Physiological synchrony in brain and body as a measure of attentional engagement

Ivo Stuldreher

Physiological synchrony in brain and body as a measure of attentional engagement

Ivo Stuldreher

Physiological synchrony in brain and body as a measure of attentional engagement

Dissertation

to obtain

the degree of doctor at the University of Twente,

on the authority of the rector magnificus,

prof. dr. ir. A. Veldkamp,

on account of the decision of the Doctorate Board,

to be publicly defended

on Friday the 19th of January 2024 at 16:45 hours

by

Ivo Vincent Stuldreher

born on the 26th of October, 1996

in Alphen aan den Rijn, the Netherlands

This dissertation has been approved by:

Promotors: prof. dr. J.B.F. van Erp

prof. dr. A.M. Brouwer

Cover design: Ivo Stuldreher

Printed by: Ipskamp Printing

Lay-out: Ivo Stuldreher

ISBN (print): 978-90-365-5936-2

ISBN (digital): 978-90-365-5937-9

URL: https://doi.org/10.3990/1.9789036559379

Graduation committee:

Chair and secretary:

prof. dr. J.N. Kok

Promotors:

prof. dr. J.B.F. van Erp	University of Twente
prof. dr. A.M. Brouwer	Radboud University

Members:

prof. dr. V. Evers	University of Twente
prof. dr. M.L. Noordzij	University of Twente
prof. dr. J.H.D.M. Westerink	Eindhoven University of Technology
prof. dr. S. Debener	University of Oldenburg
prof. dr. L.C. Parra	City Univeristy of New York

Table of Contents

Acknowledgments

Before you lies my dissertation. It feels strange typing this, as the research presented in this thesis is certainly not the result of my efforts only. How could research on physiological synchrony in groups of individuals be conducted by one individual? The answer: it cannot. There thus are plenty of people without whom this thesis would not have been here. I am grateful of all of you and would like to thank you here.

First and foremost, dear Anne-Marie. Since we first met when I applied for your internship position in 2018 we just got along. I am not sure whether it was due to our level of shared attention, but there was definitely high synchrony. This synchrony still remains today. An engineer and a psychologist working together may perhaps not be a match on paper, but in reality we are. Working with you has brought me a lot, in terms of knowledge about and experience in this field of research, but more important also in terms of having fun in the work I do. I am very grateful for the opportunities that you have given me and are still giving me. Although this dissertation has my name on the cover, the research that is presented inside is just as much yours as it is mine.

Second, dear Jan. When Anne-Marie and I approached you with the request to be my promotor your response was positive without hesitation. I feel lucky to have a promotor like you, who is always enthusiastic and supportive, but at the same time remains critical. Your support and trust cheered me on, while your critical views made my sometimes overly complex papers more accessible. You are a real visionair and thereby also helped me to look beyond the findings of my own research. By doing so you helped shape the story of this thesis and made it to the booklet that lies in front of you. Jan, you are the promotor that every PhD student would want, thank you for that.

Pursuing a PhD at TNO can be challenging at times, as the time and resources needed are not necessarily arranged in full beforehand. This challenge was more than compensated by the joy that was brought by all the colleagues I could work with. I sincerely thank everyone at TNO who helped me in my PhD trajectory, either in helping to arrange this trajectory at TNO or in contributing to the contents of this thesis. I cannot thank all of you here, but know that I appreciate the interaction with all of you. Here I would like to especially thank those indivdiuals that have contributed to the contents of this thesis. Nattapong, thank you for guiding me in my initial explorations of physiological data that in the end greatly contributed to the contents of this thesis. Daisuke, thank you for sharing insights from your PhD journey with me and providing me with the opportunity to add physiological synchrony to the toolbox of implicit measures of food experience. Charelle, pursuing a PhD at TNO is far from a clearly set-out procedure. It was really helpful that we could go down the rabbit hole

of conducting a PhD at TNO together. Good luck with the final steps towards a superb dissertation. I would also like to thank everyone from the NeuroLabNL startimpuls for including me in the group when I started this PhD trajectory.

I would also like to say thanks to all the interns that helped shape this dissertation in any form, by assisting in data collection, data analysis, writing, or in any other form. Thank you Alexandre, Ana, Anne, Emma, Jasper and Priya.

