task specificity in combination with changes in symmetry of brain activity point to bilateral and dedifferentiated, i.e., less task specific, brain activity of the motor network and activation and interrelation of several networks with age.

Most importantly these electrophysiological brain activity patterns resulted in lower classification performance in the classification of body side of task execution in OA, indicating less segregated brain network activation of the motor system. In contrast, OA showed higher classification performance with respect to the task characteristic. The study of the classifier input indicates the relevance of markers of information integration for classification performance in OA.

The current results confirm previous findings on age-related reorganization of task-related brain networks and expand them with reference to the characteristics of the task. Furthermore, the findings may have practical implications for areas of applied research such as BCI applications. Age-related differences should be taken into account in the development of BCI and neurofeedback systems if they are designed for this target group. This could include the selection of the appropriate positioning of electrodes, e.g., the use of frontal and occipital electrodes, as well as the choice of suitable features and algorithms.

Code availability

Python source code that supports the results is available from: https://github.com/christiangoelz/Code_Classification-of-visuomotor-tasks-YAOA

CRediT authorship contribution statement

C. Goelz: Analyzed data, design and implementation of the work including interpretation of the results, drafting parts of the work. **K. Mora:** Analyzed data, design and implementation of the work including interpretation of the results, drafting parts of the work. **J. Rudisch:** Design and implementation of the work including interpretation of the results, drafting parts of the work. **R. Gaidai:** Analyzed data, design and implementation of the work including interpretation of the results, drafting parts of the work. **E. Reuter:** Conceived and planned the experiments, design and implementation of the work including interpretation of the results, drafting parts of the work. **B. Godde:** Conceived and planned the experiments, design and implementation of the work including interpretation of the results, drafting parts of the work. **C. Reinsberger:** Design and implementation of the work including interpretation of the results, drafting parts of the work. **C. Voelcker-Rehage:** Conceived and planned the experiments, design and implementation of the work including interpretation of the results, drafting parts of the work. **S. Vieluf:** Conceived and planned the experiments, design and implementation of the work including interpretation of the results, drafting parts of the work.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank Janine Ohmann and Sandra Fellehner for their support during data collection.

All authors approved the final version of the manuscript and agreed to be accountable for all aspects of the work.

Funding

The research was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation, VO 1432/7-1 - SPP 1184 and DFG Project-ID 416228727 - SFB 1410). The study was supported within the framework of the equal opportunities concept 2 of the Paderborn University and by the Heinz Nixdorf Westfalian Foundation.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.neunet.2021.04.029>.

References

- Abiri, R., Borhani, S., Sellers, E. W., Jiang, Y., & Zhao, X. (2019). A comprehensive review of EEG-based brain-computer interface paradigms. *Journal of Neural Engineering*, 16(11001).
- Ahn, M., & Jun, S. C. (2015). Performance variation in motor imagery brain-computer interface: A brief review. *Journal of Neuroscience Methods*, 243, 103–110.
- Al Zoubi, O., Ki Wong, C., Kuplicki, R. T., Yeh, H.-w., Mayeli, A., Refai, H., et al. (2018). Predicting age from brain EEG signals—A machine learning approach. *Frontiers in Aging Neuroscience*, 10(184).

- Ang, K. K., Chin, Z. Y., Wang, C., Guan, C., & Zhang, H. (2012). Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b. *Frontiers in neuroscience*, 6(39).
- Barachant, A., Bonnet, S., Congedo, M., & Jutten, C. (2010). Common spatial pattern revisited by Riemannian geometry. In *2010 IEEE international workshop on multimedia signal processing* (pp. 472–476). IEEE.
- Baravalle, R., Guisande, N., Granado, M., Rosso, O. A., & Montani, F. (2019). Characterization of visuomotor/imaginary movements in EEG: An information theory and complex network approach. *Frontiers in Physics*, 7.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B-Methodological*, 57, 289–300.
- Berghuis, K. M. M., Fagioli, S., Maurits, N. M., Zijdewind, I., Marsman, J. B. C., Hortobágyi, T., et al. (2019). Age-related changes in brain deactivation but not in activation after motor learning. *NeuroImage*, 186, 358–368.
- Blankertz, B., Sannelli, C., Halder, S., Hammer, E. M., Kübler, A., Müller, K. R., et al. (2010). Neurophysiological predictor of SMR-based BCI performance. *NeuroImage*, 51, 1303–1309.
- Brunton, B. W., Johnson, L. A., Ojemann, J. G., & Kutz, J. N. (2016). Extracting spatial-temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition. *Journal of Neuroscience Methods*, 258, 1–15.
- Cabeza, R., Anderson, N. D., Locantore, J. K., & McIntosh, A. R. (2002). Aging gracefully: compensatory brain activity in high-performing older adults. *NeuroImage*, 17, 1394–1402.
- Carp, J., Park, J., Hebrank, A., Park, D. C., & Polk, T. A. (2011). Age-related neural dedifferentiation in the motor system. *PloS One*, 6(e29411–e29411).
- Cassady, K., Ruitenberg, M., Reuter-Lorenz, P., Tommerdahl, M., & Seidler, R. (2020). Neural dedifferentiation across the lifespan in the motor and somatosensory systems. *Cerebral Cortex*, 30, New York, N.Y.
- Chen, M. L., Fu, D., Boger, J., & Jiang, N. (2019). Age-related changes in vibrotactile EEG response and its implications in BCI applications: A comparison between older and younger populations. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27, 603–610.
- Cichy, R. M., & Pantazis, D. (2017). Multivariate pattern analysis of MEG and EEG: A comparison of representational structure in time and space. *NeuroImage*, 158, 441–454.
- Davis, S., Dennis, N., Daselaar, S., Fleck, M., & Cabeza, R. (2008). Que PASA? The posterior-anterior shift in aging. *Cerebral Cortex*, 18(5), 1201–1209.
- Donoho, D., & Gavish, M. (2014). The optimal hard threshold for singular values is. *IEEE Transactions on Information Theory*, 60.
- Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, 9, 474–480.
- Götz, C., Voelcker-Rehage, C., Mora, K., Reuter, E.-M., Godde, B., Dellnitz, M., et al. (2018). Improved neural control of movements manifests in expertise-related differences in force output and brain network dynamics. *Frontiers in Physiology*, 9(1540).
- Gómez-Pilar, J., Corrales, R., Luis, L., Alvarez, D., & Hornero, R. (2016). Neurofeedback training with a motor imagery-based BCI: neurocognitive improvements and EEG changes in the elderly. *Medical & Biological Engineering & Computing*, 54.
- Gramfort, A., Luessi, M., Larson, E., Engemann, D., Strohmeier, D., Brodbeck, C., et al. (2013). MEG and EEG data analysis with MNE-Python. *Front. Neurosci.*, 7, 267.
- Grosjean-Wentrup, M., & Buss, M. (2008). Multi-class common spatial patterns and information theoretic feature extraction. *IEEE Transactions on Bio-Medical Engineering*, 55, 1991–2000.
- Heuninckx, S., Wenderoth, N., Debaere, F., Peeters, R., & Swinnen, S. P. (2005). Neural basis of aging: the penetration of cognition into action control. *The Journal of Neuroscience*, 25, 6787–6796.
- Hipp, J. F., Engel, A. K., & Siegel, M. (2011). Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron*, 69, 387–396.
- Hong, X., Liu, Y., Sun, J., & Tong, S. (2016). Age-related differences in the modulation of small-world brain networks during a Go/NoGo task. *Frontiers in Aging Neuroscience*, 8(100).
- Hyvärinen, A., & Oja, E. (1997). A fast fixed-point algorithm for independent component analysis. *Neural Computation*, 9, 1483–1492.
- Jas, M., Engemann, D. A., Bekhti, Y., Raimondo, F., & Gramfort, A. (2017). Autoreject: Automated artifact rejection for MEG and EEG data. *NeuroImage*, 159, 417–429.
- Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: Gating by inhibition. *Frontiers in Human Neuroscience*, 4(186).
- Kasahara, K., DaSalla, C. S., Honda, M., & Hanakawa, T. (2015). Neuroanatomical correlates of brain-computer interface performance. *NeuroImage*, 110, 95–100.
- Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, 16, 606–617.
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164, 177–190.
- Middendorf, M., McMillan, G., Calhoun, G., & Jones, K. S. (2000). Brain-computer interfaces based on the steady-state visual-evoked response. *IEEE Transactions on Rehabilitation Engineering*, 8, 211–214.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9, 97–113.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in Python. *ThE Journal of Machine Learning Research*, 12, 2825–2830.
- Pfurtscheller, G. (2001). Functional brain imaging based on ERD/ERS. *Vision Research*, 41, 1257–1260.
- Ramos-Murguialday, A., Broetz, D., Rea, M., Läer, L., Yilmaz, Ö., Brasil, F. L., et al. (2013). Brain-machine interface in chronic stroke rehabilitation: A controlled study. *Annals of Neurology*, 74, 100–108.
- Reuter-Lorenz, P. A., & Cappell, K. A. (2008). Neurocognitive aging and the compensation hypothesis. *Current Directions in Psychological Science*, 17, 177–182.
- Reuter-Lorenz, P. A., & Lustig, C. (2005). Brain aging: reorganizing discoveries about the aging mind. *Current Opinion in Neurobiology*, 15, 245–251.
- Roland, J., Miller, K., Freudenburg, Z., Sharma, M., Smyth, M., Gaona, C., et al. (2011). The effect of age on human motor electrocorticographic signals and implications for brain-computer interface applications. *Journal of Neural Engineering*, 8(46013).
- Saha, S., & Baumert, M. (2020). Intra- and inter-subject variability in EEG-based sensorimotor brain computer interface: A review. *Frontiers in Computational Neuroscience*, 13(87).
- Sailer, A., Dichgans, J., & Gerloff, C. (2000). The influence of normal aging on the cortical processing of a simple motor task. *Neurology*, 55, 979–985.
- Sala-Llonch, R., Bartrés-Faz, D., & Junqué, C. (2015). Reorganization of brain networks in aging: a review of functional connectivity studies. *Frontiers in Psychology*, 6(663).
- Sannelli, C., Vidaurre, C., Müller, K. R., & Blankertz, B. (2019). A large scale screening study with a SMR-based BCI: Categorization of BCI users and differences in their SMR activity. *PLoS One*, 14, Article e0207351.
- Schirrmeister, R. T., Springenberg, J. T., Fiederer, L. D. J., Glasstetter, M., Eggensperger, K., Tangermann, M., et al. (2017). Deep learning with convolutional neural networks for EEG decoding and visualization. *Human Brain Mapping*, 38, 5391–5420.
- Seidler, R. D., Bernard, J. A., Burutolu, T. B., Fling, B. W., Gordon, M. T., Gwin, J. T., et al. (2010). Motor control and aging: Links to age-related brain structural, functional, and biochemical effects. *Neuroscience & Biobehavioral Reviews*, 34, 721–733.
- Shiraishi, Y., Kawahara, Y., Yamashita, O., Fukuma, R., Yamamoto, S., Saitoh, Y., et al. (2020). Neural decoding of electrocorticographic signals using dynamic mode decomposition. *Journal of Neural Engineering*, 17.
- Siegel, M., Donner, T., & Engel, A. (2012). Spectral fingerprints of large scale neuronal interactions. *Nature Reviews Neuroscience*, 13, 121–134.
- Tu, J. H., Rowley, C. W., Luchtenburg, D. M., Brunton, S. L., & Kutz, J. N. (2014). On dynamic mode decomposition: Theory and applications. *Journal of Computational Dynamics*, 1(391).
- Van Diepen, R. M., Foxe, J. J., & Mazaheri, A. (2019). The functional role of alpha-band activity in attentional processing: the current zeitgeist and future outlook. *Current Opinion in Psychology*, 29, 229–238.
- Vieluf, S., Mora, K., Götz, C., Reuter, E. M., Godde, B., Dellnitz, M., et al. (2018). Age- and expertise-related differences of sensorimotor network dynamics during force control. *Neuroscience*, 388, 203–213.
- Voelcker-Rehage, C., Reuter, E.-M., Vieluf, S., & Godde, B. (2013). Influence of age and expertise on manual dexterity in the work context. In C. M. Schlick, E. Frieling, & J. Wegge (Eds.), *The Bremen-hand-study@jacobs BT - age-differentiated work systems* (pp. 391–415). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Ward, N. S., & Frackowiak, R. S. J. (2003). Age-related changes in the neural correlates of motor performance. *Brain : A Journal of Neurology*, 126, 873–888.
- Yordanova, J., Falkenstein, M., & Kolev, V. (2020). Aging-related changes in motor response-related theta activity. *International Journal of Psychophysiology*, 153, 95–106.
- Zich, C., Debener, S., De Vos, M., Frerichs, S., Maurer, S., & Kranczioch, C. (2015). Lateralization patterns of covert but not overt movements change with age: An EEG neurofeedback study. *NeuroImage*, 116, 80–91.

Published Research Article II

RESEARCH

Open Access



Classification of age groups and task conditions provides additional evidence for differences in electrophysiological correlates of inhibitory control across the lifespan

Christian Goelz¹, Eva-Maria Reuter², Stephanie Fröhlich³, Julian Rudisch³, Ben Godde⁴, Solveig Vieluf^{1,5} and Claudia Voelcker-Rehage^{3*}

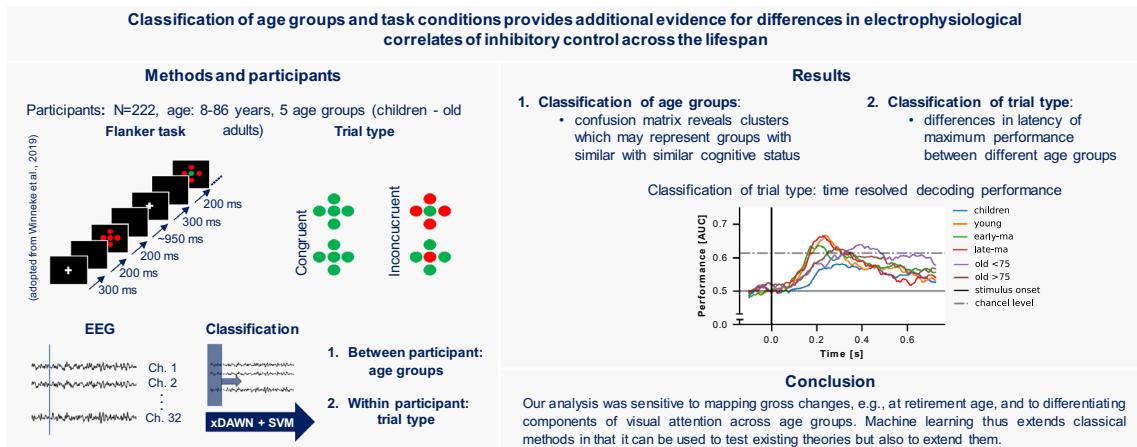
Abstract

The aim of this study was to extend previous findings on selective attention over a lifetime using machine learning procedures. By decoding group membership and stimulus type, we aimed to study differences in the neural representation of inhibitory control across age groups at a single-trial level. We re-analyzed data from 211 subjects from six age groups between 8 and 83 years of age. Based on single-trial EEG recordings during a flanker task, we used support vector machines to predict the age group as well as to determine the presented stimulus type (i.e., congruent, or incongruent stimulus). The classification of group membership was highly above chance level (accuracy: 55%, chance level: 17%). Early EEG responses were found to play an important role, and a grouped pattern of classification performance emerged corresponding to age structure. There was a clear cluster of individuals after retirement, i.e., misclassifications mostly occurred within this cluster. The stimulus type could be classified above chance level in ~95% of subjects. We identified time windows relevant for classification performance that are discussed in the context of early visual attention and conflict processing. In children and older adults, a high variability and latency of these time windows were found. We were able to demonstrate differences in neuronal dynamics at the level of individual trials. Our analysis was sensitive to mapping gross changes, e.g., at retirement age, and to differentiating components of visual attention across age groups, adding value for the diagnosis of cognitive status across the lifespan. Overall, the results highlight the use of machine learning in the study of brain activity over a lifetime.

Keywords Machine learning, Decoding, EEG/ERP, Flanker, Selective attention, Development

*Correspondence:
Claudia Voelcker-Rehage
claudia.voelcker-rehage@uni-muenster.de
Full list of author information is available at the end of the article

Graphical Abstract



1 Introduction

Selective attention (as part of inhibitory control) describes the ability to focus on relevant information while simultaneously suppressing irrelevant or distracting input which is essential for the accomplishment of complex tasks [1]. This ability changes throughout the lifespan. Whereas selective attention develops in children due to the differentiation of brain areas and networks, the opposite is noticeable in older adults, namely, a reduction in selective attention likely related to dedifferentiation processes in the brain [2, 3]. Focusing on the neuronal response to sensory stimuli measured with electroencephalography (EEG) differences in the distribution, amplitude, and latency of event-related potentials (ERP) were reported for different age groups [3]. Comparing six different age groups Reuter et al. [4] confirmed a u-shaped function of ERP markers for encoding and processing speed (i.e., P1, N1, N2, and P3 latencies), markers of visual processing and attention (i.e., P1 and N1 amplitudes) as well as gradual changes in markers of cognitive processing (N2, P3 amplitudes, and P3 distribution). Moreover, results suggest that different neural mechanisms underly performance in children and older adults [4]. The u-shaped function in previous ERP findings suggests that ERP components are similar between children and older adults despite fundamentally different mechanisms, i.e., differentiation in children versus dedifferentiation in older adults. It is unclear whether these differences are reflected in distinctive brain activation patterns and to what extent lifelong changes in electrophysiological markers can be detected at the level of individual trials.

Recently, the use of machine learning techniques to study experimental effects in EEG studies gained popularity as a complement to classical ERP analyses. These methods are referred to as multivariate pattern analysis (MVPA) or decoding approaches and are based on classification algorithms developed in the field of brain-computer interfaces (BCI) [5]. The main idea is to train a machine learning model based on single-trial EEG data that allows to classify a certain behavior or experimental condition. This involves the automatic detection of generalizable multivariate patterns associated with the behavior or experimental condition. Targeting the information content on a single-trial level with respect to an experimental condition rather than averaged activation on single electrodes and time windows, such approaches can be seen as complementary to classical univariate ERP analyses [5]. Classification approaches are less dependent on a priori assumptions (e.g., selection of electrodes or time windows), and naturally simplify the problem of multiple comparisons [6]. In this way, these methods have higher sensitivity by exploiting the interdependence of EEG signals while omitting the information loss due to trial averaging. Moreover, additional analyses allow characterizing the cortical representation of a high number of stimuli, e.g., their dynamics and similarity [5, 7]. Nevertheless, these methods should be considered as a complement to the classical univariate methods, since directional effects cannot be represented. To study the neuronal response to sensory stimuli on single-trial level decoding approaches are often used in a time-resolved manner to investigate the time course of information density in relation to the stimulus. This includes determining the

times at which it is possible to determine the nature of a stimulus based on the neural response to it. Determining the dynamics of cortical stimulus processing is thus captured in a data-driven manner, providing a complement to classical ERP analyses. For example, visual object perception [8] or working memory were investigated using decoding approaches [9]. Recently, Vahid et al. [10] identified predictive neurophysiological processes related to N1 and N2 time windows on a single-trial level associated with selective attention using machine learning methods and highlighted the possibility of such methods to validate and form a new hypothesis in a data-driven way. Moreover, López-García et al. [11] successfully applied time-resolved decoding to study selective attention in a flanker paradigm and confirmed interference processing being reflected in the N2 time window on a single-trial level with this approach. Comparing decoding performance across groups could provide new insights into differences in the cortical representation of stimuli or tasks that would not have been possible with conventional EEG analyses. So far, this has been done for different age groups [, 12, 13] and patient groups [14]. The use of machine learning to derive generalizable principles from the complex interaction of time series data could, therefore, add value to the study of lifelong changes in the neural representation of selective attention by confirming hypotheses in a data-driven manner and exploring important implications for its applications [, 15, 16].

Extending a previous analysis by Reuter et al. [4], we aimed to use machine learning to further explore selective attention across the lifespan by decoding the neural representation of inhibitory control using classification in different age groups. We further aimed to explore the classification of age groups to infer age-related changes in selective attention. We hypothesized to classify the type of stimulus and group membership above chance in all

groups. We further assumed different decoding trajectories in decoding stimulus types.

2 Methods

2.1 Data set

We used data from a total of 222 participants originally collected in three experimental studies each focusing on a different age group. Data sets comprised of EEG data recorded during a modified Flanker task [4]. The full data set includes 92 data sets collected in the setting of the Bremen-Hand-Study@Jacobs [, , 17–19] (Study 1), 81 data sets as part of the re-LOAD project [20, 21] (Study 2) and 49 data sets as part of the CEBRA project [22] (Study 3). The data were first analyzed in a comprehensive manner in Reuter et al. [4] including all 222 participants. All adult participants gave their written informed consent. For children, guardians gave their written informed consent and children agreed to participate. For Study 1 and Study 3 the German Psychological Society and for Study 2 the Ethics Committee of the Faculty of Humanities of the Saarland University, Germany, granted ethical approval. Participants older than 65 scored higher than 27 in the Mini-Mental State Examination (MMSE, [23]) or at least 23 in the Montreal Cognitive Assessment (MoCA, [24, 25]). Participants are separated in the following age categories [4]: children (8 to 10 years), young adults (20 to 29 years), early middle-aged adults (36 to 48 years), late middle-aged adults (55 to 64), old adults < 75 (66 to 75 years), very old adults > 75 (76 to 83 years). We excluded eight participants from further analysis as they had less than 35 correct trials in one of the conditions. Due to poor EEG data quality, we further excluded five participants. Group characteristics included in the final data set are displayed in Table 1.

2.2 Experimental procedures

All participants performed a modified version of the Flanker task previously reported in Reuter et al.

Table 1 Group characteristics

Group	N (female)	Age	Number of trials									
			Incongruent					Congruent				
			Mean	Std	N	Mean	Std	Min	Max	N	Mean	Std
Children	46 (23)	9.32	0.65		3262	70.91	15.30	36	98	3026	65.78	13.27
Young	39 (34)	22.85	2.50		2554	65.49	24.31	40	109	2476	63.49	22.59
Early-ma	21 (12)	42.62	3.61		2038	97.05	10.33	79	125	1951	92.90	8.39
Late-ma	25 (14)	59.04	2.39		2378	95.12	10.65	71	109	2372	94.88	9.01
Old < 75	40 (36)	71.93	3.04		2355	58.88	19.73	40	111	2263	56.58	14.42
Very old > 75	38 (30)	78.14	1.94		2632	69.26	21.07	39	105	2506	65.95	21.60

Late-ma late middle-aged, *early-ma* early middle-aged, *std* standard deviation, *min* minimum, *max* maximum

[17], Winneke et al. [19, 56] and summarized in Reuter et al. [4]. The stimuli consisted of four circles surrounding a target circle in the middle. The target circle was either set to red or green and the task was to press the corresponding button with the index or middle finger of the right hand as fast as possible. The surrounding (flanking) targets were either set to blue (neutral condition) to the same color as the target (congruent condition) or the opposite color, i.e., green target and red flanker and vice versa (incongruent condition).

The experimental procedures were identical between all studies except for trial number and stimulus duration. In Study 1 and Study 3, participants performed 300 trials (approx. 100 trials per condition), whereas in Study 2, they performed 150 trials (approx. 50 trials per condition) in randomized order. Stimuli were presented for 200 ms in Study 1 and Study 3, whereas in Study 2, stimuli were presented for 500 ms. Each trial started with a white fixation cross (300 ms), next a blank screen (200 ms) was presented followed by the presentation of the stimulus and a variable intertrial interval of about 950 ms (i.e., 800 ms to 1100 ms). Participants did a minimum of 20 practice trials and were asked to respond as fast and precisely as possible. Only congruent (no inhibitory control) and incongruent (inhibitory control) conditions as well as correct trials, i.e., trials with a correct response between 100 ms and 1200 ms after stimulus onset, were considered in the following analyses.