My PhD journey has been relatively smooth. In part, this is due to the people that I thanked before, but in part also due to everyone around me. First, my dear friends, I am so grateful to have all of you in my life. You truly bring joy to life, whether on an exorbitant sailing yacht near the shores of the Croatian coast, on our awesome camping spot at Lowlands or at home drinking a Weihenstephaner and having a chat. You have all cheered me on to get to where I am now and I hope we remain cheering each other on for many years to come.

My dear family, you are my proudest supporters. The last year has been an emotional one, which underlined the importance of each and everyone of you to me. You always cheered me on in life and in my scientific journey. Dad, you must be the most loyal reader of my papers. Mum, you always listen to me with full interest when I enthusiastically tell you about some obscure scientific finding. Coen, you inspire me with the way you use music to bring people in-sync. Noan, what a happy little boy you are, certainly due to your loving parents Kelly and Coen. You bring so much happiness to everyone around you. Oma, I know how proud you are and that fills me with joy.

Last, my dear Fenna. I can only end these acknowledgments by thanking you. You have jokingly mentioned that you did not need to be thanked as I would not have needed your support throughout this PhD. I certainly did need it and would like to express my gratitude and love for you as follows. In this dissertation I describe that the signals that can be measured from brains and bodies of people can show synchronous behavior, which we explain by mechanisms of shared attention. This means that in this thesis synchrony in measures such as heart rate or brain activation is not the result of any social mechanism, such as empathy, social closeness or love, but instead is the result of multiple individuals focusing their attention on the same information in the outside world. Although I stand by these findings, you are the person that makes me wonder if really there is not more. I feel so strongly in-sync with you, my dear Fenna, that certainly there must be more than shared attention that can cause synchrony to occur. Whether it indeed is due to an unknown aspect in our love for each other, or just due to our extremely high level of shared attention, it doesn't matter for me. Let us stay in-sync with each other as long as we may live.

Summary

Attentional engagement – the emotional, cognitive and behavioral connection with information to which the attention is focused – is important in all settings where humans process information. As it is difficult to estimate the attentional engagement of others, measures of attentional engagement could be helpful to, for instance, support teachers in online classrooms, or individuals working together in teams. In this thesis we aim to exploit the similarity in neurophysiological responses across individuals as an implicit and continuous measure of attentional engagement. We refer to this similarity in physiological responses recorded from brain or body across individuals as physiological synchrony. We build upon studies indicating that physiological synchrony may act as an index of attentional engagement. Though physiological synchrony is said to reflect attentional engagement, it is unclear how exactly modulations of attention affect physiological synchrony in brains, here monitored through the electroencephalogram (EEG), and especially in bodies, here monitored through electrodermal activity (EDA) and heart rate. Furthermore, it is unclear under which limitations that come with measuring in real-life conditions, physiological synchrony remains a valid measure of attentional engagement. In this thesis we aim to uncover how different types of attention modulation are captured by physiological synchrony in brains and bodies (part I) and to what extent physiological synchrony may be used as tool to monitor attention in real-life settings (part II).

I: attentional modulations and physiological synchrony in brains and bodies

In part I, we address the research question: "How do different manipulations of attention affect physiological synchrony in brains and bodies?"

In our first study we simultaneously monitored EEG, EDA and heart rate of individuals who all heard the exact same audio track consisting of an audiobook with interspersed auditory events such as affective sounds and beeps. Individuals were instructed to either attend to the audiobook or to attend to and keep track of the interspersed events. We computed inter-subject correlations as a measure of physiological synchrony. Using this dataset, in **Chapter 2** we investigated to what extent inter-subject correlations are higher when computed among individuals with the same instruction regarding the focus of attention compared to individuals with a different instruction regarding the focus of attention. We find that inter-subject correlations in all three measures are higher among individuals with the same focus of attention than among individuals with a different focus of attention. In EEG and heart rate, inter-subject correlations predict performance on a post-stimulus test. Using the same dataset, in **Chapter 3** we investigated how well inter-subject correlations predict the occurrence of attentionally engaging moments in time. By doing so we aimed to study the respective influences of bottom-up (stimulus driven) and top-down attention on the occurrence of physiological synchrony. We find that the occurrence of stimuli interspersed in the audiobook could be detected based on inter-subject correlations and that both bottom-up and top-down attention play a role.