2.3 EEG recordings and preprocessing

EEG was recorded with the same 32 Biosemi active electrode system (ActiveTwo, BioSemi, Amsterdam, Netherlands) throughout all studies. Electrodes were placed according to the 1020 system (Jasper, 26). Six additional electrodes recorded vertical and horizontal eye movements as well as mastoid potentials. The sampling rate was set to 2048 Hz and an online band pass filter between 0.16 and 100 Hz was used. Prior to classification we downsampled the data to 256 Hz and filtered between 1 Hz and 40 Hz using the default FIR filter implemented in MNE-Python (version 1.1). Next, we cut the data to segments of 900 ms, i.e., –100 ms to 800 ms from stimulus onset.

2.4 Machine learning

For classification we relied on a combination of spatial filtering and classification using support vector machines (SVMs) with radial basis function (rbf) kernel to classify the age group and stimulus type. Spatial filtering allows to extract induced spatial patterns, i.e., neuronal responses to external stimuli at single-trial level with not phase locked dynamics, and is thus advantageous compared to the creation of averages across trials [27, 28]. To

enhance signal to noise ratio, i.e., to enhance ERP responses we used the xDAWN algorithm [29]. The xDAWN algorithm was originally developed for P300 evoked potentials in the BCI context and subsequently extended to any type of ERP (see Cecotti and Ries [27] for an overview). Compared to other spatial filtering methods like principal component analysis or independent component analysis xDAWN was shown to be more suitable for the analysis of ERPs aiming at estimating temporal and spatial signatures [30]. Based on the pre-processed EEG segments (cf.2.3), five spatial filters were trained using the training data only (cf.2.4.1). Spatially filtered EEG signals were then classified using a SVM with rbf kernel. Scikit-learn (version 1.1.1), MNE (version 1.1), and imbalance-learn (version 0.10.1) were used to implement the machine-learning pipelines.

2.4.1 Classification of age group

We classified group membership on a trial-by-trial basis using all participants data to build a model capable of predicting the associated age group for each trial. We trained and tested our xDAWN+SVM model using stratified tenfold cross validation. We, therefore, randomly split the data 10 times using all trials (N trials per fold: 21538.50 ± 363.91) of 188 participants for training and all trials (N trials per fold: 2981.30 ± 365.50) of 21 participants for testing our model preserving the percentage of samples for each class, i.e., age group. To account for class inequality, we randomly subsampled the training data to the minority class in each fold. As such the same number of trials per age group was present in the training data. In addition to a model using all time samples, we aimed to capture time-resolved decoding performance. That is, we trained and tested repeatedly based on 20 data points with 19 overlapping data points and thus iterated from the beginning to the end of the trial to draw conclusions about decoding performance over time.

2.4.2 Classification of stimulus type

Stimulus type classification (congruent/incongruent) was performed within participants, i.e., analogous to the group-level procedure, we trained and tested a xDAWN+SVM model for each participant individually. We divided all trials 10 times randomly in a training set (N trials: 113.61 ± 33.44) and a testing set (N trials: 28.90 ± 8.33) preserving the percentage of samples for each class. We trained and tested the models in a windowing approach using 20 data points with 19 data points overlap to infer decoding performance over time and to be able to compare this between age groups to determine the most discriminative features, i.e., time windows.

3 Statistics

We calculated statistics with python using scipy (version 1.8.0) and statsmodels (version 0.13.2). For group classification we report the confusion matrices including precision and recall. In addition, we calculated the accuracy score for each time window. To test if the model of age group classification performed above chance level, we calculated the threshold over which the classification results could be considered significant. We, therefore, relied on the method described in Combrisson and Jerbi [31]. Compared to a permutation approach which randomly shuffles the labels (e.g., 1000 times) with the aim to create a null distribution against which the significance of a model can be tested, this method uses a binomial cumulative distribution to estimate the significance threshold of a classifier. Since permutation approaches are computationally very expensive, especially for large data sets and time-resolved decoding, we relied on this approach due to its computational efficiency and suitability. Due to the large number of trials and our time-resolved decoding approach, it was not feasible to generate a robust null distribution using a high number of group-level permutations. We had a sufficiently large database at the group level, which was more tolerant of deviations from the assumption of a binomial distribution due to cross-validation parameters, such as classifier type and feature space [32]. Besides, comparable results between this approach and a permutation approach were reported [31].

For task classification at an individual level, we reported the area under the receiver operating characteristic curve (AUC) as a mean over all cross-validation folds for each time window. We relied on this metric rather than accuracy as López-García et al. [11] emphasized higher sensitivity of this score compared to accuracy for task classification over time. Since the number of trials within a subject was much smaller than on group levels and for this reason the assumption of a binomial distribution could not be accepted, we used the permutation approach on this level to test the models of stimulus type classification against chance [33]. Therefore, we shuffled the labels 1000 times for each participant creating a null distribution at the group level. From this distribution, we derived the threshold over which the classifier performance could be considered significant at an alpha level of 0.05. We extracted the maximum AUC score and timepoint of maximum AUC score for each participant. In the case of normal distributed data, we conducted one-way ANOVAs to assess the effects of age group on maximum classification performance and timepoint followed by *t* tests for post hoc comparisons. Otherwise, we used nonparametric Kruskal-Wallis test followed by Dunn's tests for post hoc comparisons. For all tests, the alpha level was set to 0.05 and false discovery rate was used to correct for multiple comparisons [34].

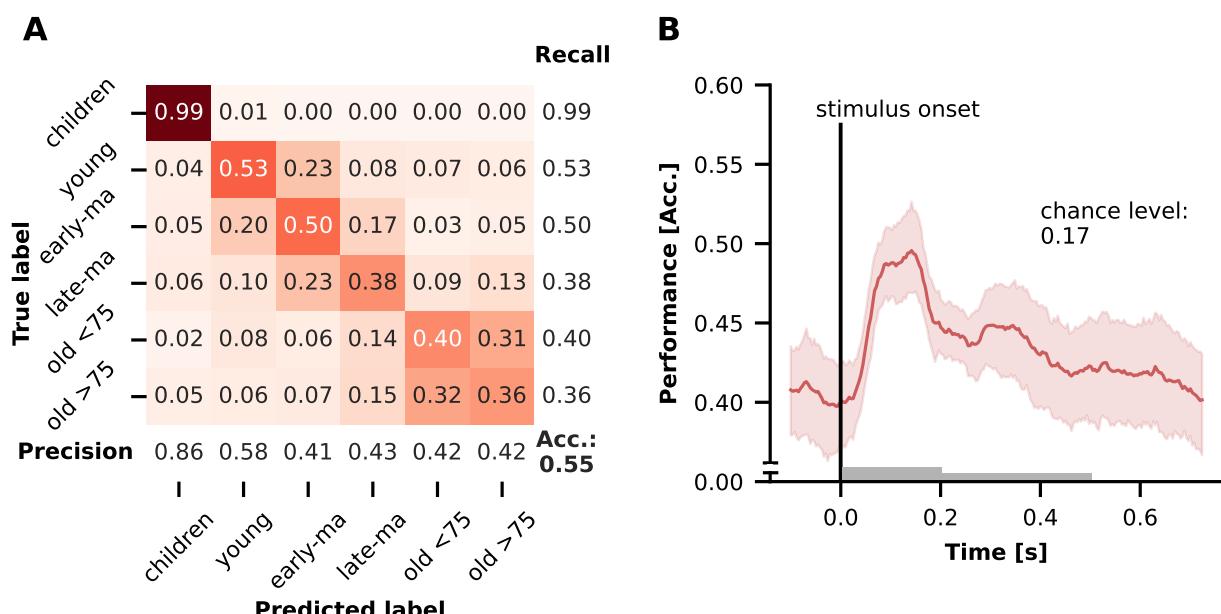


Fig. 1 Classification Performance of group membership classification. **A** Confusion matrix including precision, recall for each group as well as the accuracy score (Acc.). **B** Accuracy as a function of time as mean over folds with 95% confidence interval. The gray bars correspond to the stimulus duration. Late-ma late middle-aged, early-ma early middle-aged, Acc accuracy

4 Results

4.1 Classification of group membership

The classification of group membership is presented in Fig. 1. Classification was well above chance level (accuracy = 0.55, chance level: 0.17 at $p=0.05$). Visualizing misclassifications in a confusion matrix (see Fig. 1A) a clear group structure emerged. While the classification of the children's group worked best, trials in the remaining groups were increasingly classified into adjacent age groups. Here, a larger number of misclassifications can

be observed between the younger and early-/late-middle aged adults and a clear cut off to the older adults <75 . Within the two older groups old >75 and old <75 , however, the highest number of misclassifications can be observed. With respect to time-resolved decoding classification performance was up to 10% higher during the task. The classification performance increased shortly after stimulus onset, with the maximum between 100 ms and 200 ms which is where N1 and P1 are typically reported (see also Fig. 2D). Thereafter, the classification

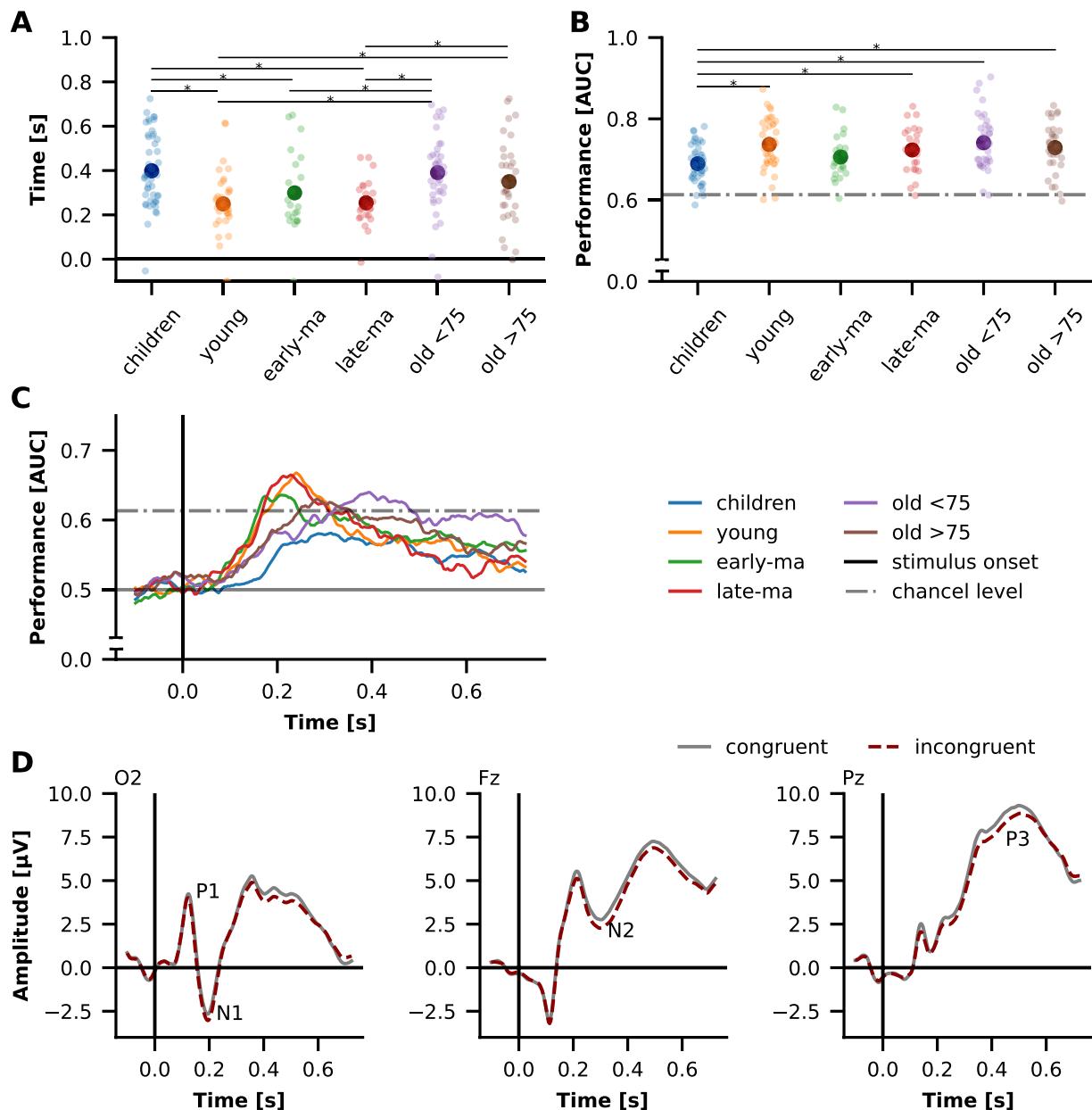


Fig. 2 Performance for decoding of stimulus type. **A** Time of maximum decoding performance. **B** Maximum decoding performance. Each point represents one participant, large points represent the group average. *Denotes a significant difference. **C** Group means decoding trajectories. **D** Whole data set ERPs at electrodes O2, Fz, Pz. AUC area under the receiver operating characteristic curve, late-ma late middle-aged, early-ma early middle-aged

performance decreased until the end of the trial (see Fig. 1B).

4.2 Results of task classification and their differences between groups

Task Classification was possible above chance level for 95.75% of all participants (chance level: 0.61 at $p=0.05$, see Fig. 2B). Classification failed in three children, two young adults, and one participant in each of the other groups. The AUC score significantly differed between the groups [$F(5206)=4.805, p < 0.001$]. Post-hoc comparison revealed significant lower scores in children compared to young adults, late middle-aged adults, old adults < 75 and very old adults > 75 ($p < 0.05$).

While the maximum AUC score occurred around 200 ms to 300 ms in the young, early middle-aged, and late middle-aged adults (see Fig. 2A, C), this was later (at approximately 400 ms) in the children and both groups of older adults (old adults < 75 and very old adults > 75). Overall, however, this timepoint was more dispersed among the children and old. The timepoints corresponded to the ERP components N1, N2, and P3 (see Fig. 2D). Kruskal-Wallis test revealed significant group differences for this [$H(5)=35.575, p < 0.001$]. Children's AUC scores peaked significantly later than AUC scores of young, early middle-aged, and late middle-aged adults. In addition, the maximum AUC score occurred later in old adults < 75 compared to early middle-aged adults. Further old adults < 75 and very old adults > 75 had later peak AUC scores than the young and the late middle-aged adults (all $p < 0.05$, see Fig. 2A).

5 Discussion

In this study, we aimed to extend previous results [4] on the development of selective attention (as one part of inhibitory control) across the lifespan. We used classification algorithms on EEG data recorded during a flanker task to infer group differences in the dynamics of central processing of selective attention by decoding group membership and the type of stimulus presented at the individual level. Both, decoding of group membership and decoding of stimulus type were significantly above chance. By studying the decoding performance over time, we found that differentiation of groups was performed best early after stimulus onset at around 100 ms to 200 ms. Furthermore, we found different trajectories of decoding the stimulus type between groups. While decoding performance in younger adults to late middle-aged adults was maximal at about 200 ms, this was highly variable in children and older adults. On average, the period around 400 ms was most important for decoding the stimulus type in these groups.

5.1 Classification of group memberships peaked early after stimulus onset

The classification of group membership was above chance level. Precision and recall of the group of children were highest. Misclassified trials in the other groups were mainly assigned to adjacent age groups. Overall, this reflected the age structure of the sample. While the gap between children and the next older group of young adults was quite large, it was smaller for the other groups. However, it should be emphasized that misclassification in the two oldest age groups of old adults < 75 and very old adults > 75 years was quite high in comparison and occurred mainly within these groups. This is certainly partly a reflection of the data structure but could also be an indicator for the high variability of neural activity within these age groups as aging trajectories are highly individual [35]. It should also be noted that there was almost no misclassification between the two oldest groups and children. While ERP markers in the literature on lifetime changes of selective attention are often described as u-shaped [3, 4], our results may highlight that different mechanisms at the beginning and the end of this u-shaped trajectory, i.e., differentiation in children and dedifferentiation in older individuals, are reflected in different brain activation patterns. Furthermore, late middle-aged adults were less likely to be misclassified, although the age gap to the adjacent groups was comparable to that between the two older groups. Late middle-aged adults were also rarely classified as older, which could indicate strong changes in cognitive performance after reaching retirement age (65 years in Germany). We are unable to shed light on this due to a lack of data on participants' retirement. However, a decline in cognitive abilities and cerebral perfusion after retirement is reported in the literature [36–38]. Our results suggest that this is also detectable at the neural level based on EEG recordings which might argue for the diagnostic value of task-related EEG.

Classification performance increased shortly after stimulus onset, peaked at 100 ms and 200 ms and gradually decreased until trial offset. While group differences based on averaged ERP data were described in the components N1, P1, N2, and P3 in previous research [4], we were able to show here that even at the single-trial level group-differentiating information is present. Instead of selected timepoints and electrodes, our analysis further allowed us to determine in a data-driven way at which timepoints the EEG signal was best differentiable between groups. The peak in decoding performance was present at 100 ms to 200 ms in which P1 and N1 ERP components are present (see Fig. 2D) indicating that the early components contribute most to the classification. These ERP components are typically discussed in

relation to sensory inhibition and visual attention [39, 40]. Other studies have found P1 and N1 components in children and older adults to be of greater amplitude and latency [41–44]. In the previous study, large amplitude differences were found between age groups in N1 and P1 components, which was interpreted as a reduction in intracortical inhibition in children and older adults, and as increased visual attention in older adults [4]. The high decoding performance found just in this period underlines these previous findings at the level of single-trials. However, maturation processes in childhood and changes in scalp to scull conductivity over the lifespan could have had an influence [45].

In general, classification performance during the task was up to 10% higher than before the onset of the stimulus. This highlights the added value of task-related EEG. Thus, task-related EEG could be used in clinical contexts to predict, for example, the cognitive status of patients. This could be especially of interest in contexts, such as the classification of mild cognitive impairment, where resting state EEG alone could have limited power [46, 47].

5.2 Trajectories of decoding stimulus type differed between group

Peak performance of stimulus type decoding at individual level was above chance level for over 95% of all participants which is in line with previous ERP results finding differences between stimulus types in N2 amplitude and P3 amplitude and latency [3, 4]. There were no group differences in maximum decoding performance between all adult groups. In children, on the other hand, decoding performance was significantly lower. The classifier was less able to distinguish between trials that required inhibitory control and those that did not in this group. At the behavioral level, lower performance in attentional control tasks has been reported in children and it is proposed that the ability to process interfering information develops slowly in children [48]. In fact, this could also be shown in the present data at the behavioral level [4]. Here, especially in children, a large interference effect was shown, i.e., a lower accuracy in the incongruent compared to the congruent condition, which could indicate that inhibitory control has not yet fully been released. Moreover, it was shown from ERP amplitudes of N2 and P3 components that these are less differentiated between stimuli with different attentional demands in older adults but also in children confirming this assumption on the neural level [4, 49]. However, only marginal interaction effects between stimulus type and age group were found based on classical ERP analyses [4]. Using machine learning we go beyond these earlier ERP results, as we found different time curves or trajectories of classification performance between the age groups, characterized

by different time windows of maximum decoding performance. Here, we found high variability in children and the two oldest age groups suggesting a high degree of individuality in visual attention in these age groups which may reflect growth in children or deterioration in older adults. For younger to late middle-aged adults, the trajectories were very similar. The highest decoding performance was observed in the time frame around 200 ms, suggesting the importance of this time frame for discriminating between congruent and incongruent conditions. This is consistent with findings on decoding stimulus type in tasks that capture inhibitory control [10, 11]. ERP components N1 and N2 also occur during these periods (see Fig. 2D and Reuter et al. [4]). N1 was shown to be modulated by early visual attentional processes [50] and N2 was discussed as marker of conflict processing in flanker tasks [51]. In contrast, maximum decoding performance in the two oldest age groups was delayed and peaked at approximately 300 ms to 400 ms, suggesting that later timing and components are critical for classification between stimulus categories in these age groups. This difference in the time windows that allow to differentiate between trials with inhibitory control and trials without was not reflected in the previous ERP analyses [4] and could indicate the general slowing of cognitive processes [52].

While decoding or classification approaches have been used in aging research to map neural distinctiveness in the visual [2] or motor system [13, 53] related to dedifferentiation, we show here that changes in the dynamics of stimulus processing can also be identified using machine learning methods. The methods made it possible to map the processing of stimuli at the individual level. On one hand, this could provide additional diagnostic value and be tested in clinical applications. On the other hand, the identified differences between the age groups could have implications for the development of technical systems for the automatic identification of attentional states.

5.3 Methodological considerations

By combining different data sources, it was possible to access a large data set for our analyses based on machine learning. However, there are small differences between the individual studies due to methodological differences in Study 2. In particular, the presentation length of the stimulus was longer in this study. Although, subjects of Study 2 were included in three, the younger and the two older, age groups and we cannot observe many misclassifications between these age groups in the group classification, we assume that these differences did not have a large impact on the results [4]. Furthermore, the data set of Study 2 consisted of female participants only. For this reason, female participants were overrepresented in

the young adult group and the two older groups, but the influence of gender in this data set was already estimated to be small [4]. Last, the number of trials performed was lower in Study 2. To have a database as large as possible as a basis for the machine learning analyses and to reduce selection errors, we decided against rectifying a certain number of trials. Furthermore, the influence of the number of trials can be assumed to be small [10]. Another influencing factor we would like to point out is the influence of different noise levels in the EEG data, which might have influenced the classification and thus the group comparisons. It is unclear and data set specific how large the effect is on the classification performance. Our previous analyses with different preprocessing strategies showed comparable results. Unlike the raw data, we further used xDAWN spatial filtering to maximize signal to signal and noise ratio at the individual level. Finally, we would like to point out that this analysis should be seen as a complement to previous ERP results, since we are examining the information content of stimulus processing to study differences in the cortical representation of inhibitory control across the lifespan but cannot make any statement about the direction of the effects. However, using machine learning methods in this work, we were free of assumptions regarding the localization of effects in space (electrodes) and time (time windows) and were able to achieve a higher sensitivity to stimulus effects across the lifespan. This approach can be attributed to decoding or MVPA approaches and could be used in studies with higher number of stimuli as starting point for further analyses to study the representational structure of stimuli as it is done in the representation similarity analysis (RSA) framework [7]. Thereby, the performance of classifiers serves as a measure for the difference between two brain activity patterns. Cross validated metrics, as used in this study, are considered advantageous [54].

6 Conclusion

In summary, we were able to extend previous results using machine learning techniques to detect age and task differences in cognitive processing on a single-trial level. This is especially crucial for a step behind classical ERP components and a more direct link between behavior and neural dynamics [55]. The data-driven approach used in this research particularly highlights early attentional processes in the classification of age groups and suggests the benefit of task-related EEG data in the classification of different age groups, which could be used in clinical contexts. With respect to information processing in selective attention our analyses could confirm the relevance of time windows corresponding to N1 and N2 components reported in ERP studies. Furthermore, the

time windows relevant for inhibitory control differed between groups, i.e., later time windows were relevant in older adults, suggesting that different processes are important for selective attention at different ages. Overall, we showed that using machine learning compared to a priori selected electrodes and timepoints, we were able to obtain assumption-free insights into differences in inhibitory control over the lifespan. Machine learning thus represents an extension to classical methods that can be used to test existing theories but also to extend them.

Abbreviations

BCI	Brain–computer interface
EEG	Electroencephalography
ERP	Event-related potential
MMSE	Mini-mental state examination
MoCA	Montreal Cognitive Assessment
MVPA	Multivariate pattern analysis
rbf	Radial basis function
RSA	Representation similarity analysis
SVM	Support vector machine

Acknowledgements

We thank Patrick Jersch, Janine Ohmann, Sandra Fellner, and other student assistants for their support during data collection. We would also like to thank Henning Budde, Lena Hübner, Flora Koutsandreou, and Axel Winneke for the work on which this study is based.

Author contributions

CG: conceptualization, software, formal analysis, writing—original draft. EMR: conceptualization, investigation, data curation, writing—review and editing. SF: writing—review and editing. JR: writing—review and editing. BG: conceptualization, writing—review and editing. SV: conceptualization, investigation, supervision, writing—review and editing. CV-R: conceptualization, investigation, supervision, project administration, writing—review and editing. All authors read and approved the final manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL. The research was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation, BU 1837/5-2, VO 1432/7-1, VO 1432/15-1, GO-802/8-1, Project ID 416228727-SFB 1410).

Availability of data and materials

The data sets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Python source code is available:

- Project name: ML_selective_attention.
- Project home page https://github.com/christiangoelz/ML_selective_attention.
- Archived version: <https://doi.org/10.5281/zenodo.7551173>.
- Operating system(s): Platform independent.
- Programming language: Python.
- Other requirements: Python 3.7 or higher, for packages see file ‘requirements.txt’.

Declarations

Competing interests

The authors declare that they have no competing interests.

Author details

¹Institute of Sports Medicine, Paderborn University, Paderborn, Germany.

²Department of Sport and Health Sciences, Technical University of Munich,

Munich, Germany.³Department of Neuromotor Behavior and Exercise, Institute of Sport and Exercise Sciences, University of Münster, Wilhelm-Schickard-Str. 8, 48149 Münster, Germany.⁴School of Business, Social and Decision Sciences, Constructor University, Bremen, Germany.⁵Division of Epilepsy and Clinical Neurophysiology, Department of Neurology, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA.

Received: 20 January 2023 Accepted: 1 April 2023

Published online: 08 May 2023

References

- Diamond A (2013) Executive functions. *Annu Rev Psychol* 64(1):135–168. <https://doi.org/10.1146/annurev-psych-113011-143750>
- Carp J, Park J, Polk TA, Park DC (2011) Age differences in neural distinctiveness revealed by multi-voxel pattern analysis. *Neuroimage* 56(2):736–743. <https://doi.org/10.1016/j.neuroimage.2010.04.267>
- Mueller V, Brehmer Y, von Oertzen T, Li S-C, Lindenberger U (2008) Electrophysiological correlates of selective attention: a lifespan comparison. *BMC Neurosci* 9(1):18. <https://doi.org/10.1186/1471-2202-9-18>
- Reuter E-M, Vieluf S, Koutsandréou F, Hübner L, Budde H, Godde B, Voelcker-Rehage C (2019) A non-linear relationship between selective attention and associated ERP markers across the lifespan. *Front Psychol*. <https://doi.org/10.3389/fpsyg.2019.00030>
- Grootswagers T, Wardle SG, Carlson TA (2017) Decoding dynamic brain patterns from evoked responses: a tutorial on multivariate pattern analysis applied to time series neuroimaging data. *J Cogn Neurosci* 29(4):677–697. https://doi.org/10.1162/jocn_a_01068
- Fahrenfort JJ, van Driel J, van Gaal S, Olivers CNL (2018) From ERPs to MVPA using the Amsterdam decoding and modeling toolbox (ADAM). *Front Neurosci*. <https://doi.org/10.3389/fnins.2018.00368>
- Kriegeskorte N, Mur M, Bandettini P (2008) Representational similarity analysis—connecting the branches of systems neuroscience. *Front Syst Neurosci* 2:4. <https://doi.org/10.3389/neuro.06.004.2008>
- Cauchoux M, Barragan-Jason G, Serre T, Barbeau EJ (2014) The neural dynamics of face detection in the wild revealed by MVPA. *J Neurosci* 34(3):846–854. <https://doi.org/10.1523/JNEUROSCI.3030-13.2014>
- Bae G-Y, Luck SJ (2018) Dissociable decoding of spatial attention and working memory from EEG oscillations and sustained potentials. *J Neurosci* 38(2):409–422. <https://doi.org/10.1523/JNEUROSCI.2860-17.2017>
- Vahid A, Mückschel M, Stober S, Stock A-K, Beste C (2020) Applying deep learning to single-trial EEG data provides evidence for complementary theories on action control. *Commun Biol* 3(1):112. <https://doi.org/10.1038/s42003-020-0846-z>
- López-García D, Sobrado A, Peñalver JMG, Górriz JM, Ruz M (2020) Multivariate pattern analysis techniques for electroencephalography data to study flanker interference effects. *Int J Neural Syst* 30(7):2050024. <https://doi.org/10.1142/S0129065720500240>
- Csizmadia P, Czigler I, Nagy B, Gaál ZA (2021) Does creativity influence visual perception?—An event-related potential study with younger and older adults. *Front Psychol* 12:742116. <https://doi.org/10.3389/fpsyg.2021.742116>
- Goelz C, Mora K, Rudisch J, Gaidai R, Reuter E, Godde B, Reinsberger C, Voelcker-Rehage C, Vieluf S (2021) Classification of visuomotor tasks based on electroencephalographic data depends on age-related differences in brain activity patterns. *Neural Netw* 142:363–374. <https://doi.org/10.1016/j.neunet.2021.04.029>
- Bae GY, Leonard CJ, Hahn B, Gold JM, Luck SJ (2020) Assessing the information content of ERP signals in schizophrenia using multivariate decoding methods. *Neuroimage Clin* 25:102179. <https://doi.org/10.1016/j.jneuroim.2020.102179>
- Brunton BW, Beyeler M (2019) Data-driven models in human neuroscience and neuroengineering. *Curr Opin Neurobiol* 58:21–29. <https://doi.org/10.1016/j.conb.2019.06.008>
- Bzdok D, Yeo BTT (2017) Inference in the age of big data: future perspectives on neuroscience. *Neuroimage* 155:549–564. <https://doi.org/10.1016/j.neuroimage.2017.04.061>
- Reuter E-M, Voelcker-Rehage C, Vieluf S, Lesemann FP, Godde B (2017) The P3 parietal-to-frontal shift relates to age-related slowing in a selective attention task. *J Psychophysiol* 31(2):49–66. <https://doi.org/10.1027/0269-8803/a000167>
- Voelcker-Rehage C, Reuter EM, Vieluf S, Godde B (2013) Influence of age and expertise on manual dexterity in the work context: the bremen-hand-study@Jacobs. In: Schlick CM, Frieling E, Wegge J (eds) *Age-differentiated work systems*. Springer, Berlin, pp 391–415. https://doi.org/10.1007/978-3-642-35057-3_17
- Winneke AH, Godde B, Reuter E-M, Vieluf S, Voelcker-Rehage C (2012) The association between physical activity and attentional control in younger and older middle-aged adults. *GeroPsych* 25(4):207–221. <https://doi.org/10.1024/1662-9647/a000072>
- Hübner L, Godde B, Voelcker-Rehage C (2018) Acute exercise as an intervention to trigger motor performance and EEG beta activity in older adults. *Neural Plast* 2018:4756785. <https://doi.org/10.1155/2018/4756785>
- Hübner L, Godde B, Voelcker-Rehage C (2018) Older adults reveal enhanced task-related beta power decreases during a force modulation task. *Behav Brain Res* 345:104–113. <https://doi.org/10.1016/j.bbr.2018.02.028>
- Koutsandréou F, Wegner M, Niemann C, Budde H (2016) Effects of motor versus cardiovascular exercise training on children's working memory. *Med Sci Sports Exerc* 48(6):1144–1152. <https://doi.org/10.1249/mss.0000000000000869>
- Folstein MF, Folstein SE, McHugh PR (1975) "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 12(3):189–198. [https://doi.org/10.1016/0022-3956\(75\)90026-6](https://doi.org/10.1016/0022-3956(75)90026-6)
- Julayanont P, Nasreddine ZS (2017) Montreal Cognitive Assessment (MoCA): concept and clinical review. In: Larner AJ (ed) *Cognitive screening instruments: a practical approach*. Springer International Publishing, Cham, pp 139–195. https://doi.org/10.1007/978-3-319-44775-9_7
- Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H (2005) The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* 53(4):695–699. <https://doi.org/10.1111/j.1532-5415.2005.53221.x>
- Jasper HH (1958) The ten-twenty electrode system of the International Federation. *Electroencephalogr. Clin. Neurophysiol.* 10:371–375
- Cecotti H, Ries AJ (2017) Best practice for single-trial detection of event-related potentials: application to brain-computer interfaces. *Int J Psychophysiol* 111:156–169. <https://doi.org/10.1016/j.ijpsycho.2016.07.500>
- King J-R, Gwilliams L, Holdgraf C, Sassenhagen J, Barachant A, Engemann D, Larson E, Gramfort A (2018) Encoding and decoding neuronal dynamics: methodological framework to uncover the algorithms of cognition
- Rivet B, Souloumiac A, Attina V, Gibert G (2009) xDAWN Algorithm to enhance evoked potentials: application to brain-computer interface. *IEEE Trans Biomed Eng* 56(8):2035–2043. <https://doi.org/10.1109/TBME.2009.2012869>
- Rivet B, Souloumiac A (2013) Optimal linear spatial filters for event-related potentials based on a spatio-temporal model: asymptotical performance analysis. *Signal Process* 93(2):387–398. <https://doi.org/10.1016/j.sigpro.2012.07.019>
- Combrisson E, Jerbi K (2015) Exceeding chance level by chance: the caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. *J Neurosci Methods* 250:126–136. <https://doi.org/10.1016/j.jneumeth.2015.01.010>
- Noirhomme Q, Lesenfants D, Gomez F, Soddu A, Schrouff J, Garraux G, Luxen A, Phillips C, Laureys S (2014) Biased binomial assessment of cross-validated estimation of classification accuracies illustrated in diagnosis predictions. *Neuroimage Clin* 4:687–694. <https://doi.org/10.1016/j.jneuroim.2014.04.004>
- Ojala M, Garriga GC (2010) Permutation tests for studying classifier performance. *J Mach Learn Res* 11:1833–1863
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Methodol* 57(1):289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Raz N, Ghisletta P, Rodrigue KM, Kennedy KM, Lindenberger U (2010) Trajectories of brain aging in middle-aged and older adults: regional and individual differences. *Neuroimage* 51(2):501–511. <https://doi.org/10.1016/j.neuroimage.2010.03.020>

36. Celidoni M, Dal Bianco C, Weber G (2017) Retirement and cognitive decline: a longitudinal analysis using SHARE data. *J Health Econ* 56:113–125. <https://doi.org/10.1016/j.jhealeco.2017.09.003>
37. Rogers RL, Meyer JS, Mortel KF (1990) After reaching retirement age physical activity sustains cerebral perfusion and cognition. *J Am Geriatr Soc* 38(2):123–128. <https://doi.org/10.1111/j.1532-5415.1990.tb03472.x>
38. Rohwedder S, Willis RJ (2010) Mental retirement. *Int J Econ Perspect* 24(1):119–138. <https://doi.org/10.1257/jep.24.1.119>
39. Pires L, Leitão J, Guerrini C, Simões MR (2014) Event-related brain potentials in the study of inhibition: cognitive control, source localization and age-related modulations. *Neuropsychol Rev* 24(4):461–490. <https://doi.org/10.1007/s11065-014-9275-4>
40. Wild-Wall N, Falkenstein M, Hohnsbein J (2008) Flanker interference in young and older participants as reflected in event-related potentials. *Brain Res* 1211:72–84. <https://doi.org/10.1016/j.brainres.2008.03.025>
41. Anderer P, Semlitsch HV, Saletu B (1996) Multichannel auditory event-related brain potentials: effects of normal aging on the scalp distribution of N1, P2, N2 and P300 latencies and amplitudes. *Electroencephalogr Clin Neurophysiol* 99(5):458–472
42. Itier RJ, Taylor MJ (2004) Effects of repetition and configural changes on the development of face recognition processes. *Dev Sci* 7(4):469–487. <https://doi.org/10.1111/j.1467-7687.2004.00367.x>
43. Kropotov J, Ponomarev V, Tereshchenko EP, Müller A, Jäncke L (2016) Effect of aging on ERP components of cognitive control. *Front Aging Neurosci*. <https://doi.org/10.3389/fnagi.2016.00069>
44. Tomé D, Barbosa F, Nowak K, Marques-Texeira J (2015) The development of the N1 and N2 components in auditory oddball paradigms: a systematic review with narrative analysis and suggested normative values. *J Neural Transm* 122:375–391
45. Wendel K, Väistönen J, Seemann G, Hyttinen J, Malmivuo J (2010) The influence of age and skull conductivity on surface and subdermal bipolar EEG leads. *Comput Intell Neurosci* 2010:397272. <https://doi.org/10.1155/2010/397272>
46. Farina FR, Emek-Savaş DD, Rueda-Delgado L, Boyle R, Kiiski H, Yener G, Whelan R (2020) A comparison of resting state EEG and structural MRI for classifying Alzheimer's disease and mild cognitive impairment. *Neuroimage* 215:116795. <https://doi.org/10.1016/j.neuroimage.2020.116795>
47. Fröhlich S, Kutz DF, Müller K, Voelcker-Rehage C (2021) Characteristics of resting state EEG power in 80+-year-olds of different cognitive status. *Front Aging Neurosci* 13:675689. <https://doi.org/10.3389/fnagi.2021.675689>
48. Waszak F, Li SC, Hommel B (2010) The development of attentional networks: cross-sectional findings from a life span sample. *Dev Psychol* 46(2):337–349. <https://doi.org/10.1037/a0018541>
49. Anguera JA, Gazzaley A (2012) Dissociation of motor and sensory inhibition processes in normal aging. *Clin Neurophysiol* 123(4):730–740
50. Gomez Gonzalez CM, Clark VP, Fan S, Luck SJ, Hillyard SA (1994) Sources of attention-sensitive visual event-related potentials. *Brain Topogr* 7(1):41–51. <https://doi.org/10.1007/bf01184836>
51. Clayson PE, Larson MJ (2013) Psychometric properties of conflict monitoring and conflict adaptation indices: response time and conflict N2 event-related potentials. *Psychophysiology* 50(12):1209–1219. <https://doi.org/10.1111/psyp.12138>
52. van Dinteren R, Arns M, Jongsma MLA, Kessels RPC (2014) P300 Development across the lifespan: a systematic review and meta-analysis. *PLoS ONE* 9(2):e87347. <https://doi.org/10.1371/journal.pone.0087347>
53. Carp J, Park J, Hebrank A, Park DC, Polk TA (2011) Age-related neural differentiation in the motor system. *PLoS ONE* 6(12):e29411. <https://doi.org/10.1371/journal.pone.0029411>
54. Walther A, Nili H, Ejaz N, Alink A, Kriegeskorte N, Diedrichsen J (2016) Reliability of dissimilarity measures for multi-voxel pattern analysis. *Neuroimage* 137:188–200. <https://doi.org/10.1016/j.neuroimage.2015.12.012>
55. Bridwell DA, Cavanagh JF, Collins AGE, Nunez MD, Srinivasan R, Stober S, Calhoun VD (2018) Moving beyond ERP components: a selective review of approaches to integrate EEG and behavior. *Front Hum Neurosci* 12:106. <https://doi.org/10.3389/fnhum.2018.00106>
56. Winneke AH, Hübner L, Godde B, Voelcker-Rehage C (2019) Moderate cardiovascular exercise speeds up neural markers of stimulus evaluation during attentional control processes. *J Clin Med*. <https://doi.org/10.3390/jcm8091348>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com

Published Research Article III



Electrophysiological signatures of dedifferentiation differ between fit and less fit older adults

Christian Goetz¹ · Karin Mora² · Julia Kristin Stroehlein¹ · Franziska Katharina Haase¹ · Michael Dellnitz² · Claus Reinsberger¹ · Solveig Vieluf¹

Received: 28 February 2020 / Revised: 3 November 2020 / Accepted: 27 November 2020
© The Author(s) 2021

Abstract

Cardiorespiratory fitness was found to influence age-related changes of resting state brain network organization. However, the influence on dedifferentiated involvement of wider and more unspecialized brain regions during task completion is barely understood. We analyzed EEG data recorded during rest and different tasks (sensory, motor, cognitive) with dynamic mode decomposition, which accounts for topological characteristics as well as temporal dynamics of brain networks. As a main feature the dominant spatio-temporal EEG pattern was extracted in multiple frequency bands per participant. To deduce a pattern's stability, we calculated its proportion of total variance among all activation patterns over time for each task. By comparing fit ($N = 15$) and less fit older adults ($N = 16$) characterized by their performance on a 6-min walking test, we found signs of a lower task specificity of the obtained network features for the less fit compared to the fit group. This was indicated by fewer significant differences between tasks in the theta and high beta frequency band in the less fit group. Repeated measures ANOVA revealed that a significantly lower proportion of total variance can be explained by the main pattern in high beta frequency range for the less fit compared to the fit group [$F(1,29) = 12.572$, $p = .001$, partial $\eta^2 = .300$]. Our results indicate that the dedifferentiation in task-related brain activation is lower in fit compared to less fit older adults. Thus, our study supports the idea that cardiorespiratory fitness influences task-related brain network organization in different task domains.

Keywords Spatio-temporal coherent patterns · Electroencephalography · Dynamic mode decomposition · Cardiorespiratory fitness · Older adults

Introduction

Age-related changes in brain network activity are characterized by dedifferentiated and compensatory involvement of wider and more unspecialized brain regions during task completion which relates to a decline of sensory, motor and cognitive skills (Baltes and Lindenberger 1997; Park et al. 2004; Sala-Llonch et al. 2015). Despite the general

tendency towards a decline, brain activity and network interaction in older adults were shown to be highly dependent on the individual's lifestyle (Smith and Thelen 2003). One influencing factor might be cardiorespiratory fitness, which has the potential to diminish the described age effects in resting state networks (Voss et al. 2016). Here, we aimed to detect characteristics of brain network activity representing rest- and task-related specificity of information processing in elderly with different cardiorespiratory fitness levels.

Dedifferentiation was detected in task-related fMRI studies investigating reduced neural specialization as well as compensatory involvement of task relevant brain areas in older adults, which is considered part of dedifferentiation in this work (Sala-Llonch et al. 2015). Furthermore, it was shown that regional hyperactivation in older compared to younger participants was related to changes in both fMRI- and EEG-derived task-related functional networks

Supplementary information The online version of this article (<https://doi.org/10.1007/s11571-020-09656-9>) contains supplementary material, which is available to authorized users.

✉ Solveig Vieluf
vieluf@sportmed.upb.de

¹ Institute of Sports Medicine, Paderborn University, Warburger Str. 100, 33098 Paderborn, Germany

² Department of Mathematics, Paderborn University, Paderborn, Germany

(Geerligs et al. 2012, 2014). These changes also showed up across different tasks leading to a less task-specific activation of task-related brain network processes (Dennis and Cabeza 2011). Moreover, age-related changes seem to be multidimensional and complex. Chen et al. (2017) and Nobukawa et al. (2019) found age-related reorganization in the dynamics of brain networks, indicated by either higher complexity or irregularity in brain network patterns (Nobukawa et al. 2019) or changes in variations of functional connectivity over time (Chen et al. 2017). These findings might relate to increased background activity or neural noise throughout task execution (Hong and Rebec 2012). Increases in neuronal noise have been suggested as contributing factors to cognitive deficits and have been linked to neurobiological mechanisms (such as a decline in the dopaminergic neuromodulation) associated with a decrease in the distinction of cortical representations due to dedifferentiated brain activation (Li et al. 2001; Sala-Llonch et al. 2015).

Douw et al. (2014) related a higher interconnected functional modular topology of MEG derived resting state brain networks to cardiorespiratory fitness in middle aged participants. The findings indicate that cardiorespiratory fitness has an influence on brain networks and might bear the potential to counteract age-related changes, i.e., dedifferentiation and compensatory mechanisms. Indeed, Stillman et al. (2019) reviewed effects of physical activity and fitness on fMRI derived resting state brain networks in older adults and reported opposing effects on age-related changes. In this context, several authors reported an increase of connectivity within resting state networks in older adults due to cardiorespiratory fitness, which suggests less dedifferentiated brain functioning (Prakash et al. 2011; Voss et al. 2010, 2016). However, less is known about the influence of cardiorespiratory fitness on task-related brain network processes and dedifferentiation across different tasks in older adults. The influence of cardiovascular fitness on the decrease in specificity of information processing in the sensory, motor, and cognitive areas, in which age-related decline is reported, seems to be particularly relevant for everyday life and is yet unclear. Especially the investigation of dynamic processes of central information processing could provide new insights into this impact.

EEG allows to study brain dynamics with high temporal resolution capture changes of the temporal characteristics of functional networks. Due to the high complexity of age-related reorganization of brain activity and its interaction with cardiorespiratory fitness, we chose a holistic approach which combines the representation of spatial and temporal brain activity patterns. Dynamic mode decomposition (DMD) is an algorithm that was developed in the field of fluid dynamics (Rowley et al. 2009; Schmid and Sesterhenn 2008). It was recently applied to various other fields,

including neuroscientific data (Brunton et al. 2016; Casorso et al. 2019; Götz et al. 2018; Kunert-Graf et al. 2019; Vieluf et al. 2018). DMD combines the properties of spatial and temporal decomposition methods enlarging classical functional connectivity (FC) approaches based on bivariate connectedness between voxels or electrodes, which is frequently done in literature (Sporns 2013).

Utilizing DMD, we therefore aimed to study the influence of cardiorespiratory fitness on task-related brain network activity by assessing coherent spatio-temporal patterns of EEG in rest as well as tasks representing the sensory, motor, and cognitive domains, respectively. Albeit exploratory, we expected electrophysiological signatures of dedifferentiation, i.e., less task specificity in different tasks in fit compared to less fit individuals. In addition, we hypothesized that fit older adults have lower levels of neural noise, which translates into a higher prominence of dominant brain network patterns.

Materials and methods

The data were collected during an intervention study, which was registered at the German Clinical Trials Register (DRKS00014921) and took place at Paderborn University. The study protocol was approved by the ethics committee of the University of Muenster. Written informed consent to participate in the study was obtained by each participant before the experiments. No compensation was offered.

Participants

Participants were recruited via local newspaper and social media advertisements as well as by personal contact with organizations providing leisure activities for seniors. Participants were included if they (1) were above 60 years old, (2) free of diagnosed neurological or mental diseases and (3) right-handed. Half of the participants participated in a golf training and half continued their normal activities prior to the recording. In the context of this study golf was considered as part of the daily activities. All included participants reported subjective memory complaints in daily life but had no diagnosed form of dementia or its preclinical manifestation mild cognitive impairment, and were therefore considered healthy. All participants scored below 13 on a neuropsychological test battery (see Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-cog) Table 1). In sum, a total of 41 participants between 60 and 77 years of age participated in this study (age: 67 ± 4.16 , gender: 22 ♀, 19 ♂; see Table 1). All participants had normal or corrected to normal vision.

Screening

In preceding appointments, participants' cardiorespiratory fitness was assessed with a 6 min walking test (Enright 2003). Participants had to walk 6 min as far as possible with a fast and constant pace. The distance was assessed as marker of cardiorespiratory fitness as this test was shown to be a reliable and valid way to test physical endurance in older adults (Rikli and Jones 1998; Zhang et al. 2017). Prior to the EEG measurement and in relation to the domains of the main tasks, maximum voluntary contraction (MVC) as well as reaction time (RT) and tactile threshold (TT) were assessed. MVC measurements consisted of three 5 s lasting maximum precision grip trials with 60 s rest in between (Vieluf et al. 2013). RT was assessed with auditory stimuli, i.e., 60 spoken letters presented via speakers. As soon as a letter was presented, participants had to press a foot switch immediately with the right foot. The foot switch was positioned in a standardized position next to the right foot. Reaction time was measured in relation to stimulus onset (Voelcker-Rehage and Alberts 2007). The TT was detected on the non-dominant hand as a 2-point discrimination test. Participants were asked to distinguish between one-point and two-point stimulation. The distance between the stimulation points was successively increased by 1 mm starting at a minimum distance of 1 mm. TT was achieved when the participant could clearly distinguish 7 out of 10 stimulations presented (Finnell et al. 2004).