Rather than only by explicit instructions on what to focus attention on, in real-life settings attention also varies across and within individuals implicitly, due to variations in personal trait across individuals and variations in the momentary attentional state within individuals. In **Chapter 4** we investigated whether inter-subject correlations in EEG capture interpersonal differences in attentional processing that originate from differences in personal trait, in this study food neophobia, the fear to buy or try novel foods. Indeed, we find higher inter-subject correlations in EEG among individuals with higher levels of food neophobia when watching a movie about the origin and production of a novel food. In **Chapter 5** we investigated whether momentary variations in attention within individuals could be captured by inter-subject correlations in EDA and heart rate. We computed inter-subject correlations during ten movies that were presented over the course of a night in which the monitored individuals were sleep deprived. Inter-subject correlations in EDA predicted performance on consecutive vigilant attention tasks. This indicates that inter-subject correlations can reflect variations in general attention.

II: physiological synchrony from lab to life

In part II we address the research question: “To what extent may physiological synchrony be used as tool to monitor attention in real-life settings?”

In **Chapter 6** we investigated whether inter-subject correlations as a measure of attentional engagement can also be captured with wearables instead of high-end lab-grade equipment. Inter-subject correlations in EDA and heart rate measured using wearables were found to distinguish between individuals with different selective focus of attention with similar, if not higher, accuracies than in EDA and heart rate measured using high-end lab-grade equipment. For monitoring physiological synchrony we are thus not dependent on expensive, immobile laboratory-grade equipment. This indicates that precise signals with high sample-rate containing information across the entire frequency spectrum are not essential for monitoring of physiological synchrony. In **Chapter 7** we further investigated the preconditions in terms of group size and recording duration for successful capturing of inter-subject correlations in EDA and heart rate recorded with wearable devices. As expected, increases in both group size and recording duration increase the percentage of participants with significant inter-subject correlations. It was found that the total amount of data (i.e., the group size times the recording duration) determines the percentage of participants with significant inter-subject correlations, where it does not matter by what ratio of group size and recording duration this amount of data is achieved.

For application in real-life settings, inter-subject correlations may be combined with novel machine learning techniques. In **Chapter 8** we combined inter-subject correlations in EEG, EDA and heart rate with unsupervised learning algorithms with the aim of clustering individuals with the same selective focus of attention. We used the data of the study presented in Chapter 2 and attempted to cluster the two groups of individuals that each have a different instruction regarding their selective focus of attention, without using labels regarding attentional instruction for any of the individuals. Using EEG, accuracies are as high as 85% and higher than expected based on chance. The clustering approach was not successful when using EDA or heart rate alone. However, combining inter-subject correlations in EEG, EDA and heart rate in a multimodal approach results in a maximum accuracy of 96%. Furthermore, results are less dependent on the specific algorithm used.

In **Chapter 9** we moved towards an actual real-life setting, examining inter-subject correlations in EDA and heart rate data that was previously recorded among students in the classroom in two separate studies. As proof of concept, we show that inter-subject correlations are higher among students in the same compared to different classrooms, but only in one of the two studies. In this study more data were available with more individuals of whom EDA and heart rate were recorded per classroom. Results were also more robust for heart rate than for EDA. Overall, inter-subject correlations can be monitored in the classroom, but only with the evident remark that sufficient amounts of data are available.

In **Chapter 10** the implications of our findings regarding the overall aims of this thesis are discussed. In Part I inter-subject correlations were found to reflect multiple types of attention modulations, affected by both sensory and top-down mechanisms of attention. This thesis adds to previous work by showing that not only synchronous brain, but also synchronous body metrics reflect this attentional processing, where synchronous brains and bodies also reflect complementary aspects of attention. Top-down focus of attention is best reflected by synchrony in EEG measures. Emotionally salient events attracting attention are best reflected by synchrony in body measures, being EDA and heart rate. The findings presented in Part II are promising for physiological synchrony to be used in real-life settings, as physiological synchrony can be successfully captured using wearable devices in real-life settings and can be combined with novel unsupervised learning algorithms. Still, the limitations in terms of sufficient data that are required for robust monitoring and in terms of limited variance in attention explained should be considered. To advance physiological synchrony as a tool to monitor attention in real-life settings, future work should focus on the applied scientific, methodological and ethical questions that remain unanswered. Future work could for instance investigate to what extent shared attentional engagement underlies the occurrence of physiological synchrony in settings with social interaction. It could also focus on the development novel multimodal metrics of physiological synchrony. Last, future work could investigate the ethics of the use of physiological synchrony as tool to monitor attentional engagement in each use-case that is explored.