EEG experiments

All EEG measurements were recorded with an actiCap electrode cap and BrainAmp standard amplifiers (Brain Products, Munich, Germany). We recorded brain activity at 128 electrodes with a sampling rate of 500 Hz. Ground and reference electrodes were fixed at FPz and FCz, respectively. Impedances were kept below 15 k Ω .

At first, EEG was recorded four minutes in a rest condition in supine position with eyes closed. During the tasks, participants sat 80 cm in front of a screen (23'', 1920 × 1080 pixels at 60 Hz, AOC, Taipei, Taiwan). Their right arm rested comfortably on an armrest grasping a force transducer with index finger and thumb in precision grip (1022-C3-20 kg, SOEMER, Lennestadt-Elspe, Germany). Participants placed their left index finger on a braille device (P11, Metec Ag, Stuttgart Germany). In addition, speakers were placed approximately 50 cm behind the participants as well as a footswitch (StealthSwitch SS1R4 Pro USB, StealthSwitch, Highwood, IL, United States) next to their right foot. The tasks consisted of a motor task (force control), a cognitive task (auditive 2

back), and a sensory task (sensory oddball task) lasting 90 s. Each task was presented 2 times.

Motor task

Participants had to apply force to a force transducer with their index finger and thumb in precision grip to match a visually presented target (Voelcker-Rehage et al. 2006). The visual target was a line that moved from the right to the left on the screen for 90 s and changed level every 3 s in randomized order between heights that represents 10%, 20% and 30% of their individual MVC (see Fig. 1). Mean force level was set to 20% of participants MVC throughout the experiment. Participants were given online feedback on a screen 80 cm in front of them. Target line was displayed in blue whereas a red curser represented the applied force (see Fig. 1d).

Cognitive task

Participants were asked to listen to a sequence of letters presented via two speakers behind them and press the foot switch with the right foot, if a letter appeared again two letters later (2-back; see Fig. 1c). The sequences consisted of 60 letters in predefined randomized selections with 20% matching rate (Bopp and Verhaeghen 2018; Gajewski and Falkenstein 2014).

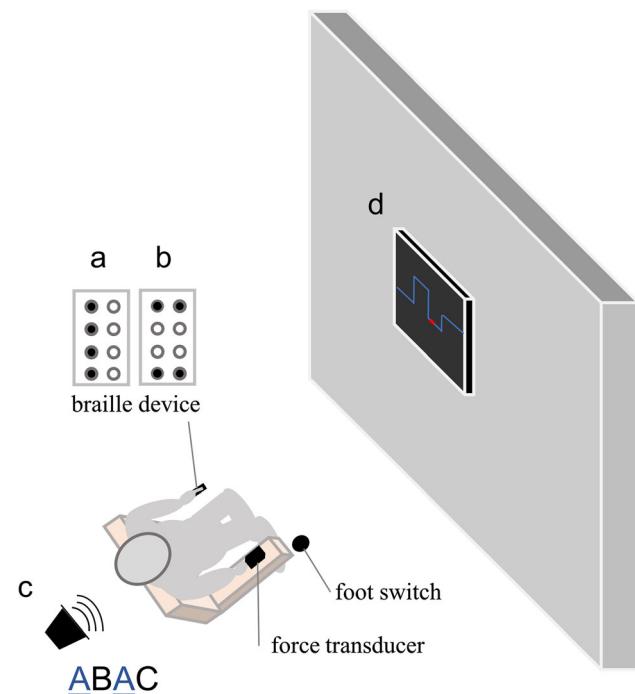


Fig. 1 Experimental setup of the sensory (a, b), n-back (c), and motor task (d)

Sensory task

The braille device presented two stimuli (see Fig. 1a and 1b) to the participants with a randomized inter stimulus interval of 0.8–1.2 s (mean 1 s) and a duration of 0.5 s each. In this passive oddball design no response was required from the participants. Pattern (a) was set to 80% occurrence and pattern (b) to 20% (Reuter et al. 2012, 2013, 2014).

Data analysis

Data analysis was performed with BrainVision Analyzer 2.1.2 (Brain Products, Munich, Germany), MATLAB 2018b (Mathworks, Natick, MA, United States), EEGLAB 14.1.2 (Brunner et al. 2013), as well as Brainstorm 3.4 (Tadel et al. 2011) software packages. EEG data were cut from task onset to offset. Next, data were filtered with a zero-phase shift 4th order Butterworth filter from 1 to 30 Hz and an additional notch filter at 50 Hz implemented in Brain Vision Analyzer 2.1.2 was applied. After down sampling to 200 Hz, data were inspected for bad channels as well as electrode bridges using Alschuler's et al. (2014) algorithm, which is based on the electrical distance of neighboring electrodes. To avoid false positives, coherence between adjacent channels was additionally checked for each pair of channels, where a value of 1 indicates complete coherence and a value of 0 indicates complete incoherence. Bad channels or channel pairs identified by Alschuler's et al. (2014) algorithm and with a coherence above 0.97 were rejected. Participants whose rejection rate was above 15% or exhibited a bridged reference were excluded from further analysis. Seven participants had to be excluded here. From the remaining participants, we excluded on average 4% of the electrodes per participant. Then, data was re-referenced to common average and segmented into non-overlapping segments of 150 data points (0.75 s). Before performing an ICA on the re-concatenated data in EEGLab (AMICA, Palmer et al. 2011) for artifact rejection, segments containing non stereotypical artifacts identified by visual inspection were excluded. ICA components reflecting unambiguous artifacts such as eye blinks, heartbeat or muscle activity were marked as bad (mean rejection rate: $19 \pm 5\%$) and excluded. After reconstruction of the signals all rejected channels were interpolated using spherical spline interpolation (Perrin et al. 1989) and checked again visually for segments containing artifacts.

In order to extract spatio-temporal coherent patterns within the EEG we decomposed the preprocessed data segments with the exact DMD algorithm described in

Brunton et al. (2016) after Tu et al. (2014). Specifically, we constructed the two data matrices $X \in \mathbb{R}^{n \times (m-1)}$ and $X' \in \mathbb{R}^{n \times (m-1)}$, where X is the original data matrix and X' is the data matrix shifted by one time point containing $m - 1$ datapoints of n electrodes. Moreover, the relationship between these two matrices can be expressed with a linear operator A describing the underlying dynamical process such that

$$X' = AX.$$

The DMD of X is then obtained via eigen-decomposition of A . In other words by analyzing the relationship between X and X' in a given time window (see Brunton et al. 2016 for details of the algorithm) it is possible to approximate the linked spatial and temporal characteristics. Therefore, we can obtain an approximation \hat{X} of the observed measurement X by defining a dynamical model $\hat{X} = \Phi \exp(\Omega t)z$.

The matrix Φ contains the dynamic modes, i.e., eigenvectors of A (see Theorem 1 in Tu et al. 2014) and $\Omega = \frac{\log(A)}{\Delta t}$ reveals the dynamics of the system, where the diagonal matrix A contains the DMD eigenvalues, i.e., eigenvalues of A , on its diagonal. The variable t denotes time whereby $\Delta t = 0.005$ s. The constant z can be obtained from the initial conditions $x_1 = \Phi z$. Moreover, it is possible to obtain the oscillation frequencies in cycles per second (Hz) from Ω :

$$f = \frac{\text{imag}(\Omega)}{2\pi}.$$

We conjecture that the dimension of the underlying system is larger than the one obtainable from the original data matrix X . Thus to avoid an underestimation, we propose to increase the dimension of the data matrix via the delay embedding method applied by Tu et al. (2014) and described in Brunton et al. (2016) as well as in Cohen (2018). In order to estimate the optimal values of the stacking factor h and the window size w we applied an error analysis on participant data of five randomly selected participants as described in Brunton et al. (2016). Based on these results we decided to choose a stacking depth of $h = 5$ and $w = 150$ data points.

To obtain the influence of each electrode in the DMD mode, the absolute DMD values associated with the θ (4 to < 7 Hz), α (7 to < 12 Hz), low β (12 to < 16 Hz) and high β (16 to < 30 Hz) frequencies were determined obtaining DMD mode magnitudes. Each DMD mode reflects a spatially coherent structure associated with a certain dynamic behavior, i.e., growth or decay and oscillation, where

magnitude indicates the element's (electrode) involvement in the mode. Next, we calculated a singular value decomposition (SVD) over all frequency specific modes over all time windows per participant to reduce the dimension of all time windows to its main features. Besides we calculated the proportion of variance among all activation patterns over time for each SVD mode:

$$R^2 = \frac{s_i^2}{\sum_j s_j^2} \times 100,$$

where s_j are the singular values obtained via SVD and s_i denotes the i -th singular value corresponding to the i -th SVD mode. Since the first SVD mode explained most of the variance of the activation patterns in all subjects ($M = 81.7\% \pm 1.1\%$) we extracted only the first mode to obtain the dominant feature of all DMD modes (DMD main mode) in all analyzed time windows. Moreover, we calculated the proportion of variance among all activation patterns explained by the obtained DMD main mode. The higher the proportion the more variation among all activation patterns during task completion is captured by this dominant mode which is considered representative for stability or prominence of this pattern.

As DMD is closely related to source separation, we used sLoreta to visualize the obtained spatial DMD maps in source space to demonstrate its relation. We created a three layer Boundary Element (BEM) forward model using the Brainstorm toolbox based on a template of the McConnell Brain Imaging Center (MNI/ICBM152) and fitted the electrodes using recorded individual electrode positions per subject (Fonov et al. 2009; Gramfort et al. 2010; Pascual-Marqui 2002; Tadel et al. 2011).

Statistical analysis

All statistical analyses were conducted using SPSS 24 for Windows (IBM, Armonk, NY, United States), Matlab 2018b (Mathworks, Natick, MA, United States) with the additional Brainstorm package (Tadel et al. 2011) as well as R with RStudio (version 1.1.456) and the additional cramer-package (Baringhaus and Franz, 2004). To test differences between fit and less fit participants we divided them based on their performance in the 6-min walking test with a median split for men and women separately to account for gender differences (Bohannon 2007) in two groups. In order to characterize brain network activity and test for task specificity of the obtained DMD maps we compared the task- and rest-specific DMD mode maps pairwise with a permutation t -test for dependent measurements. The permutation approach was chosen as a nonparametric alternative and offers the advantage of calculating the exact test statistic. For this we chose a Monte

Carlo approach with 10,000 randomizations, i.e., all possible values of the test statistics have been determined with a 10,000-fold random reordering of the data, the distribution of the t -test statistics under the null hypothesis (Maris 2012; Maris and Oostenveld 2007). By choosing this nonparametric statistic, we intended to account for the high dimensionality of the EEG data causing possible inaccuracies in test assumption requirements. By using an exact test statistics we intended a reduction of type I and type II errors (Maris and Oostenveld 2007). All p values were corrected with false discovery rate to account for type I errors (Benjamini and Hochberg 1995). For a statistical evaluation of the group differences, we exploratively compared the multivariate distribution of obtained t -values in each frequency band with Cramér tests between the groups. This nonparametric two sample test works on Euclidean interpoint distances to test the equality of the underlying distributions (Baringhaus and Franz, 2004).

We further analyzed the variance of all DMD modes explained by the DMD main mode with repeated measurement analysis of variance (ANOVA) with the between factor group (2; fit, less fit) and the within factor task (4; rest, motor, cognitive, sensory) for each frequency band (θ , α , low β , high β). Shapiro-Wilk and Box tests showed no violation of normal distribution and homogeneity of error variances (all $p > .05$). There was homogeneity of the error variances, as assessed by Levene's test except for the variables in the theta and alpha band during n-back and sensory tasks. We nevertheless report those values, as ANOVA was shown to be a robust test statistics in groups with almost the same size larger than 10 (Box 1954; Schmider et al. 2010). Significant interactions and main effects were followed by Bonferroni corrected pairwise comparisons. Independent t -tests and Mann-Whitney U tests, in case of violation of normal distribution, were used to check for differences in characteristics between the groups. The level of significance was set a priori to $\alpha = 0.05$ for all tests.

Results

The personal characteristics of the final sub-sample are shown in Table 1. Mann-Whitney U tests and independent t -tests showed no significant differences between the groups of fit and less fit participants for personal characteristics or screening variables, except for the walked distance in the six minute walking test (fit: $M = 765$ m, $SD = 127$ m; less fit: $M = 549$ m, $SD = 100$ m, $t(29) = -5.25$, $p < .001$).

Table 1 Comparison of characteristics of the fit and less fit group

Parameter	Less fit N = 16	Fit N = 15	Statistical value	p-value
Sex	8♀, 8♂	8♀, 7♂		
Age (years)	M = 68.56, SD = 4.13	M = 66.26, SD = 3.67	t = 1.637	.114
Height (cm)	M = 173.75, SD = 10.93	M = 171.93, SD = 10.02	t = 0.481	.634
Weight (kg)	M = 78.38, SD = 15.06	M = 79.60, SD = 10.76	t = -.259	.898
ADAS-Cog	M = 7.13, SD = 2.5	M = 5.87, SD = 3.14	t = 1.230	.225
Tactile Threshold (mm)	Median = 3, SD = 0.72	Median = 3, SD = 0.70	U = 103	.459
RT (s)	M = 1.03, SD = 0.29	M = 0.91, SD = 0.19	t = 1.290	.207
MVC (N)	M = 53.91, SD = 18.30	M = 62.32, SD = 21.45	t = -1.177	.249
PASE score	Median = 144.9, SD = 46.75	Median = 161.35, SD = 50.40	U = 93	.600
Six-minute walking (m)	M = 563.29, SD = 115.6 (norm: ♀ ≥ 60 years: 475 (95%-CI [448, 503])/♂ ≥ 60 years: 560 (95%-CI [511, 609]))	M = 750.21, SD = 138.5	t = -5.249	.000*

*Indicates a significant difference. Norm values of six-minute walking test from Bohannon (2007). MVC Maximum voluntary contraction, ADAS-cog Alzheimer's disease assessment scale-cognitive subscale, PASE physical activity scale for the elderly

DMD mode maps

Figure 2 illustrates the DMD topographical maps with an estimated source localization. Note that the source localization is only presented for visualization purposes here. Each DMD mode reflects a pattern of correlation in space at certain frequencies (Kutz et al. 2016). Low values indicate low contribution to the mode whereas high values indicate high contribution to that mode. Noticeable DMD main mode maps at rest showed frequency specific characteristics. The θ activity was dominant over frontal-central areas and showed highest values in the motor task. Dominant α activity over occipital regions during rest was suppressed and shifted to temporo-parietal areas. The low β activation showed the highest values at rest above the central areas and became more extensive in temporal direction during tasks. While the high β-band showed a similar central activation during rest, the activity in this frequency band changed towards a dominant activation over frontal with a shift towards centro-temporal areas that seemed to be most pronounced during the motor task (see Fig. 2).

Figure 3a-d illustrates the comparisons between the tasks for the fit and less fit groups separately in the analyzed frequency bands (θ, α, low β and high β) as topographic t-maps. For simplicity and clarity of presentation only minimum and maximum t- and corresponding p-values as well as their electrode position are presented in Tables 2 and 3. Means and standard deviations per electrode are available in the supplementary material (Online Resource 1).

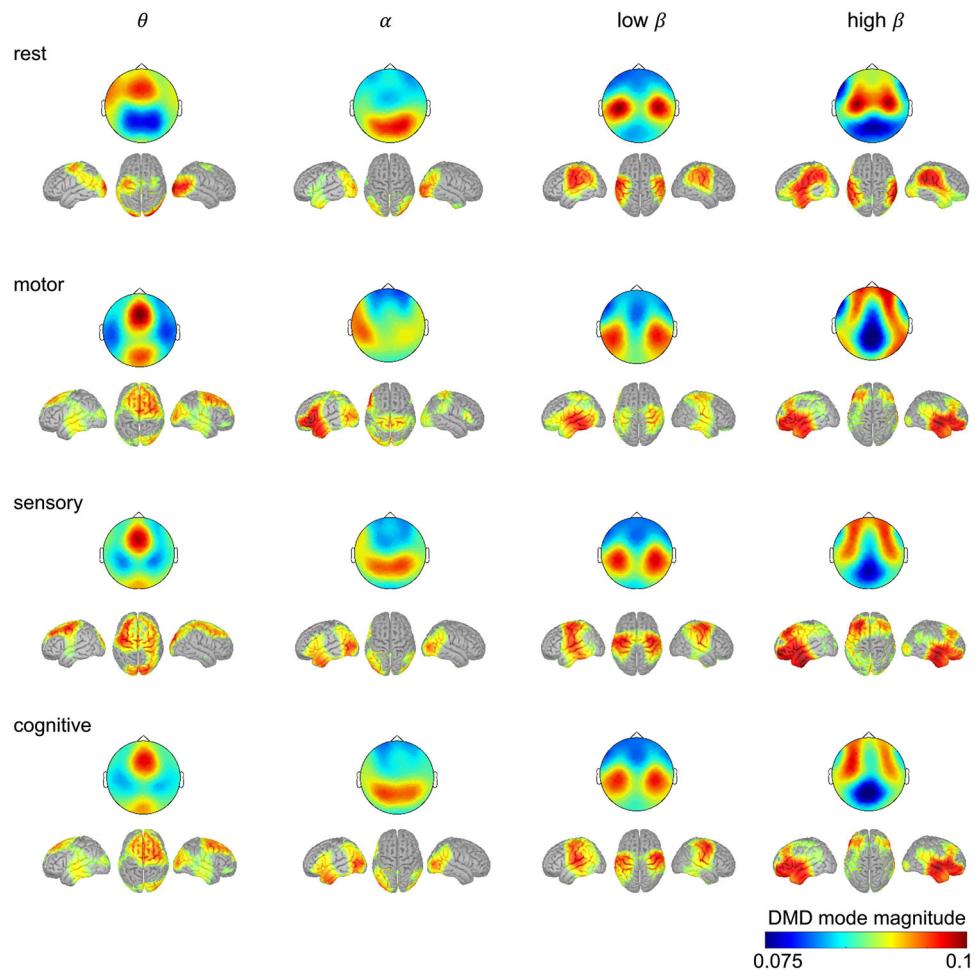
For both groups and frequency bands, DMD main mode maps differed between each task and rest with the strongest differences in the θ-band and in the high β-band. Both

groups θ-band main mode during task was characterized by higher activity over parieto-occipital areas and lower activity over fronto-temporal areas compared to rest. In contrast, in the high β-band, the differences between task execution and rest are marked by lower activation over (parieto-) central and higher activation over frontal and occipital areas in both groups.

Comparing the t-distributions with Cramér tests between the two groups revealed significant differences for all frequency bands (all $p < .001$). Studying the t-maps in group comparison, finer spatial differences can be observed. The differences between task and rest seem to be similar between the groups in the θ-band and in the high β-band. However, the differences between tasks and rest might be stronger in α- and low β-bands in the fit compared to the less fit group. This is especially observed in the α-band with lower parieto-occipital/frontal and higher temporal-central activation during motor task compared to the rest condition. Similar trends could be observed in the low β-band with higher parieto-occipital and lower central activation in the cognitive task compared to (see Fig. 3; Table 2).

There were likewise fewer pronounced differences between the task conditions with the highest expression between the motor task and the two other tasks in both groups. The differences, however, seemed to predominate in the fit group, especially in the θ-band and in the high β-band in the motor task compared to the sensory and cognitive tasks. In the fit group the θ-band is characterized by reduced activation of parieto-occipital and higher activation of temporo-parietal areas in the n-back task as well as higher temporo-frontal and a reduced activation of central and parietal areas of the sensory compared to the motor task. High β-band differences were present in form of

Fig. 2 DMD main mode features during rest with eyes closed and the three different task conditions (motor, sensory, cognitive) in the frequency ranges of θ (4 to < 7 Hz), α (7 to < 12 Hz), low (12 to < 16 Hz) and high β (16 to < 30 Hz) as well as their source representation. Each row represents a condition and each column represents a frequency band, thus there are four topographic maps per condition. Maps represent the mean over all participants



activation over central areas and sporadically decreased over frontal or occipital areas of the cognitive and sensory tasks compared to the motor task. There were likewise only small differences in both groups in task comparisons in the low β -band. Comparison of cognitive and sensory task revealed almost no differences in both groups (see Table 3).

Explained variance

Results of DMD main mode explained variance are illustrated in Fig. 3e–h. Repeated measures ANOVA revealed a significant effect of task in the θ -band [$F(3, 87) = 26.704, p < .001$, partial $\eta^2 = .479$] but none for group [$F(1, 29) = 3.174, p = .085$, partial $\eta^2 = .099$]. Post hoc tests revealed significant differences between the motor and the n-back task independent of group ($p = .023$).

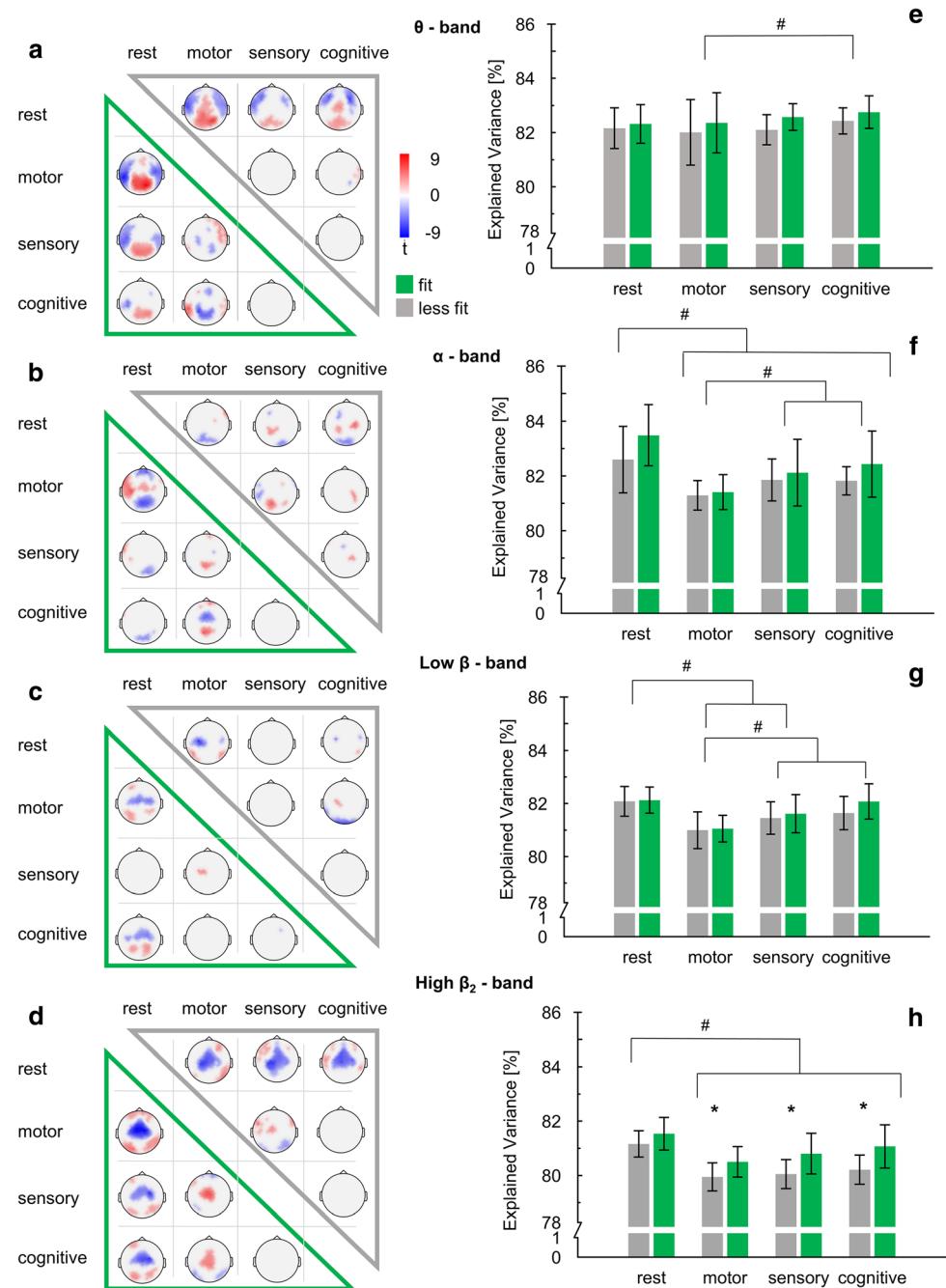
Moreover, there was a significant main effect of task in the α -band [$F(3, 87) = 4647, p = .005$, partial $\eta^2 = .138$] but no significant effect of group [$F(1, 29) = 2.79, p = .106$, partial $\eta^2 = .088$]. All tasks differed significantly from rest (sensory: $p < .001$; motor: $p < .001$; cognitive:

$p = .002$). Sensory and motor differed significantly ($p = .008$) as well as motor and n-back ($p < .001$).

In the low β -band there was a significant effect of task [$F(3, 87) = 31.399, p < .001$, partial $\eta^2 = .520$], but neither an effect of group [$F(1, 29) = 1.034, p = .318$, partial $\eta^2 = .030$] nor an interaction between task and group [$F(3, 87) = 1.204, p = .313$, partial $\eta^2 = .040$] were significant. The sensory task and motor task explained variance showed significantly lower values compared to rest ($p < .001$). Motor task was characterized by significantly lower values compared to the sensory task ($p < .001$) and cognitive task ($p = .001$) as well as higher values of the cognitive compared to the sensory task ($p = .020$).

Repeated measures ANOVA revealed a significant main effect of task [$F(3, 87) = 41.855, p < .001$, partial $\eta^2 = .600$] and group [$F(1, 29) = 12.572, p = .001$, partial $\eta^2 = .300$] in the high β -band. There was no significant interaction of task and group [$F(3, 87) = 2.004, p = .119$, partial $\eta^2 = .070$]. Bonferroni-corrected post hoc tests demonstrated significantly higher values in rest compared to all tasks (all $p < .001$).

Fig. 3 a–d Statistical *t*-maps of significant differences of DMD main mode features in θ , α , low- and high β between the conditions divided in fit (green) and less fit (grey) participants. Only *t*-values with corresponding corrected *p* value $< .05$ are visualized. The opposite side of each group served as second term in the *t*-test for each group. **e–h** DMD main mode goodness of fit expressed in % of variance they explain as group mean and standard deviation. * indicates significant pairwise comparisons of the main effect group. # indicates significant pairwise comparisons of the main effect task



Discussion

Utilizing the multivariate analysis method dynamic mode decomposition (DMD), we aimed to find characteristics of task specific brain network patterns and their differences in fit compared to less fit elderly. We analyzed EEG data recorded during sensory, motor, and cognitive tasks in comparison to rest. DMD revealed frequency-dependent spatio-temporal scalp patterns that appear at rest and three different tasks. In fit older adults, these patterns showed trends of higher task specificity pointing to less

dedifferentiated brain activity. Furthermore, we found this pattern's proportion of total variance explained higher in fit participants which might indicate less neural noise during task execution. The results jointly support the idea that physical fitness reduces the impact of dedifferentiation.

We identified main features of frequency specific brain network dynamics expressed in task-specific EEG patterns. As DMD can be conceptualized as a combination of well-known spectral analysis methods like the Fourier Transform with the Fast Fourier Transform algorithm (FFT) and spatial decomposition with principle component analysis

Table 2 Minimum and maximum *t*-values as well as corresponding *p*-value and their electrode position (loc.) of the group specific comparison between task conditions and resting state condition

vs.	Rest											
	Fit						Less fit					
	<i>t</i> _{min}	<i>p</i>	Loc.	<i>t</i> _{max}	<i>p</i>	Loc.	<i>t</i> _{min}	<i>p</i>	Loc.	<i>t</i> _{max}	<i>p</i>	Loc.
θ												
Motor	– 8.64	< .001*	TP7	9.32	< .001*	PO4	– 6.52	.005*	FTT9h	7.32	.005*	PPO6h
Sensory	– 5.65	< .001*	TP7	5.34	< .001*	POO1	– 5.57	.009*	F5	3.79	.024*	PPO6h
Cognitive	– 5.70	.01*	CP5	4.58	.01*	PPO6h	– 8.23	.006*	FFC6h	4.19	.007*	Cz
α												
Motor	– 6.82	.003*	AFp2	6.56	.003*	CCP5h	– 4.25	.002*	PO3	5.00	.020*	F10
Sensory	– 4.67	.038*	O2	5.49	.038*	FT9	– 4.29	.023*	O2	4.43	.024*	FT9
Cognitive	– 4.12	.050*	POO1	4.24	.047*	FT9	– 4.61	.009*	O2	5.92	.009*	FT9
Low β												
Motor	– 5.10	.021*	C4	3.99	.030*	AFF5h	– 7.44	.006*	C3	4.92	.009*	P7
Sensory	– 5.83	.003*	T7	5.80	.003*	PO8	– 3.57	.129	C3	4.05	.129	P7
Cognitive	– 4.61	.013*	FCC4h	4.07	.013*	P4	– 5.49	.003*	C3	4.37	.030*	P8
High β												
Motor	– 9.73	.009*	C2	6.07	.003*	O1	– 7.09	.003*	CCP3h	4.80	.003*	PPO10h
Sensory	– 7.90	.010*	FCC4h	3.61	.017*	PO8	– 7.62	.003*	CP3	3.62	.013*	AF7
Cognitive	– 7.73	.002*	FCC4h	4.83	< .001*	PPO6h	– 6 – 57	.003*	CCP1h	4.95	.003*	TP8

*Indicates a statistically significant difference

(PCA) our results are comparable to well-known EEG-power characteristics modified throughout the tasks reflecting markers of frequency specific network interaction and modulation dominant throughout the tasks. However, in addition DMD links the assessed spatial and temporal properties providing a low-dimensional representation of the underlying time-variable complex system.

In examining the DMD derived EEG patterns we aimed to study selectivity of neural responses, or dedifferentiation, across rest and three different tasks. The dominant fronto-central θ activation found here has been discussed as being sensitive to cognitive involvement and was dominant in the task conditions compared to rest (Jensen and Mazaheri 2010; Onton et al. 2005; Siegel et al. 2012). Occipital α activity being suppressed in the task conditions in this study has been discussed as a marker for the suppression of the visual network during visual attention. In general, the activity of α could play a role in so-called gating by inhibition processes, i.e., the selective activation of task-relevant and suppression of task-irrelevant areas and networks (Jensen and Mazaheri 2010). Moreover, high β was discussed as a marker of large-scale coupling and sensorimotor information integration showing wider distribution in the tasks compared to the rest condition (Siegel et al. 2012). Overall, the expression of these patterns was different in rest and the three tasks and can be regarded as

reflection of task-specific network activation processes linked to aforementioned processes. Indeed, comparing these patterns across tasks both groups showed differences between rest and all tasks in all frequency bands, most dominant in θ - and high β -bands. The role of these two frequency bands could indicate the importance of cognitive control and large-scale coupling in maintaining functioning in older age pointing to compensatory activity as it was found by Vlahou et al. (2014) and Knyazev et al. (2015).

In line with the influence of cardiorespiratory fitness on resting state networks we hypothesized to find task specificity of these task-related brain network patterns less pronounced in less fit participants compared to fit participants. By comparing information processing between a motor, cognitive, and sensory task, we were able to study the influence of cardiorespiratory fitness on the reduction of neuronal specialization between tasks. These task domains should correspond to the domains in which age-related decline is reported and thus provide a high relevance for everyday life. In fact, the differences between the tasks seem to be more pronounced in the fit group compared to the less fit group. In other words, task-related neural responses of fit compared to less fit participants showed higher differentiation across task. This finding thus might indicate opposing effects of cardiorespiratory fitness on age-related reduction of neuronal dedifferentiation

Table 3 Minimum and maximum t -values as well as corresponding p -value and their electrode position (loc.) of the group specific comparison between all task conditions

vs.	Motor											
	Fit						Less fit					
	t_{\min}	p	Loc.	t_{\max}	p	Loc.	t_{\min}	p	Loc.	t_{\max}	p	Loc.
θ												
Sensory	– 3.29	.032*	PO4	4.69	.031*	FFT10h	– 3.36	.181	PPO6h	3.42	.181	T8
Cognitive	– 7.48	.002*	P3	6.63	.002*	TP7	– 4.13	.029*	P6	4.16	.029*	TP10
α												
Sensory	– 4.26	.006*	P9	5.25	.006*	PPO1h	– 4.44	.018	TPP7h	5.92	.018*	PPO1h
Cognitive	– 5.49	.006*	FCz	6.84	.006*	PPO1h	– 4.31	.031*	F10	4.98	.029*	CPP6h
Low β												
Sensory	– 3.12	.131	AF3	4.61	.043*	C1	– 3.61	.100	P9	4.03	.100	F9
Cognitive	– 3.10	.240	Fp2	3.76	.24	CP1	– 8.02	.004*	O9	4.43	.016*	CPP6h
High β												
Sensory	– 4.15	.013*	AFp2	6.10	.003*	Cz	– 3.81	.037*	PPO10h	6.41	.013*	FT7
Cognitive	– 4.13	.014*	PPO9h	4.50	.014*	TP7	– 4.68	.070	Iz	3.87	.070	F10
vs.	Sensory											
	Fit						Less fit					
	t_{\min}	p	Loc.	t_{\max}	p	Loc.	t_{\min}	p	Loc.	t_{\max}	p	Loc.
θ												
Cognitive	– 2.93	.292	FFC6h	3.11	.292	T7	– 3.66	.146	TP10	3.65	.146	FC4
α												
Cognitive	– 3.08	.606	FFC4h	2.33	.292	F9	– 4.05	.013	FFC1h	4.23	.013*	CCP4h
Low β												
Cognitive	– 3.61	.039*	FFC4h	3.16	.221	PO3	– 2.72	.155	Iz	3.47	.155	P8
High β												
Cognitive	– 3.73	.103	Cz	3.42	.103	PPO1h	– 3.28	.219	C1	2.36	.44	AF8

*Indicates a statistically significant difference

found in literature in motor, cognitive as well as visual tasks (Carp et al. 2011; Park et al. 2004; Rajah and D’Esposito 2005) and is in line with findings showing these effects on resting state networks (Voss et al. 2016). However, there were no clear significant differences between the sensory and the cognitive task in both groups. The undetectable difference indicates a similar degree of dedifferentiation that is independent of the cardiorespiratory fitness level. On the other side it has been shown that both tasks are highly dependent on the cognitive resources of working memory, this may have masked task-specific differences between the two tasks (Dehghan Nayyeri et al. 2019). It is also noteworthy that the clearest differences between the tasks were again found in the theta and high beta range. This could again show the influence of cognitive control processes and large-scale coupling in maintaining function in older adults and could also indicate that these processes are influenced by cardiorespiratory fitness.

Based on recent reports on age-related changes in the dynamics of reorganization processes linking increased levels of neural noise and dedifferentiation, we characterized each DMD main mode in terms of the proportion of the total variance of all activation patterns explained by the dominant pattern. By analyzing this as stability or prominence of the main DMD mode throughout task execution, we intended to take into account the changes in the dynamics of age-related reorganization reported in the literature (Chen et al. 2017; Li et al. 2001; Nobukawa et al. 2019). Furthermore, we aimed to investigate differences in this dynamic between subjects with different levels of cardiorespiratory fitness. Expecting lower levels of neural noise in the fit group, we found lower values in high β main mode explained variance in the less fit group independent of the task. Carefully interpreted, this could lead to a more targeted, task-specific information processing throughout task execution. With this finding, the present study can

contribute to previous literature describing effects of cardiorespiratory fitness on age-related cognitive changes and brain network dynamics. To the best of our knowledge, no study has yet investigated the influence of cardiorespiratory fitness on dynamic brain network processes in older adults in the context of dedifferentiation. The comparison of brain networks during a variety of tasks and analyzing the similarities and differences allows to gain insights into functional organizations of task relevant activities. Our results point to less neural noise throughout task execution in the fit group and extend the existing knowledge about the influence of cardiorespiratory fitness on dynamic processes and age-related changes in these processes.

On a methodological level, decomposing EEG spatial temporal dynamics with DMD is dependent on delay embedding, i.e., stacking factor, as well as number of electrodes and chosen window size. We picked the parameters according to an error analyses over several participants to choose the optimal parameters. Moreover, delay embedding was used by other authors before (Brunton et al. 2016; Cohen 2017). As we measured high density EEG, we decided to check for bridging artifacts. Therefore, we used an algorithm proposed by Alschuler et al. (2014) which identifies bridges based on the electrical distance distribution of the signals. However, Alschuler et al. (2014) note that this algorithm might be too conservative. We therefore decided to double check this step and calculated the coherence between neighboring channels.

In order to achieve a high level of comparability between tasks and groups of participants, we decided to reduce dimensionality of the variables with SVD and chose the most important characteristic, i.e., the first principle component. We decided not to study further components since we describe a coarse phenomenon such as dedifferentiation.

Moreover, we used the 6-min walking test to asses cardiorespiratory fitness instead of VO_{2max} measurements. This test is highly influenced by motivational aspects. However, there is a strong relation between the 6-min walking test and VO_{2max} and it is used as a common standard for indirect measurement of cardiorespiratory performance (Zhang et al. 2017) and the practicability as well as the motivation of the participants was very high for the selected test. Although all participants can be seen as rather fit in comparison to norm values reported in literature (Bohannon 2007) we found differences between fit and less fit in line with dedifferentiation. When measurements took place, participants had either learned golf within the last 22 weeks or continued with their normal daily activities. The golf training could have obviously influenced the results. In this study we were more interested in the long-term effects associated with cardiorespiratory fitness, reflected by the 6-min walking test, than in short-term

effects caused by golf training. Moreover, Voss et al. (2016) pointed out that the influence of cardiorespiratory fitness and short-term physical activity on brain networks are independent phenomena. Of course, longitudinal recordings and a more objective measurement of cardiorespiratory fitness and daily activity would extend presented findings. As this study was part of a bigger intervention study, sample size was fixed a priori. The primary outcome of this study was the ADAS-Cog based on the randomized controlled trial of Lautenschlager et al. (2008). The standardized mean difference for ADAS-Cog was -1.22. The drop-out rate was set to 20%. It was estimated that a sample size of 46 participants (23 in each group) would provide 95% power for detecting a significant group difference. A healthy young control group would be beneficial in order to categorize our findings.

Conclusion

In applying DMD to continuous EEG recordings during rest and three different tasks, we considered both topological properties and the temporal dynamics of task-related brain networks. Thus, we identified electrophysiological signatures of age-related brain reorganization processes in fit and less fit older adults. Fit participants showed higher task specificity, i.e., more differentiated brain activation patterns, as well as higher prominence of these patterns, indicating less neural noise throughout task execution. Our findings support the idea that physical fitness manifests in task-related brain network activation patterns that are in line with reduced dedifferentiation in older adults.

Acknowledgements We thank Franziska van den Bongard and Roman Gaidai for support during data collection.

Author's contribution CG, JKS, FKH, CR and SV contributed to the study conception and design. CG and FKH set up the experiments. Data collection were performed by CG, JKS and FKH. CG and KM analyzed data. All authors interpreted results, drafted parts of the work, approved the final version of the manuscript, and agreed to be accountable for all aspects of the work.

Funding Open Access funding enabled and organized by Projekt DEAL. The study was supported within the framework of the equality concept of the Paderborn University through funding line 2.

Data availability Data and material are available from the corresponding author, SV, upon reasonable request.

Compliance with ethical standards

Conflict of interest Authors declare to have no competing interests.

Code availability Custom code is available from the corresponding author, SV, upon reasonable request.

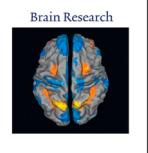
Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alschuler DM, Tenke CE, Bruder GE, Kayser J (2014) Identifying electrode bridging from electrical distance distributions: a survey of publicly-available EEG data using a new method. *Clin Neurophysiol* 125:484–490. <https://doi.org/10.1016/j.clinph.2013.08.024>
- Baltes P, Lindenberger U (1997) Emergence of a powerful connection between sensory and cognitive functions across the adult life span: a new window to the study of cognitive aging? *Psychol Aging* 12:12–21. <https://doi.org/10.1037/0882-7974.12.1.12>
- Baringhaus L, Franz C (2004) On a new multivariate two-sample test. *J Multivar Anal* 88(1):190–206. [https://doi.org/10.1016/S0047-259X\(03\)00079-4](https://doi.org/10.1016/S0047-259X(03)00079-4)
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol* 57:289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Bohannon RW (2007) Six-minute walk test: a meta-analysis of data from apparently healthy elders. *Top Geriatr Rehabil* 23:155–160. <https://doi.org/10.1097/01.TGR.0000270184.98402.ef>
- Bopp KL, Verhaeghen P (2018) Aging and n-back performance: a meta-analysis. *J Gerontol B Psychol* 25:2252. <https://doi.org/10.1093/geronb/gby024>
- Box GEP (1954) Some theorems on quadratic forms applied in the study of analysis of variance problems. II. Effects of inequality of variance and of correlation between errors in the two-way classification. *Ann Math Stat* 25:484–498. <https://doi.org/10.1214/aoms/1177728717>
- Brunner C, Delorme A, Makeig S (2013) Eeglab: an open source matlab toolbox for electrophysiological research. *Biomed Tech (Berl)*. <https://doi.org/10.1515/bmt-2013-4182>
- Brunton BW, Johnson LA, Ojemann JG, Kutz JN (2016) Extracting spatial-temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition. *J Neurosci Methods* 258:1–15. <https://doi.org/10.1016/j.jneumeth.2015.10.010>
- Carp J, Park J, Hebrank A, Park DC, Polk TA (2011) Age-related neural dedifferentiation in the motor system. *PLoS ONE* 6:e29411. <https://doi.org/10.1371/journal.pone.0029411>
- Casorso J, Kong X, Chi W, van de Ville D, Yeo BT, Liégeois R (2019) Dynamic mode decomposition of resting-state and task fMRI. *NeuroImage* 194:42–54. <https://doi.org/10.1016/j.neuroimage.2019.03.019>
- Chen Y, Wang W, Zhao X, Sha M, Liu YN, Zhang X, Ma J, Ni H, Ming D (2017) Age-related decline in the variation of dynamic functional connectivity: a resting state analysis. *Front Aging Neurosci* 9:203. <https://doi.org/10.3389/fnagi.2017.00203>
- Cohen MX (2017) Comparison of linear spatial filters for identifying oscillatory activity in multichannel data. *J Neurosci Methods* 278:1–12. <https://doi.org/10.1016/j.jneumeth.2016.12.016>
- Cohen MX (2018) Using spatiotemporal source separation to identify prominent features in multichannel data without sinusoidal filters. *Eur J Neurosci* 48:2454–2465. <https://doi.org/10.1111/ejn.13727>
- Dehghan Nayyeri M, Burgmer M, Pfeiderer B (2019) Impact of pressure as a tactile stimulus on working memory in healthy participants. *PLoS ONE* 14:e0213070–e0213070. <https://doi.org/10.1371/journal.pone.0213070>
- Dennis NA, Cabeza R (2011) Age-related dedifferentiation of learning systems: an fMRI study of implicit and explicit learning. *Neurobiol Aging* 32:2318.e17–2318.e30. <https://doi.org/10.1016/j.neurobiolaging.2010.04.004>
- Douw L, Nieboer D, van Dijk BW, Stam CJ, Twisk JWR (2014) A healthy brain in a healthy body: brain network correlates of physical and mental fitness. *PLoS ONE* 9:e88202. <https://doi.org/10.1371/journal.pone.0088202>
- Enright PL (2003) The six-minute walk test. *Respir Care* 48:783–785
- Finnell JT, Knopp R, Johnson P, Holland PC, Schubert W (2004) A calibrated paper clip is a reliable measure of two-point discrimination. *Acad Emerg Med* 11:710–714. <https://doi.org/10.1197/j.aem.2003.11.022>
- Fonov VS, Evans AC, McKinstry RC, Almli CR, Collins DL (2009) Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage* 47:S102. [https://doi.org/10.1016/S1053-8119\(09\)70884-5](https://doi.org/10.1016/S1053-8119(09)70884-5)
- Gajewski PD, Falkenstein M (2014) Age-related effects on ERP and oscillatory EEG-dynamics in a 2-back task. *J Psychophysiol* 28:162–177. <https://doi.org/10.1027/0269-8803/a000123>
- Geerligs L, Saliasi E, Maurits NM, Lorist MM (2012) Compensation through increased functional connectivity: neural correlates of inhibition in old and young. *J Cognit Neurosci* 24:2057–2069. https://doi.org/10.1162/jocn_a_00270
- Geerligs L, Maurits NM, Renken RJ, Lorist MM (2014) Reduced specificity of functional connectivity in the aging brain during task performance. *Hum Brain Mapp* 35:319–330. <https://doi.org/10.1002/hbm.22175>
- Götz C, Voelcker-Rehage C, Mora K, Reuter E-M, Godde B, Dellnitz M, Reinsberger C, Vieluf S (2018) Improved neural control of movements manifests in expertise-related differences in force output and brain network dynamics. *Front Physiol* 9:1540. <https://doi.org/10.3389/fphys.2018.01540>
- Gramfort A, Papadopoulou T, Olivi E, Clerc M (2010) OpenMEEG: opensource software for quasistatic bioelectromagnetics. *Biomed Eng Online* 9:45. <https://doi.org/10.1186/1475-925X-9-45>
- Hong S, Rebec G (2012) A new perspective on behavioral inconsistency and neural noise in aging: compensatory speeding of neural communication. *Front Aging Neurosci* 4:27. <https://doi.org/10.3389/fnagi.2012.00027>
- Jensen O, Mazaheri A (2010) Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Front Hum Neurosci* 4:186. <https://doi.org/10.3389/fnhum.2010.00186>
- Knyazev GG, Volf NV, Belousova LV (2015) Age-related differences in electroencephalogram connectivity and network topology. *Neurobiol Aging* 36:1849–1859. <https://doi.org/10.1016/j.neurobiolaging.2015.02.007>
- Kunert-Graf JM, Eschenburg KM, Galas DJ, Kutz JN, Rane SD, Brunton BW (2019) Extracting reproducible time-resolved resting state networks using dynamic mode decomposition. *Front Comput Neurosci* 13:75. <https://doi.org/10.3389/fncom.2019.00075>
- Kutz JN, Brunton SL, Brunton BW, Proctor JL (2016) Dynamic mode decomposition: data-driven modelling of complex systems. SIAM, Philadelphia
- Lautenschlager NT, Cox KL, Flicker L, Foster JK, van Boekxmeer FM, Xiao J, Greenop KR, Almeida OP (2008) Effect of physical activity on cognitive function in older adults at risk for

- Alzheimer disease: a randomized trial. *JAMA* 300:1027–1037. <https://doi.org/10.1001/jama.300.9.1027>
- Li S-C, Lindenberger U, Sikström S (2001) Aging cognition: from neuromodulation to representation. *Trends Cogn Sci* 5:479–486. [https://doi.org/10.1016/S1364-6613\(00\)01769-1](https://doi.org/10.1016/S1364-6613(00)01769-1)
- Maris E (2012) Statistical testing in electrophysiological studies. *Psychophysiol* 49:549–565. <https://doi.org/10.1111/j.1469-8986.2011.01320.x>
- Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164:177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>
- Nobukawa S, Kikuchi M, Takahashi T (2019) Changes in functional connectivity dynamics with aging: a dynamical phase synchronization approach. *NeuroImage* 188:357–368. <https://doi.org/10.1016/j.neuroimage.2018.12.008>
- Onton J, Delorme A, Makeig S (2005) Frontal midline EEG dynamics during working memory. *NeuroImage* 27:341–356. <https://doi.org/10.1016/j.neuroimage.2005.04.014>
- Palmer J, Kreutz-Delgado K, Makeig S (2011) AMICA: an adaptive mixture of independent component analyzers with shared components. https://sccn.ucsd.edu/~jason/amica_a.pdf. Accessed 3 December 2019
- Park DC, Polk TA, Park R, Minear M, Savage A, Smith MR (2004) Aging reduces neural specialization in ventral visual cortex. *Proc Natl Acad Sci USA* 101:13091–13095. <https://doi.org/10.1073/pnas.0405148101>
- Pascual-Marqui RD (2002) Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. *Method Find Exp Clin* 24(Suppl D):5–12
- Perrin F, Pernier J, Bertrand O, Echallier JF (1989) Spherical splines for scalp potential and current density mapping. *Electroencephalogr Clin Neurophysiol* 72:184–187. [https://doi.org/10.1016/0013-4694\(89\)90180-6](https://doi.org/10.1016/0013-4694(89)90180-6)
- Prakash RS, Voss MW, Erickson KI, Lewis JM, Chaddock L, Malkowski E, Alves H, Kim J, Szabo A, White SM, Wójcicki TR, Klamm EL, McAuley E, Kramer AF (2011) Cardiorespiratory fitness and attentional control in the aging brain. *Front Hum Neurosci* 4:229. <https://doi.org/10.3389/fnhum.2010.00229>
- Rajah MN, D’Esposito M (2005) Region-specific changes in prefrontal function with age: a review of PET and fMRI studies on working and episodic memory. *Brain* 128:1964–1983. <https://doi.org/10.1093/brain/awh608>
- Reuter E-M, Voelcker-Rehage C, Vieluf S, Godde B (2012) Touch perception throughout working life: effects of age and expertise. *Exp Brain Res* 216:287–297. <https://doi.org/10.1007/s00221-011-2931-5>
- Reuter E-M, Voelcker-Rehage C, Vieluf S, Winneke AH, Godde B (2013) A parietal-to-frontal shift in the P300 is associated with compensation of tactile discrimination deficits in late middle-aged adults. *Psychophysiol* 50:583–593. <https://doi.org/10.1111/pyps.12037>
- Reuter E-M, Voelcker-Rehage C, Vieluf S, Winneke AH, Godde B (2014) Extensive occupational finger use delays age effects in tactile perception—an ERP study. *Atten Percept Psychophys* 76:1160–1175. <https://doi.org/10.3758/s13414-014-0634-2>
- Rikli R, Jones J (1998) The reliability and validity of a 6-minute walk test as a measure of physical endurance in older adults. *J Aging Phys Activ* 6:363–375. <https://doi.org/10.1123/japa.6.4.363>
- Rowley CW, Mezić I, Bagheri S, Schlatter P, Henningson DS (2009) Spectral analysis of nonlinear flows. *J Fluid Mech* 641:115–127. <https://doi.org/10.1017/S0022112009992059>
- Sala-Llonch R, Bartrés-Faz D, Junqué C (2015) Reorganization of brain networks in aging: a review of functional connectivity studies. *Front Psychol* 6:663. <https://doi.org/10.3389/fpsyg.2015.00663>
- Schmid P, Sesterhenn J (2008) Dynamic mode decomposition of numerical and experimental data. In: APS division of fluid dynamics meeting abstracts, MR.007
- Schmid E, Ziegler M, Danay E, Beyer L, Buehner M (2010) Is it really robust? Reinvestigating the robustness of ANOVA against the normal distribution. *Eur J Res Methods Behav Soc Sci* 6:15–147. <https://doi.org/10.1027/1614-2241/a000016>
- Siegel M, Donner TH, Engel AK (2012) Spectral fingerprints of large-scale neuronal interactions. *Nat Rev Neurosci* 13:121–134. <https://doi.org/10.1038/nrn3137>
- Smith LB, Thelen E (2003) Development as a dynamic system. *Trends Cogn Sci* 7:343–348. [https://doi.org/10.1016/S1364-6613\(03\)00156-6](https://doi.org/10.1016/S1364-6613(03)00156-6)
- Sporns O (2013) Structure and function of complex brain networks. *Dialogues Clin Neurosci* 15:247–262
- Stillman CM, Donofry SD, Erickson KI (2019) Exercise, fitness and the aging brain: a review of functional connectivity in aging. *Arch Sci Psychol*. <https://doi.org/10.31296/aop.v3i4.98>
- Tadel F, Baillet S, Mosher JC, Pantazis D, Leahy RM (2011) Brainstorm: a user-friendly application for MEG/EEG analysis. *Comput Intell Neurosci* 2011:879716. <https://doi.org/10.1155/2011/879716>
- Tu J, Rowley C, Luchtenburg D, Brunton S, Kutz J (2014) On dynamic mode decomposition: theory and applications. *JCD* 1:391–421. <https://doi.org/10.3934/jcd.2014.1.391>
- Vieluf S, Godde B, Reuter E-M, Voelcker-Rehage C (2013) Age-related differences in finger force control are characterized by reduced force production. *Exp Brain Res* 224:107–117. <https://doi.org/10.1007/s00221-012-3292-4>
- Vieluf S, Mora K, Götz C, Reuter E-M, Godde B, Dellnitz M, Reinsberger C, Voelcker-Rehage C (2018) Age- and expertise-related differences of sensorimotor network dynamics during force control. *Neurosci* 388:203–213. <https://doi.org/10.1016/j.neuroscience.2018.07.025>
- Vlahou EL, Thurm F, Kolassa I-T, Schlee W (2014) Resting-state slow wave power, healthy aging and cognitive performance. *Sci Rep* 4:5101. <https://doi.org/10.1038/srep05101>
- Voelcker-Rehage C, Alberts JL (2007) Effect of motor practice on dual-task performance in older adults. *J Gerontol B Psychol Sci Soc Sci* 62:P141–P148. <https://doi.org/10.1093/geronb/62.3.P141>
- Voelcker-Rehage C, Stronge AJ, Alberts JL (2006) Age-related differences in working memory and force control under dual-task conditions. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn* 13:366–384. <https://doi.org/10.1080/138255890969339>
- Voss MW, Prakash RS, Erickson KI, Basak C, Chaddock L, Kim JS, Alves H, Heo S, Szabo AN, White SM, Wójcicki TR, Mailey EL, Gothe N, Olson EA, McAuley E, Kramer AF (2010) Plasticity of brain networks in a randomized intervention trial of exercise training in older adults. *Front Aging Neurosci* 2:32. <https://doi.org/10.3389/fnagi.2010.00032>
- Voss MW, Weng TB, Burzynska AZ, Wong CN, Cooke GE, Clark R, Fanning J, Awick E, Gothe NP, Olson EA, McAuley E, Kramer AF (2016) Fitness, but not physical activity, is related to functional integrity of brain networks associated with aging. *NeuroImage* 131:113–125. <https://doi.org/10.1016/j.neuroimage.2015.10.044>
- Zhang Q, Lu H, Pan S, Lin Y, Zhou K, Wang L (2017) 6MWT performance and its correlations with VO₂ and handgrip strength in home-dwelling mid-aged and older Chinese. *Int J Environ Res Public Health* 14:473. <https://doi.org/10.3390/ijerph14050473>

Published Research Article IV



Classification characteristics of fine motor experts based on electroencephalographic and force tracking data

R. Gaidai^{a,1}, C. Goelz^{a,1}, K. Mora^{b,c}, J. Rudisch^d, E.-M. Reuter^e, B. Godde^f, C. Reinsberger^a, C. Voelcker-Rehage^d, S. Vieluf^{a,g,*}

^a Institute of Sports Medicine, Paderborn University, Paderborn, Germany

^b Remote Sensing Centre for Earth System Research, Leipzig University, Leipzig, Germany

^c German Centre for Integrative Biodiversity Research (iDiv), Leipzig, Germany

^d Department of Neuromotor Behavior and Exercise, Institute of Sport and Exercise Sciences, University of Münster, Münster, Germany

^e Department of Sport and Health Sciences, Technical University of Munich, Munich, Germany

^f Department of Psychology & Methods, Jacobs University Bremen, Bremen, Germany

^g Division of Epilepsy and Clinical Neurophysiology, Department of Neurology, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA

ARTICLE INFO

Keywords:

Fine motor control
Dynamic mode decomposition
Machine learning
Decoding
Expertise

ABSTRACT

The application of machine learning techniques provides a data-driven approach for a deeper understanding of the development and expressions of expertise. In extension to the common procedure of comparing experts' and novices' performances in expertise-domain-related tasks we applied conventional classification algorithms. We distinguished between tasks for each participant and between groups, i.e., experts or novices, based on electroencephalographic (EEG) activity patterns and force output variables during four different force modulation tasks. The tasks under investigation involved sinusoidal and steady force tracking tasks, which were performed with the left and right hand. Classification of tasks based on EEG patterns as well as force output was possible with high accuracy in novices and experts, whereas classification of group membership, i.e., experts or novices, was at chance level. In follow-up analyses, we found a high degree of individuality in the EEG patterns of the experts, implying the long-term development of specialized central processing during fine motor tasks in fine motor experts. Taken together, the results suggest that continuous practice in the work context leads to the development of a highly individual and task-specific central control pattern.

1. Introduction

Many years of professional experience in the work context contribute to the development of an outstanding and highly automatized performance level, i.e., expertise in domain-specific areas (Ericsson et al., 2006). A common approach to assess expertise-related differences is to compare experts' and novices' performances in expertise-domain-related laboratory tasks. In this context, fine motor experts have often been studied because their respective domains can be represented under laboratory conditions. Experts' compared to novices' behavioral performance was observed to be more precise and executed faster in precision mechanics (Vieluf et al., 2012; Vieluf et al., 2013; Vieluf et al.,

2018), musicians (Krampe and Ericsson, 1996; Krampe, 2002) and surgeons (Law et al., 2004) when the task under investigation corresponds to the experts' domain. Moreover, experts' performance was less variable in terms of intra- and intertrial variability and more complex (Komar et al., 2015; Vieluf et al., 2018). On the neuronal level a comparison of experts and novices shows manifold differences in both structure and function of certain brain areas and their complex interaction, i.e., network behavior, pointing to more efficient information processing (Binder et al., 2017; Götz et al., 2018; Vieluf et al., 2018).

Assessing differences of central processing and behavior, experts and novices can be classified by using machine learning techniques. Such data-driven approaches allow us to build models based solely on

Abbreviations: CSP, common spatial patterns; dev, mean deviation; DMD, dynamic mode decomposition; ICA, independent component analysis; int var, intertrial variability; LDA, linear discriminant analysis; LH, left hand; MVC, maximum voluntary contraction; RH, right hand; SVM, support vector machine; UMAP, uniform manifold approximation and projection; var, absolute variability.

* Corresponding author at: Paderborn University, Institute of Sports Medicine, Warburger Str. 100, 33098 Paderborn, Germany.

E-mail address: vieluf@sportmed.uni-paderborn.de (S. Vieluf).

¹ Both authors contributed equally.

measurements and to learn general principles (Brunton and Beyeler, 2019; Bzdok and Yeo, 2017). Machine learning, thus, complements the characterization of expertise-related differences using classical statistical approaches by automatically extracting generalizable components from the complex interaction of features. Such models find practical application in the automatic classification of individual performance states. In a previous study, we used EEG data to classify fine motor tasks. We found differences in classification performance between different age groups that were associated with neurophysiological differences between the groups (Goelz et al., 2021). Building upon this, decoding expertise on a group level as well as decoding fine motor tasks using machine learning methods could provide information about expertise and task-related characteristics of central information processing. Previous classification of experts was done based on different types of data and in different expertise contexts. Using support vector machines, experts and novices were classified based on their hand movements during a simulated surgical procedure (Watson, 2014). Likewise, it was possible to classify the level of expertise of neurosurgeons based on force application and acceleration of the instruments used during a neurosurgical procedure in virtual reality (Winkler-Schwartz et al., 2019). Artists and non-artists could be classified using the periodicity in the frequency spectra of EEG data recorded during visual observation and mental imagery of paintings (Shourie, 2016). In addition, three stages of expertise of goalkeepers in soccer could be classified with a high accuracy based on gaze behavior (Hosp et al., 2021). Moreover, classification was used to show individuality of finger movement patterns in expert flute players (Albrecht et al., 2014) as well as in timing and movement variability in pianists (Caramiaux et al., 2018).

It remains unclear whether such a classification can also be performed successfully using a task that reflects more generalized functions or skills related to specific fields of expertise such as fine motor skills. To the best of our knowledge, no study so far investigated whether differences in classification of fine motor tasks are also evident for groups with different expertise levels. A deeper understanding of this would help to

better characterize expertise-related differences in fine motor tasks and has practical implications for areas where data-driven models are used to automatically diagnose performance levels or classify tasks at the individual level.

In this study, novices and fine motor experts performed force tracking tasks, i.e., sine and steady force tracking with the left and right hand, while their brain activity was recorded with EEG. We classified the type of task based on behavioral and EEG parameters and investigated group differences in classification performance. Furthermore, we aimed to classify the level of expertise, i.e., membership in the group of experts or novices. Based on our previous results in groups of different ages, we expected to gain insight into neurophysiological characteristics of fine motor expert performance.

2. Results

2.1. Group differences in force tracking and EEG activity

Results of the force tracking performance of the sine and steady target forces (see Fig. 1A) with the right and left hands are summarized in Fig. 1C and Table 1. For simplicity, the tasks are referred to as steady right, sine right, steady left, and steady left. Performance of experts was more accurate (lower mean absolute deviation from target force level) in steady left compared to novices. There was no significant difference in mean deviation for steady right, sine left, or sine right. Variability within trials was smaller in experts compared to novices in steady left but not in steady right, sine right, or sine left. Intertrial variability was lower in experts in steady left (see Table 1).

Similar to our previous studies (Götz et al., 2018; Vieluf et al., 2018), we used dynamic mode decomposition (DMD) to calculate EEG markers representing brain network characteristics (Brunton et al., 2016) and extracted DMD mean modes per task and frequency band (θ , α , β_1 and β_2) from the EEG. Group differences in EEG activity are displayed in Fig. 1D as statistical t-maps. There were no group differences on EEG level

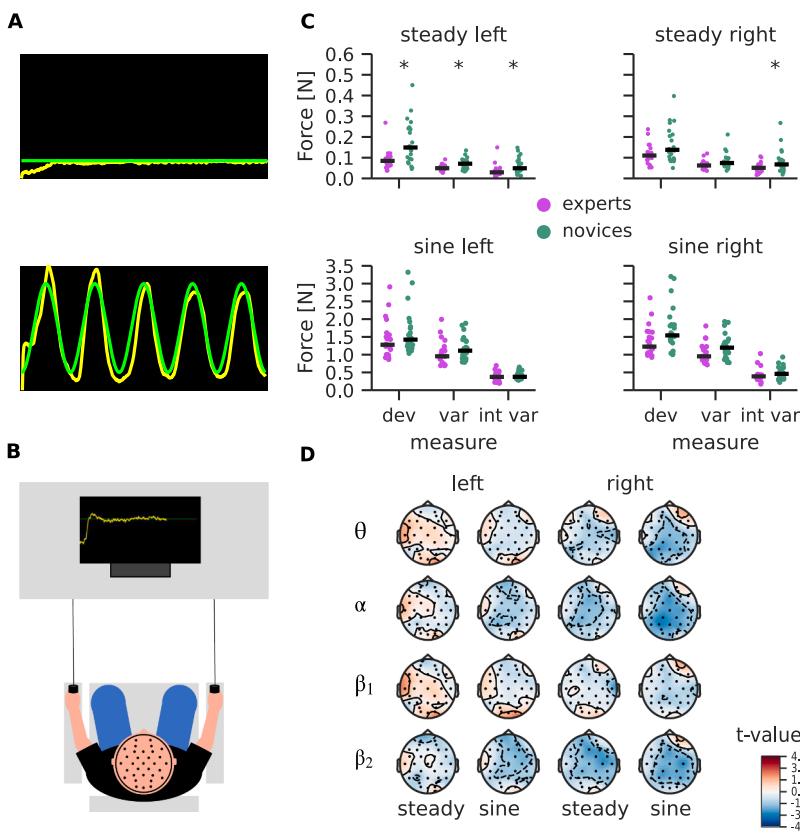


Fig. 1. A: Target force (green) and exemplary applied force (yellow). B: Experimental setup. C: Force tracking performance per group (dev: mean deviation, var: absolute variability, int var: intertrial variability). Each dot corresponds to one participant. Black bars reflect the median value. * indicates a p-value < 0.05. D: Statistical t-maps of the group comparison of fine motor experts vs. novices. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1

Group differences in force tracking. dev: mean deviation, var: absolute variability, int var: intertrial variability, Mdn: median, IQR: interquartile range, M: mean, SD: standard deviation, U: test statistic of the Mann-Whitney-U test (in case of not normally distributed data), t: test statistic of the t-test (in case of normally distributed data), p: probability value.

		Experts	Novices	Statistical values
dev [N]	Steady right	Mdn = 0.110 IQR = 0.128	Mdn = 0.138 IQR = 0.225	U = 158.000, p = 0.117
	Sine right	Mdn = 1.226 IQR = 1.632	Mdn = 1.542 IQR = 1.844	U = 154.000, p = 0.108
	Steady left	Mdn = 0.085 IQR = 0.104	Mdn = 0.149 IQR = 0.240	U = 112.000, p = 0.028*
	Sine left	Mdn = 1.279 IQR = 1.556	Mdn = 1.424 IQR = 1.775	U = 165.000, p = 0.122
var [N]	Steady right	Mdn = 0.062 IQR = 0.020	Mdn = 0.074 IQR = 0.031	U = 151.000, p = 0.107
	Sine right	Mdn = 0.953 IQR = 0.355	Mdn = 1.197 IQR = 0.403	U = 151.000, p = 0.107
	Steady left	M = 0.049 SD = 0.025	M = 0.071 SD = 0.037	t(41) = -2.990, p = 0.028*
	Sine left	Mdn = 0.954 IQR = 0.396	Mdn = 1.112 IQR = 0.447	U = 163.000, p = 0.121
int var [N]	Steady right	Mdn = 0.051 IQR = 0.038	Mdn = 0.067 IQR = 0.045	U = 139.000, p = 0.079
	Sine right	Mdn = 0.389 IQR = 0.123	Mdn = 0.459 IQR = 0.195	U = 163.000, p = 0.121
	Steady left	Mdn = 0.029 IQR = 0.016	Mdn = 0.049 IQR = 0.046	U = 121.000, p = 0.031*
	Sine left	Mdn = 0.374 IQR = 0.224	Mdn = 0.376 IQR = 0.101	U = 214.000, p = 0.689

comparing DMD mean mode magnitude between the groups in all force tracking tasks and frequency bands.

2.2. Task classification

For task classification we trained and tested a machine learning model for each participant individually. Results are displayed in Fig. 2 as confusion matrices and Table 2 including accuracy, F1, precision and recall for each group as mean over folds and group (see supplementary material Table S4 for individual results). Classification of tasks based on EEG features was above chance level in all individuals of both groups (chance level ~ 0.25). The same was true for classification based on force tracking. Neither classification based on EEG features nor based on force tracking features differed between the groups (see Table 2 for statistical values).

Besides models for each participant, we trained a model for task classification at group level for experts and novices respectively and tested it on participants not used during model training to test if generalizable patterns can be learned. Results are displayed in Table 2 including accuracy, F1, precision and recall as mean over folds. The performance of these models based on EEG data tended strongly towards chance (chance level ~ 0.25). The model trained and tested on data from the novices showed a slightly higher classification performance. Classification performance based on force tracking parameters decreased. Here, a better classification performance was found for experts compared to novices (see Table 2 for statistical values).

2.3. Classification of group membership

Model training and testing for group classification followed the same principle as the group level task classification, i.e., a model was trained for classification of group membership and tested on participants not used during model training. As with the classification of tasks, Fig. 3 and Table 3 summarize the performance of models for group membership classification as confusion matrices and by the metrics accuracy, F1, precision, and recall. Classification based on EEG features did not perform above chance (chance level ~ 0.5). The same was true for the classification based on force tracking features. Only classification based

on force tracking features of steady left was slightly above chance. Here, the classifier performed more accurately in classifying experts than novices. Combining features of both levels as well as different classifiers and features did not improve classification performance (see supplementary material Table S3).

2.4. Feature space characteristics

To examine the EEG feature space, we used uniform manifold approximation and projection (UMAP) (McInnes et al., 2018). Visualization of EEG feature space revealed patterns of individual characteristics in both groups. As shown in Fig. 4A and 4C, the EEG characteristics of each trial formed a cluster structure that can be assigned to individual participants. Comparing this individuality between the groups of experts and novices, it was more pronounced in the experts. This is illustrated by considering the distances of the cluster centroids as shown in Fig. 4B and 4D. What was striking here is an overall larger distance of the clusters in relation to each other which was present in a larger mean distance of the cluster centroids to each other in experts (see Fig. 4F). In addition, we compared the mean cluster size between the groups and found a smaller cluster size, i.e., more compact clusters, within the experts (see Fig. 4E). On the behavioral level, however, we did not find such a structure.

3. Discussion

In this study we gained insights into central and behavioral aspects of experts' fine motor performance by using machine learning techniques. Classification of different tasks at the individual level in both groups could be performed with high accuracy, indicating that indeed task-relevant features were chosen for classification. In contrast, the classification of group membership was at chance level. Follow-up analyses suggest that experts' brain activation patterns were characterized by higher individuality than those of novices. Overall, the results support the assumption that continuous practice in the work context results in the development of a highly individual and task-specific central control pattern.

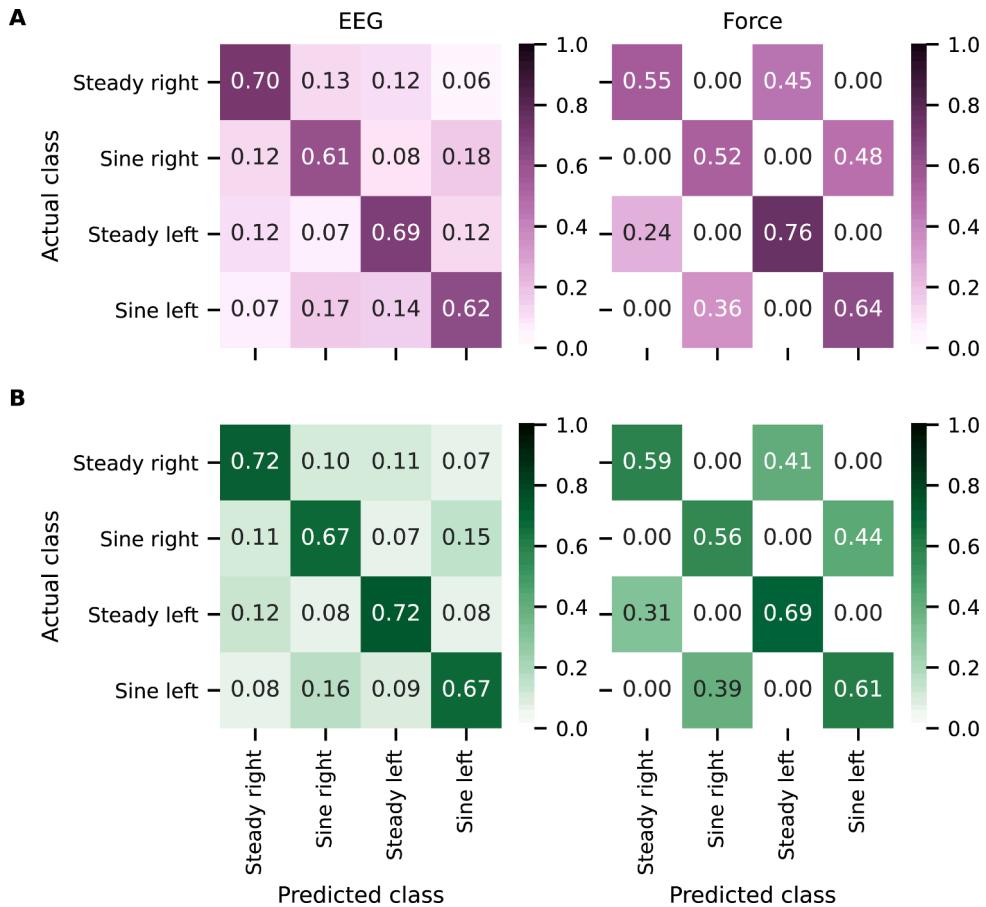


Fig. 2. Confusion matrices of task classification of fine motor experts (A) and novices (B) based on EEG features (left) and force tracking features (right). Mean over all participants.

3.1. Expertise-related differences in force tracking and EEG characteristics

Similar to our previous studies (Götz et al., 2018; Vieluf et al., 2012; Vieluf et al., 2013; Vieluf et al., 2018), we found differences in force tracking performance between experts and novices. Experts' force tracking performance was more precise and less variable within the steady left task. Intertrial variability between trials of the steady left and steady right tasks was lower in experts. Although we did not see this effect consistently across all tasks, a general pattern of higher performance levels among the expert group emerged. Based on previous results using the same cohort and an experimental procedure with longer trials, we assume that this pattern would become more pronounced with longer task execution (Götz et al., 2018; Vieluf et al., 2018). Likewise, we did not observe differences in DMD mean mode magnitudes. This result may at first seem to be inconsistent with previous results (Götz et al., 2018; Vieluf et al., 2018). However, the task context studied here is rather characterized by short trials with breaks in between. This could have had a different requirement profile than longer trials, so that the participants likely fatigued less quickly. Furthermore, it must be considered that the participants performed numerous iterations of the steady and sine force tracking tasks in previous experimental sessions within the framework of the Bremen-Hand-Study@Jacobs. However, we may exclude the possibility that faster adaptation processes may have taken place in the experts. Furthermore, as in our previous paper, we did not consider task-related values or electrode preselection in the expert-novice comparison to ensure automatic pattern recognition by the classifiers.

3.2. Classification of tasks in experts and novices

Following our previous study (Goetz et al., 2021) in which we found differences in decoding tasks between age groups, we examined the classification of different tasks based on EEG and force tracking at the individual level, i.e., training and testing a model for each participant. Classification performance at the EEG level was comparable to previous results based on the same tasks. While we found differences in the classification performance in different age groups in our previous study which might point to dedifferentiated and compensatory brain network activation, we couldn't detect any in the task decoding between novices and experts at the individual level. This could indicate that the groups do not differ in the task representations. However, it could be that expertise-related differences are more related to efficiency of information processing (Binder et al., 2017; Götz et al., 2018; Vieluf et al., 2018).

Besides task classification based on EEG, we added force tracking features here as an additional feature set, since we observed expertise effects mainly at the behavioral level. Again, we did not find any differences in classification performance between the groups for individual models. Descriptively, a worse performance of the machine learning models based on the force tracking features can be observed in comparison to models based on EEG features. However, it should be considered that the classification with force tracking features was based only on the mean deviation and the variability of the deviation between applied and target force. Furthermore, almost perfect classification ($\sim 100\%$) was observed in the classification of steady vs sine tracking, whereas the classification of the hand used for force tracking was not accurate, especially for the steady task. An extension of the force tracking feature set further increased the performance of the classifiers regarding the used hand (see supplementary material Figure S1).

Table 2

Overview of task classification results at individual and group level. Mdn: median, IQR: interquartile range, M: mean, SD: standard deviation, U: test statistic of the Mann-Whitney-U test (in case of not normally distributed data), t: test statistic of the t-test (in case of normally distributed data), p: probability value.

			Experts	Novices	Statistical values
Individual level	Accuracy	EEG	M = 0.658	M = 0.696	t(41) = 0.955, p = 0.345
		Force	SD = 0.156	SD = 0.097	
		EEG	M = 0.617	M = 0.614	t(41) = -0.163, p = 0.872
		Force	SD = 0.074	SD = 0.077	
	F1	EEG	M = 0.652	M = 0.689	t(41) = 0.913, p = 0.367
		Force	SD = 0.159	SD = 0.101	
		EEG	M = 0.590	M = 0.588	t(41) = -0.055, p = 0.956
		Force	SD = 0.091	SD = 0.091	
Precision	Accuracy	EEG	M = 0.675	M = 0.708	t(41) = 0.913, p = 0.415
		Force	SD = 0.158	SD = 0.100	
		EEG	M = 0.615	M = 0.611	t(41) = -0.139, p = 0.890
		Force	SD = 0.086	SD = 0.097	
	Recall	EEG	M = 0.658	M = 0.695	t(41) = 0.936, p = 0.355
		Force	SD = 0.156	SD = 0.097	
		EEG	M = 0.617	M = 0.614	t(41) = -0.138, p = 0.891
		Force	SD = 0.074	SD = 0.077	
Group level	Accuracy	EEG	M = 0.306	M = 0.352	t(18) = -3.597, p = 0.002*
		Force	SD = 0.034	SD = 0.021	
		EEG	Mdn = 0.517	Mdn = 0.503	U = 77.000, p = 0.045*
		Force	IQR = 0.024	IQR = 0.017	
	F1	EEG	M = 0.298	M = 0.348	t(18) = -4.095, p < 0.001*
		Force	SD = 0.031	SD = 0.023	
		EEG	M = 0.490	M = 0.418	t(18) = 4.403, p < 0.001*
		Force	SD = 0.028	SD = 0.044	
Precision	Accuracy	EEG	M = 0.308	M = 0.359	t(18) = -4.198, p < 0.001*
		Force	SD = 0.032	SD = 0.021	
		EEG	Mdn = 0.520	Mdn = 0.455	U = 78.000, p = 0.040*
		Force	IQR = 0.059	IQR = 0.099	
	Recall	EEG	M = 0.306	M = 0.352	t(18) = -3.568, p = 0.002*
		Force	SD = 0.034	SD = 0.021	
		EEG	Mdn = 0.516	Mdn = 0.504	U = 75.000, p = 0.064
		Force	IQR = 0.024	IQR = 0.016	

However, such maximization and benchmarking at this level was not the final goal of this study and the use of this enlarged feature set did not increase classification performance (see [supplementary material](#) Table S3).

To compare the generalizability of task classification between experts and novices, we also trained a classification model for each group at group level and compared their performance on participants not used during model training. Here, we found a drop in classification performance presumably caused by individuality due to differing physical and mental conditions across participants, as it is described in the literature for subject-transfer decoding ([Morioka et al., 2015](#)). While the reduction in classification performance was higher for experts for the model based on EEG data, it was the opposite for the model based on force tracking data. This could indicate higher individuality in EEG data of experts while the force tracking performance was more similar between experts. This was also reflected in the statistical comparison of force tracking performance.

In summary, the high accuracy in the classification of tasks at the individual level supports our decision on feature selection. We found no differences between experts and novices in classification performance at individual level as we found for different age groups ([Goelz et al., 2021](#)), which could indicate that the specificity of the task representation on central level is not different between experts and novices. Generalizability is different between groups and might be related to different levels of variability between participants. Subsequent studies in other contexts would be valuable, especially in areas where the effect of machine learning models is crucial, such as neurofeedback or brain-computer-interface (BCI) systems.

3.3. Classification of group membership

Classification of group membership was not above chance level. This was true for both feature levels and level combination (EEG and force tracking features). The best classification performance was achieved using the force tracking features of the steady left task, which is in line with the results of the group comparison at behavioral level. We additionally applied various other classifiers including random forest, linear discriminant analysis and automatic machine learning techniques ([Feurer et al., 2019](#); see [supplementary material](#) Table S1 for an overview of tested models) and used different feature sets (see [supplementary material](#) Table S2 for an overview of tested models). Taken together, the results of all models were comparable. The most promising results were achieved with random forest and combined EEG and force tracking features using the force tracking of the steady left task (see [supplementary material](#) Table S3).

These results differ from the studies presented in the introduction, in which it was possible to successfully classify experts and novices ([Hosp et al., 2021](#); [Shourie, 2016](#); [Watson, 2014](#); [Winkler-Schwartz et al., 2019](#)). It is striking that all these studies used tasks in the exact expertise context for classification. The expertise level of surgeons during simulated or virtual operations could be classified with a high accuracy of 83% and 84% ([Watson, 2014](#); [Winkler-Schwartz et al., 2019](#)). Artistic expertise was successfully classified with an accuracy of 97.5% based on spectral features of EEG during observation and mental imagery of paintings ([Shourie, 2016](#)). Classification of different expertise levels of goalkeepers in soccer observing game scenes was possible as well with high accuracy ([Hosp et al., 2021](#)). Compared to these studies, the professional context of the fine motor experts (including opticians, watchmakers etc.) analyzed in this study was rather broad and the chosen force tracking tasks were only a rough estimate of the expertise context. Although this experimental set-up had the advantage of controlled laboratory conditions, it might have been at the expense of an exact expertise context. Consequently, the expertise context could be a crucial component in classifying experts and novices. This could also be reflected in the fact that we achieved the best classification performance for the steady left tracking task. Holding objects statically with the left hand could be a fundamental part of the work context of the experts studied here. In addition, a larger data pool could help to build more accurate ML models (e.g., using deep learning) that are able to capture even subtle expertise-related differences in indistinct representations of

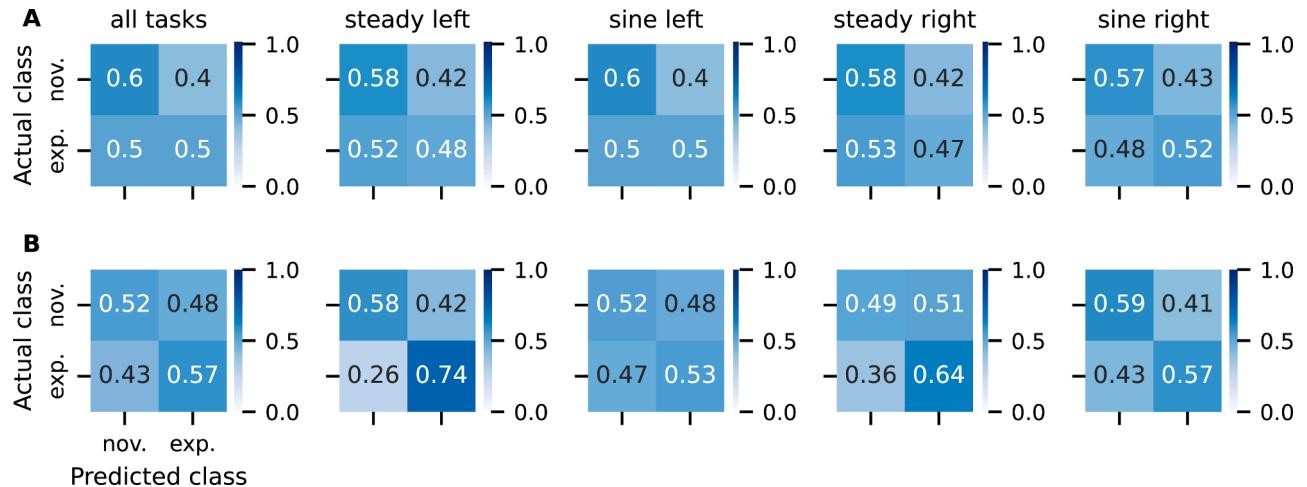


Fig. 3. Mean confusion matrices of group membership classification over all folds based on EEG features (A) and force tracking performance features (B). Exp.: experts, nov.: novices.

Table 3

Overview of group membership classification results. Mean (M) and standard deviation (SD) over all folds.

	Task	Accuracy	F1	Precision	Recall
EEG	All	M = 0.531 SD = 0.068	M = 0.517 SD = 0.077	M = 0.552 SD = 0.091	M = 0.550 SD = 0.089
	Steady left	M = 0.556 SD = 0.067	M = 0.530 SD = 0.058	M = 0.561 SD = 0.061	M = 0.563 SD = 0.062
	Sine left	M = 0.558 SD = 0.078	M = 0.539 SD = 0.094	M = 0.572 SD = 0.096	M = 0.570 SD = 0.106
	Steady right	M = 0.461 SD = 0.073	M = 0.446 SD = 0.084	M = 0.495 SD = 0.111	M = 0.482 SD = 0.104
	Sine right	M = 0.545 SD = 0.094	M = 0.535 SD = 0.099	M = 0.569 SD = 0.092	M = 0.569 SD = 0.099
Force	All	M = 0.434 SD = 0.161	M = 0.421 SD = 0.165	M = 0.509 SD = 0.122	M = 0.535 SD = 0.051
	Steady left	M = 0.618 SD = 0.114	M = 0.604 SD = 0.130	M = 0.673 SD = 0.097	M = 0.681 SD = 0.088
	Sine left	M = 0.489 SD = 0.117	M = 0.482 SD = 0.117	M = 0.551 SD = 0.092	M = 0.545 SD = 0.082
	Steady right	M = 0.473 SD = 0.180	M = 0.454 SD = 0.183	M = 0.549 SD = 0.200	M = 0.575 SD = 0.078
	Sine right	M = 0.522 SD = 0.145	M = 0.517 SD = 0.147	M = 0.590 SD = 0.100	M = 0.594 SD = 0.097

the expertise context.

Future studies could further exploit research questions targeting the context dependency of classification performance. Furthermore, the high degree of individuality described above plays a major role, influencing classification performance also in the classification of group membership. A deeper understanding would further advance the use and development of technologies based on machine learning. Such technologies could be applied in areas where it is of interest to predict the performance state of individuals. These include, for example, rehabilitative or professional contexts. Other possible fields of application could be talent acquisition or adaptive systems.

3.4. Feature space characteristics

Given the low generalizability of the machine learning models in classifying group membership in combination with good decoding performance of force tracking tasks at the individual level, we further investigated the feature space of the EEG features (see Fig. 4A and 4C). Studying the UMAP visualization, a clear assignment of trial-specific EEG features to individual subjects is possible. For both groups, there is a clustered structure with clearly assignable clusters. This structure

suggests a large individuality of the EEG markers, which could partly explain the poor classification performance of the tasks and group membership at the group level. However, we found good classification performance of the tasks within participants. Considered together, this could suggest an individual structure and representation of the tasks in each participant that is not transferable to other participants. This phenomenon is also known from the field of BCI research. Here it has been observed that the models used are poorly transferable to other participants or sessions, thus the development of generalizable models is a current field of research (Morioka et al., 2015; Xu et al., 2020).

Comparing the structure of EEG features between the groups of experts and novices the clustered structure of UMAP embedded EEG features is more prominent in the group of experts. Indeed, the distances of the cluster centroids in the UMAP embedding of the feature representation are higher in experts compared to novices (see Fig. 4B and 4D). However, we do not observe this individuality in all experts in the same way, so it would be conceivable that different expertise levels are represented in our dataset. To the best of our knowledge, individuality of brain activity states in the context of expertise research has not been explored yet. However, a high degree of individuality can be found in the movement patterns of musicians. For example, flute players and pianists showed individuality in their movements and playing, which is skill-dependent (Albrecht et al., 2014; Caramiaux et al., 2018). Likewise, high individuality of muscle synergies was found in trained powerlifters, which was related to specific neural strategies in trained athletes (Kristiansen et al., 2015). Similarly, our data suggests a specialization of brain activation during fine motor tasks. In addition, we also observed a lower dispersion of UMAP embedded EEG patterns per trial in experts (see Fig. 4E). This was reflected in a smaller cluster size per subject for experts.

This high individuality, especially in experts, could have applications for the development of generalizable algorithms in the context of BCI. In this context, the training state of users could have an influence on the transfer of BCI systems to different users.

3.5. Methodological considerations

First and foremost, we had a limited amount of data. However, the size of our dataset was not smaller than that of comparable articles mentioned previously. These studies used data of 20–50 participants and significantly fewer samples due to fewer task repetitions to classify experts and novices with conventional machine learning methods comparable to those we used (Hosp et al., 2021; Winkler-Schwartz et al., 2019). Machine learning models are generally considered to be rather data-hungry and benefit from higher data volumes, which make it

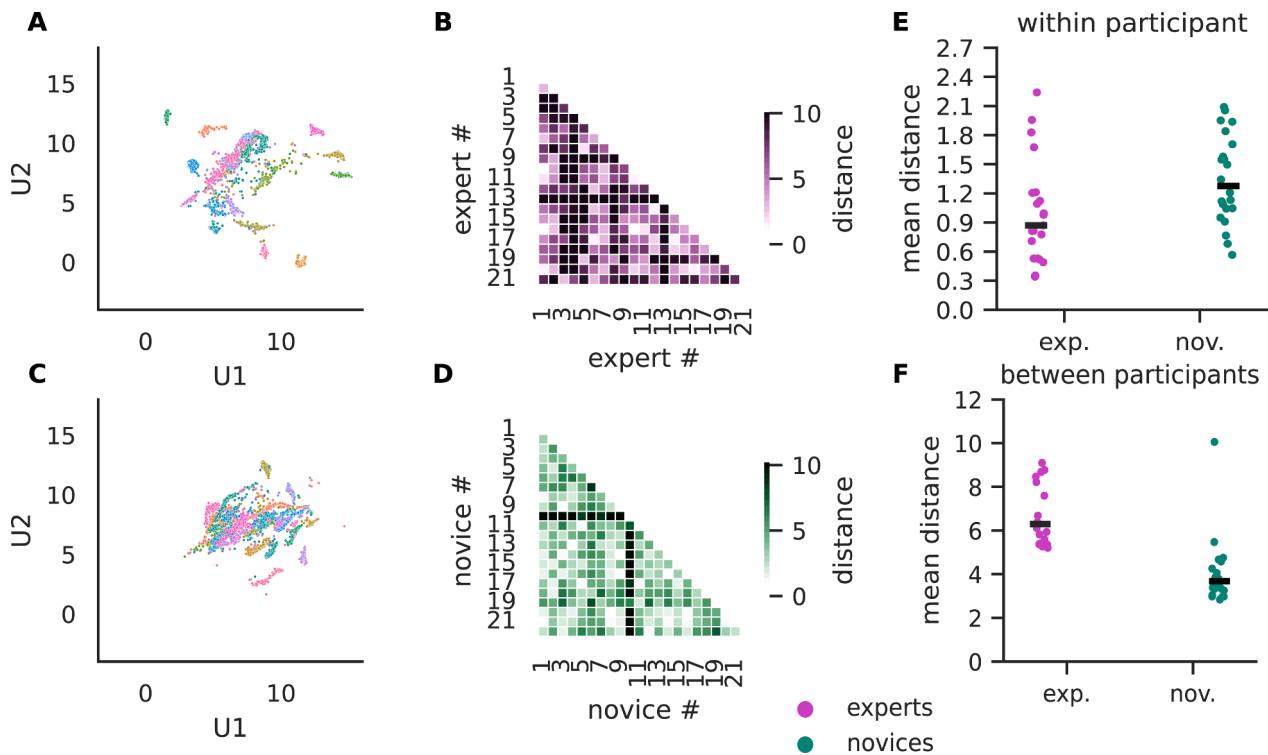


Fig. 4. A & C: UMAP embedding of EEG feature space of fine motor experts (A) and novices (C). Each color corresponds to one participant, each dot corresponds to one trial. B & D: Distance matrices of centroids of each participant cluster in UMAP embedding of experts (B) and novices (D). E: Mean distance of trials within each participant. F: Mean distance of between centroids of each participant cluster. Exp.: experts, nov.: novices. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

possible to detect even subtle differences. However, to compensate for this in terms of a data augmentation approach, we classified at the trial level. In addition, we chose conventional machine learning models that can be applied to smaller datasets. We did not use other methods like deep learning because of their suitability for much larger amounts of data. In the process of group membership classification, feature reductions were not performed to obtain a result based on the entire set of variables and ensure automatic pattern recognition by the classifiers. However, such an approach involves the risk of overfitting, since it is possible that the model is not transferable to new data as it might include redundant features (Hawkins, 2004). We therefore decided to use support vector machines (SVMs) which are effective in high dimensional space.

Using alternative machine learning approaches might result in different performance metrics. For comparison, we used various other classifiers and feature sets (see [supplementary material](#) Table S1 and Table S2) without being able to achieve significant improvement in classification. All machine learning methods used are dependent on the features selected. We decided not to use end-to-end decoding and to use features that have shown expertise-related differences in previous studies. As such the analysis of electrophysiological data was performed with DMD. This method was applied successfully in our previous work detecting expertise-related differences (Gölz et al., 2018; Vieluf et al., 2018). In addition, DMD has already proven to be a valuable method for feature extraction (Rahul-Vigneswaran et al., 2019) as well as classifying fine motor movements (Shiraishi et al., 2020). On the one hand, we did not have to make assumptions by using DMD, as the decomposition step adheres to mathematical optimization principles. On the other hand, we selected DMD modes that belong to certain frequency bands. However, a large part of the spectrum, i.e., θ -, α -, β_1 and β_2 -band, was represented and thus this assumption should not affect the analysis greatly. Finally, we would like to point out that the high individuality might reflect amplitude effects that were influenced by measurement

conditions. Since the experiments were performed by several people, we exclude that this explains the observed effect completely. Furthermore, we obtain satisfactory results in the classification of the tasks on an individual level, so that we assume that the EEG reflects the task effects sufficiently.

4. Conclusion

In this study, we used machine learning to gain data-driven insights into expressions of expertise. We found that classification of group membership based on tasks that roughly reflect the expertise context is not possible with high accuracy. In contrast to positive reports from the literature, we assume a high importance of the expertise context in classification. Furthermore, the classification of tasks is possible with high accuracy within novices and experts at individual level. By examining a low-dimensional representation of the feature space, we found a more pronounced individuality of EEG patterns in experts, suggesting more specialized neural mechanisms in fine motor experts during task performance. In addition to providing data-driven insights, these results could be relevant to the application of machine learning in the context of expertise classification as well as the generalizability of such algorithms in the context of BCI research.

5. Methods

5.1. Participants

The sample was recruited in the context of the Bremen-Hand-Study@Jacobs (Voelcker-Rehage et al., 2013) via flyers, newspaper announcements and telephone calls. Participation was voluntary and each participant provided informed consent. Participants got 8 € per hour compensation. The study was conducted in accordance with the Declarations of Helsinki. The Ethics Committee of the German Society of

Psychology approved the study.

All participants were healthy, without neurological abnormalities and had normal or corrected to normal vision and hearing. In line with their self-report, the Edinburgh Handedness Inventory (Oldfield, 1971) identified all participants as right-handed. Exclusion criteria included hobbies that require the skilled use of hands, such as playing a musical instrument or sewing. Participants were asked about their demographic and educational background, weekly working hours, daily hand use and health related conditions. Experts should have at least 10 years professional experience in a field that requires the skillful and dexterous use of hands such as precision mechanics, e.g., opticians or watchmakers, (Ericsson et al., 2006). Novices were defined as participants whose daily work does not involve extensive use of their hands, such as office or service workers. The sample consisted of 21 experts and 22 novices. Both groups were comparable in age (experts: Mdn = 54.00, IQR: 16.25; novices: Mdn = 55.50, IQR = 18.00) and gender (experts: 10 female / 11 male, novices: 13 female / 9 male). Left hand maximum voluntary contraction (MVC) did not differ between the two groups. Experts had higher MVC in the precision grip with the right hand and a higher frequency of daily hand use (see Table 4 for group values and statistical values).

5.2. Experimental procedure

The experimental setup has been described previously (Goelz et al., 2021). The experiment comprised various force tracking tasks. The task was to track a given target force as precisely as possible. The target force matched either a constant (steady) or a sinusoidal (sine) force level, which was presented as a green line on a screen (19'', 60 Hz frame rate) 80 cm away from the participants (see Fig. 1A and 1B). The steady line was fixed at 2 N and the sine wave ranged from 2 N to 12 N with a frequency of 1 Hz. The time axis (x-axis) corresponded to 5 s and the force axis (y-axis) to a scale from 0 N to 14 N. Participants manipulated a force curve displayed on the screen by pressing a force transducer with thumb and index finger in precision grip with left or right hand. The inactive hand rested either on the armrest or in the participant's lap. The force transducers (Mini-40 Model, ATI Industrial Automation, Garner, NC, USA) were attached to the armrests of a chair on which the participants were sitting. The experiment took place in four blocks of 40 trials with 5 s each and 5–7 s break in between. The first two blocks involved tracking the constant target force with the right hand (steady right) followed by the sinusoidal curve (sine right) with the same hand. The participants then performed the following two blocks with the left hand in the same order (steady left, sine left).

The applied force trajectory was recorded at a sampling rate of 120 Hz with 0.06 N resolution. EEG was recorded with a 32 active channel system (ActiveTwo, BioSemi, Amsterdam, Netherlands). Electrodes were placed according to the international 10–20 system (Jasper, 1958). Two additional electrodes, active Common Mode Sense and Driven

Right Leg were used as reference and ground electrodes and fixed next to Cz. Furthermore, six additional electrodes recorded vertical and horizontal eye movements as well as mastoid potential. The sampling rate of the signal was set to 2048 Hz and an online band-pass filter between 0.16 Hz and 100 Hz was used.

5.3. Data analysis

For the data analysis we used *Python* version 3.7.6 (Python Software Foundation, Wilmington, DE, USA) with *MNE Python* version 0.20.7 (Gramfort et al., 2013) to analyze and visualize the EEG and *umap-learn* version 0.5.1 (McInnes et al., 2018) to examine the EEG feature space. The machine learning pipelines were implemented with *scikit-learn* version 0.24.2 (Pedregosa et al., 2011). For each participant we classified which task, i.e., steady right, sine right, steady left or sine left, was performed. In addition, we classified experts vs. novices on a group level. For all classifiers we tuned the hyperparameters using grid search, i.e., we searched for the best hyperparameters among a predefined grid of values. An overview of the hyperparameter grid used can be found in Table S1 in the supplementary material. Consequently, hyperparameters were chosen that resulted in the highest classification accuracy.

5.3.1. Preprocessing

5.3.1.1. Force tracking data. First, we filtered the data of the force trajectory with a fourth order Butterworth filter with a cut-off at 30 Hz (Götz et al., 2018; Vieluf et al., 2012; Vieluf et al., 2015; Vieluf et al., 2018). Next, for each trial we calculated the mean deviation representing the mean absolute difference between target and applied force for each participant. Furthermore, we extracted the absolute variability represented by the standard deviation of the difference between target and applied force. To automatically detect outliers in the force data, we used the statistical Z-values of the mean accuracy of the individual results. Trials with $Z > 3$ were excluded from further analyses as we assumed these to be incorrectly executed.

5.3.1.2. EEG data. EEG data was resampled to 200 Hz and re-referenced to the linked mastoids. The default zero-phase FIR filter in MNE-Python was used to filter the signals between 4 and 30 Hz. We corrected ocular artifacts using independent component analysis (ICA) with the FastICA Algorithm (Hyvärinen, 1999). The resulting independent components whose source signal correlated with the signal of the electrooculography channel were automatically marked and removed. This resulted in the rejection of one component in 17 experts / 18 novices and two components in 4 experts / 4 novices. Next, we created data segments starting at the one second and ending at the four seconds after trial onset mark. Trial segments that contained artifacts were identified and removed automatically using the Autoreject pipeline (Jas et al., 2017). Moreover, we excluded all trial segments marked as

Table 4

Group characteristics - MVC: Maximum voluntary contraction, RH: right hand, LH: left hand, Mdn: median, IQR: interquartile range, M: mean, SD: standard deviation, U: test statistic of the Mann-Whitney-U test (in case of not normally distributed data), t: test statistic of the t-test (in case of normally distributed data), p: probability value.

		Experts (n = 21, female: 10)	Novices (n = 22, female: 13)	Statistical values
Age [years]		Mdn = 54.000 IQR = 16.250	Mdn = 55.500 IQR = 18.000	U = 204.000, p = 0.635
MVC [N]	LH	Mdn = 46.455 IQR = 24.519	Mdn = 42.627 IQR = 27.757	U = 204.000, p = 0.520
	RH	Mdn = 56.000 IQR = 28.386	Mdn = 48.614 IQR = 13.243	U = 147.000, p = 0.043*
Daily use of hands ¹		M = 37.381 SD = 5.536	M = 18.381 SD = 5.670	t(41) = 10.988, p < 0.001*
# of trials after preprocessing		Mdn = 156 IQR = 1	Mdn = 156 IQR = 2	U = 274.000 p = 0.296

1: Maximum value = 45; A higher value corresponds to a higher frequency of hand use.

incorrectly executed. Each three second segment was decomposed using DMD (Brunton et al., 2016). For this step we analysed windows of 0.5 s with 0.25 s overlap to account for data nonstationarities and each segment was embedded by stacking and shifting the data four times, as described in previous work (Goelz et al., 2021). The optimum of these parameters (stacking depth, window size, and overlap) were determined by an error analysis of data from five randomly chosen subjects.

5.3.2. Classification of tasks

To show that we are using task-relevant features for group membership classification and to detect any group differences, we classified the tasks based on EEG as well as behavioral measurements, i.e., force tracking, EEG feature extraction as well as classification followed strictly the procedure described in previous work (Goelz et al., 2021). The absolute DMD mode magnitudes in each frequency band and trial were projected onto the common spatial pattern (CSP) space and the logarithmic variance of the projected DMD mode magnitudes of all time windows per trial were calculated as features. To be consistent with our previous work, we used linear discriminant analysis (LDA) for the classification based on EEG. The force tracking features included the mean deviation and variability per trial. For classification based on force tracking we used support vector machines (SVMs).

For each participant, we built and evaluated a machine learning model using 10-fold nested cross-validation. In each fold the data was divided into 80% (i.e., maximum 128 trials) for training and 20% (i.e., maximum 32 trials) for testing. Due to exclusion of trials because of incorrect execution or artifacts in corresponding EEG segments the exact number of trials varied between participants but was comparable between groups (see Table 4 and supplementary material Table S4).

To test the generalizability of the models between participants, we also built and evaluated a model at the group level for experts and novices. For this we used 10-fold nested cross-validation. In each fold we used all trials of 80% of the subjects (experts: 17 / novices: 18) per group for training and the 20% of the subjects (experts: 4 / novices: 4) for testing.

Hyperparameters of SVM or LDA classifiers were tuned in each fold using ten-fold cross-validation on the training data using a grid search procedure (see supplementary material Table S1 for the hyperparameter grid used).

5.3.3. Classification of group membership

To increase the data size for the classification of group membership, we performed the classification on all force tracking trials rather than the mean values of an entire force tracking block. All force tracking tasks were used collectively as well as each force tracking task individually. For that, we used either only EEG features, only force tracking features or a combination thereof in separate iterations. The EEG features consisted of the absolute DMD mean modes per frequency band and trial assuming that expertise relevant information is detectable in spatially and temporally coherent EEG patterns, as we have shown in previous studies (Götz et al., 2018; Vieluf et al., 2018). We averaged DMD modes corresponding to the frequency bands θ - (4 to $<$ 8 Hz), α - (8 to $<$ 12 Hz), β_1 - (12 to $<$ 16 Hz) and β_2 - (16 to $<$ 30 Hz) per trial for each participant to obtain the DMD mean mode for each frequency band and trial. As force tracking features the mean deviation and variability per trial for each participant as described above were used.

To classify group membership based on EEG data and force tracking performance, we used SVMs. The model was trained and tested using 10-fold nested cross-validation. Therefore, we divided the dataset ten times into a training and test set using all trials of 80% of participants (i.e., 35 participants) for training and all trials of 20% of participants (i.e., 8 participants) to test the generalization of the classifier to unseen participants (see supplementary material Table S3). In each data partition hyperparameters of the SVM classifier were tuned dividing the training dataset again 10 times with the same procedure (see supplementary material Table S1 for the hyperparameter grid used).

5.3.4. Extraction of feature space characteristics

To visually examine the EEG feature space, we used uniform manifold approximation and projection (UMAP) (McInnes et al., 2018) and reduced the dimensionality to two dimensions. To quantify the clustering structure in the UMAP embedding, we calculated the Euclidean distance of trials within each participant. In addition, the centroids of all trials per participant were calculated and the Euclidean distance between these centroids was determined.

5.3.5. Statistical analysis

For statistical analyses, we used *scipy* version 1.6.2 (Virtanen et al., 2020), *statsmodels* version 0.12.2 (Seabold and Perktold, 2010) and *permute* version 0.2 (Millman, 2015). To compare the groups in basic characteristics as well as in their force tracking performance, we conducted *t*-tests and Mann-Whitney-*U* tests, in case of violation of normality assumption. Force tracking performance of each participant was quantified as mean over all trials in measures described above (mean deviation and variability). Intertrial variability was calculated as standard deviation over all trials of the mean tracking accuracy.

To test for group differences in EEG activity, we averaged the DMD mean mode magnitudes in each frequency band over all trials per force tracking task and participant. Next, we conducted permutation *t*-tests to compare DMD mean mode magnitudes between the groups. This statistic was chosen as a non-parametric alternative suitable for high dimensional data like EEG data (Maris and Oostenveld, 2007). For this we determined all possible values of the test statistics under the null hypothesis with a 10,000-fold random reordering of the data.

To evaluate the machine learning models, we calculated the proportion of correct predictions of the model (accuracy) as well as (macro average) F1, precision and recall values and report the mean over all folds. We conducted *t*-tests and Mann-Whitney-*U* tests, in case of violation of normality assumption, to compare performance of task classification between the groups. For all tests the alpha level was set to 0.05 and false discovery rate (Benjamini and Hochberg, 1995) was used to account for type I errors.

Funding

The research was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation, VO 1432/7-1 -SPP 1184 and DFG Project-ID 416,228,727 - SFB 1410). The study was supported within the framework of the equal opportunities concept 2 of the Paderborn University and by the Heinz Nixdorf Westfalian Foundation.

Code availability

Python source code that supports the results is available from: https://github.com/christiangoelz/FME_classification.

CRediT authorship contribution statement

R. Gaidai: Software, Formal analysis, Writing – original draft. **C. Goelz:** Software, Formal analysis, Writing – original draft. **K. Mora:** Software, Formal analysis, Writing – review & editing. **J. Rudisch:** Writing – review & editing. **E. Reuter:** Conceptualization, Investigation, Writing – review & editing. **B. Godde:** Conceptualization, Supervision, Writing – review & editing. **C. Reinsberger:** Writing – review & editing. **C. Voelcker-Rehage:** Conceptualization, Supervision, Writing – review & editing. **S. Vieluf:** Conceptualization, Investigation, Supervision, Project administration, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence

the work reported in this paper.

Acknowledgements

We thank Patrick Jersch, Janine Ohmann and other student assistants for their support during data collection.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.brainres.2022.148001>.

References

- Albrecht, S., Janssen, D., Quarz, E., Newell, K.M., Schöllhorn, W.I., 2014. Individuality of movements in music-finger and body movements during playing of the flute. *Hum. Mov. Sci.* 35, 131–144. <https://doi.org/10.1016/j.humov.2014.03.010>.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Series B (Methodol.)* 57 (1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>.
- Binder, J.C., Bezzola, L., Haueter, A.I.S., Klein, C., Kühnis, J., Baetschmann, H., Jäncke, L., 2017. Expertise-related functional brain network efficiency in healthy older adults. *BMC Neurosci.* 18 (1), 1–15. <https://doi.org/10.1186/s12868-016-0324-1>.
- Brunton, B.W., Beyeler, M., 2019. Data-driven models in human neuroscience and neuroengineering. *Curr. Opin. Neurobiol.* 58, 21–29. <https://doi.org/10.1016/j.conb.2019.06.008>.
- Brunton, B.W., Johnson, L.A., Ojemann, J.G., Kutz, J.N., 2016. Extracting spatial-temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition. *J. Neurosci. Methods* 258, 1–15. <https://doi.org/10.1016/j.jneumeth.2015.10.010>.
- Bzdok, D., Yeo, B.T.T., 2017. Inference in the age of big data: Future perspectives on neuroscience. *Neuroimage*. 155, 549–564. <https://doi.org/10.1016/j.neuroimage.2017.04.061>.
- Caramiaua, B., Bevilacqua, F., Wanderley, M.M., Palmer, C., Grahn, J.A., 2018. Dissociable effects of practice variability on learning motor and timing skills. *PLoS ONE* 13 (3), e0193580. <https://doi.org/10.1371/journal.pone.0193580>.
- Ericsson, K.A., Charness, N., Feltovich, P.J., Hoffman, R.R., 2006. *The Cambridge Handbook of Expertise and Expert Performance*, second ed. Cambridge University Press, Cambridge.
- Feurer, M., Klein, A., Eggensperger, K., Springenberg, J.T., Blum, M., Hutter, F., 2019. Auto-sklearn: Efficient and Robust Automated Machine Learning. In: Hutter, F., Kotthoff, L., Vanschoren, J. (Eds.), *Automated Machine Learning: Methods, Systems, Challenges*. Springer International Publishing, Cham, pp. 113–134.
- Göbel, C., Mora, K., Rudisch, J., Gaidai, R., Reuter, E., Godde, B., Reinsberger, C., Voelcker-Rehage, C., Vieluf, S., 2021. Classification of visuomotor tasks based on electroencephalographic data depends on age-related differences in brain activity patterns. *Neural Netw.* 142, 363–374. <https://doi.org/10.1016/j.neunet.2021.04.029>.
- Göbel, C., Voelcker-Rehage, C., Mora, K., Reuter, E.M., Godde, B., Dellnitz, M., Reinsberger, C., Vieluf, S., 2018. Improved neural control of movements manifests in expertise-related differences in force output and brain network dynamics. *Front. Physiol.* 9, 1540–1540. doi: 10.3389/fphys.2018.01540.
- Gramfort, A., Luessi, M., Larson, E., Engemann, D., Strohmeier, D., Brodbeck, C., Gaj, R., Jas, M., Brooks, T., Parkkonen, L., Hämäläinen, M., 2013. MEG and EEG data analysis with MNE-Python. *Front. Neurosci.* 7, 267–267. doi: 10.3389/fnins.2013.00267.
- Hawkins, D.M., 2004. The Problem of Overfitting. *J. Chem. Inf. Comput. Sci.* 44 (1), 1–12. <https://doi.org/10.1021/ci0342472>.
- Hosp, B.W., Schultz, F., Höner, O., Kasneci, E., Wood, G., 2021. Soccer goalkeeper expertise identification based on eye movements. *PLoS ONE* 16 (5), e0251070. <https://doi.org/10.1371/journal.pone.0251070>.
- Hyvärinen, A., 1999. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans. Neural Netw.* 10 (3), 626–634. <https://doi.org/10.1109/72.761722>.
- Jas, M., Engemann, D.A., Bekhti, Y., Raimondo, F., Gramfort, A., 2017. Autoreject: Automated artifact rejection for MEG and EEG data. *Neuroimage*. 159, 417–429. <https://doi.org/10.1016/j.neuroimage.2017.06.030>.
- Jasper, H., 1958. The ten twenty electrode system of the international federation. *Electroencephalogr. Clin. Neurophysiol.* 10, 370–375.
- Komar, J., Seifert, L., Thouwarec, R., Benguigui, N., 2015. What Variability tells us about motor expertise: measurements and perspectives from a complex system approach. *Mov. Sport Sci./Sci. Mot.* (89), 65–77. <https://doi.org/10.3917/sm.089.0065>.
- Krampe, R.T., Ericsson, K.A., 1996. Maintaining Excellence: Deliberate Practice and Elite Performance in Young and Older Pianists. *J. Exp. Psychol. Gener.* 125 (4) <https://doi.org/10.1037/0096-3445.125.4.331>, 331–331.
- Krampe, R.T., 2002. Aging, expertise and fine motor movement. *Neurosci. Biobehav. Rev.* 26 (7), 1–8. [https://doi.org/10.1016/s0149-7634\(02\)00064-7](https://doi.org/10.1016/s0149-7634(02)00064-7).
- Kristiansen, M., Madeleine, P., Hansen, E.A., Samani, A., 2015. Inter-subject variability of muscle synergies during bench press in power lifters and untrained individuals. *Scand. J. Med. Sci. Sports* 25 (1), 89–97. <https://doi.org/10.1111/sms.12167>.
- Law, B., Lomax, A.J., Atkins, M.S., Mackenzie, C.L., Kirkpatrick, A.E., 2004. Eye gaze patterns differentiate novice and experts in a virtual laparoscopic surgery training environment. In: Proceedings of the 2004 symposium on Eye tracking research & applications (ETRA '04). Association for Computing Machinery, New York, NY, USA, pp. 41–48. <https://doi.org/10.1145/968363.968370>.
- Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* 164 (1), 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>.
- McInnes, L., Healy, J., Saul, N., Grossberger, L., 2018. UMAP: Uniform Manifold Approximation and Projection. *J. Open Sour. Softw.* 3, 861. <https://doi.org/10.21105/joss.00861>.
- Millman, K.J., 2015. *Permute—a Python package for permutation tests and confidence sets*. University of California, Berkeley.
- Morioka, H., Kanemura, A., Hirayama, J.-I., Shikauchi, M., Ogawa, T., Ikeda, S., Kawanabe, M., Ishii, S., 2015. Learning a common dictionary for subject-transfer decoding with resting calibration. *NeuroImage*. 111, 167–178. <https://doi.org/10.1016/j.neuroimage.2015.02.015>.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*. 9 (1), 97–113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4).
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., 2011. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830. [https://doi.org/10.1016/0028-3932\(71\)90067-410.5555/1953048.2078195](https://doi.org/10.1016/0028-3932(71)90067-410.5555/1953048.2078195).
- Rahul-Vigneswaran, K., Sachin-Kumar, S., Mohan, N., Soman, K.P., 2019. Dynamic Mode Decomposition based feature for Image Classification. In: TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON). pp. 745–750.
- Seabold, S., Perktold, J., 2010. Statsmodels: Econometric and statistical modeling with python. In: Proceedings of the 9th Python in Science Conference. Austin, TX, pp. 61. doi: 10.25080/Majora-92bf1922-011.
- Shiraishi, Y., Kawahara, Y., Yamashita, O., Fukuma, R., Yamamoto, S., Saitoh, Y., Kishima, H., Yanagisawa, T., 2020. Neural decoding of electrocorticographic signals using dynamic mode decomposition. *J. Neural Eng.* 17 (3), 036009. <https://doi.org/10.1088/1741-2552/ab8910>.
- Shourie, N., 2016. Cepstral analysis of EEG during visual perception and mental imagery reveals the influence of artistic expertise. *J. Med. Sign. Sens.* 6 (4) <https://doi.org/10.4103/2228-7477.195088>, 203–203.
- Vieluf, S., Mahmoodi, J., Godde, B., Reuter, E.M., Voelcker-Rehage, C., 2012. The influence of age and work-related expertise on fine motor control. *GeroPsych*. 25 (4), 199–206. <https://doi.org/10.1024/1662-9647/a000071>.
- Vieluf, S., Godde, B., Reuter, E.M., Voelcker-Rehage, C., 2013. Effects of age and fine motor expertise on the bilateral deficit in force initiation. *Exp. Brain Res.* 231 (1), 107–116. <https://doi.org/10.1007/s00221-013-3673-3>.
- Vieluf, S., Temprado, J.J., Berton, E., Jirsa, V.K., Sleimen-Malkoun, R., 2015. Effects of task and age on the magnitude and structure of force fluctuations: Insights into underlying neuro-behavioral processes. *BMC Neurosci.* 16 (1), 1–17. <https://doi.org/10.1186/s12868-015-0153-7>.
- Vieluf, S., Mora, K., Götz, C., Reuter, E.M., Godde, B., Dellnitz, M., Reinsberger, C., Voelcker-Rehage, C., 2018. Age- and Expertise-Related Differences of Sensorimotor Network Dynamics during Force Control. *Neuroscience*. 388, 203–213. <https://doi.org/10.1016/j.neuroscience.2018.07.025>.
- Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, I., Feng, Y.u., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., Vijaykumar, A., Bardelli, A.P., Rothberg, A., Hilboll, A., Kloekner, A., Scopatz, A., Lee, A., Rokem, A., Woods, C.N., Fulton, C., Masson, C., Häggström, C., Fitzgerald, C., Nicholson, D.A., Hagen, D.R., Pasechnik, D.V., Olivetti, E., Martin, E., Wieser, E., Silva, F., Lenders, F., Wilhelm, F., Young, G., Price, G.A., Ingold, G.-L., Allen, G.E., Lee, G.R., Audren, H., Probst, I., Dietrich, J.P., Silterra, J., Webber, J.T., Slavić, J., Nothman, J., Buchner, J., Kulick, J., Schönberger, J.L., de Miranda Cardoso, J.V., Reimer, J., Harrington, J., Rodríguez, J.L.C., Nunez-Iglesias, J., Kuczynski, J., Tritz, K., Thoma, M., Newville, M., Kümmeler, M., Bolingbroke, M., Tartre, M., Pak, M., Smith, N.J., Nowaczyk, N., Shebanov, N., Pavlyk, O., Brodtkorb, P.A., Lee, P., McGibbon, R.T., Feldbauer, R., Lewis, S., Tygier, S., Sievert, S., Vigna, S., Peterson, S., More, S., Pudlik, T., Oshima, T., Pingel, T.J., Robitaille, T.P., Spura, T., Jones, T.R., Cera, T., Leslie, T., Zito, T., Krauss, T., Upadhyay, U., Halchenko, Y.O., Vázquez-Baeza, Y., 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods*. 17 (3), 261–272.
- Voelcker-Rehage, C., Reuter, E.M., Vieluf, S., Godde, B., 2013. Influence of age and expertise on manual dexterity in the work context: The Bremen-hand-study@Jacobs. In: *Age-Differentiated Work Systems*. Springer, Berlin, Heidelberg, pp. 391–415.
- Watson, R.A., 2014. Use of a machine learning algorithm to classify expertise: Analysis of hand motion patterns during a simulated surgical task. *Acad. Med.* 89 (8), 1163–1167. <https://doi.org/10.1097/ACM.00000000000000316>.
- Winkler-Schwartz, A., Yilmaz, R., Mirchi, N., Bissonnette, V., Ledwos, N., Siyar, S., Azarnoush, H., Karlik, B., Del Maestro, R., 2019. Machine Learning Identification of Surgical and Operative Factors Associated With Surgical Expertise in Virtual Reality Simulation. *JAMA Netw. Open*. 2 (8), e198363. <https://doi.org/10.1001/jamanetworkopen.2019.8363>.
- Xu, L., Xu, M., Ke, Y., An, X., Liu, S., Ming, D., 2020. Cross-Dataset Variability Problem in EEG Decoding With Deep Learning. *Front. Hum. Neurosci.* 14 <https://doi.org/10.3389/fnhum.2020.00103>, 103–103.