Samenvatting

Aandachtige betrokkenheid (Engels: attentional engagement, vanaf nu kortweg aandacht) - de emotionele, cognitieve en gedragsmatige verbinding met informatie waarop de aandacht is gericht - is belangrijk in alle omstandigheden waarin mensen informatie verwerken. Omdat het vaak moeilijk is om de aandacht van anderen in te schatten, zouden impliciete en continue metingen van aandacht nuttig kunnen zijn, bijvoorbeeld ter ondersteuning van leraren in online klaslokalen, of van mensen die samenwerken in teams. In dit proefschrift proberen we de overeenkomsten in neurofysiologische reacties tussen personen te benutten als maat voor aandacht. We noemen deze overeenkomsten in fysiologische metingen van de activiteit van de hersenen of het lichaam van verschillende personen fysiologische synchronie. We bouwen voort op studies die aantonen dat fysiologische synchronie zoals gemeten in hersensignalen kan fungeren als een index van aandacht. Hoewel fysiologische synchronie samengaat met aandacht is het onduidelijk hoe variatie in aandacht fysiologische synchronie exact beïnvloedt in de hersenen, hier gemeten via het elektro-encefalogram (EEG), en vooral in het lichaam, hier gemeten via huidgeleiding (elektrodermale activiteit; EDA) en hartslag. Bovendien is het onduidelijk of en hoe het meten van fysiologische synchronie een betrouwbaar instrument is om aandacht te monitoren buiten het laboratorium. In dit proefschrift onderzoeken we hoe verschillende soorten aandachtmodulatie worden gereflecteerd door fysiologische synchronie in de hersenen en het lichaam (deel I) en in hoeverre fysiologische synchronie kan worden gebruikt als hulpmiddel om aandacht te monitoren buiten het lab (deel II).

I: aandachtmodulaties en fysiologische synchronie in hersenen en lichamen

In deel I behandelen we de onderzoeksvraag: "Hoe beïnvloeden verschillende manipulaties van aandacht de fysiologische synchronie in hersenen en lichaam?".

In onze eerste studie registreerden we EEG, EDA en hartslag van proefpersonen die allemaal naar exact dezelfde audiotrack luisterden, bestaande uit een luisterboek waar korte auditieve stimuli, zoals emotionele geluiden en piepjes, doorheen werden gespeeld. Proefpersonen werden geïnstrueerd om ofwel op het luisterboek te letten of op de korte auditieve stimuli te letten. We berekenden inter-subject correlaties als maat voor fysiologische synchronie. Met behulp van deze dataset onderzochten we in **Hoofdstuk 2** in hoeverre inter-subject correlaties hoger zijn wanneer ze berekend worden tussen proefpersonen met dezelfde aandachtinstructie vergeleken met individuen met een andere aandachtinstructie. De resultaten lieten zien dat inter-subject correlaties in alle drie de maten hoger zijn bij proefpersonen met dezelfde aandachtfocus dan bij proefpersonen met een verschillende aandachtfocus. In EEG en hartslag voorspellen inter-subject correlaties de prestatie op een achteraf afgenomen test die de kennis van de stimulus toetst. Met behulp van dezelfde dataset onderzochten we in **Hoofdstuk 3** hoe goed inter-subject correlaties momenten van verhoogde aandacht in de tijd voorspellen. Daarmee probeerden we de invloed van zowel bottom-up stimulus gedreven als top-down aandacht op fysiologische synchronie te bestuderen. We ontdekten dat het presenteren van stimuli tijdens het audioboek gedetecteerd kon worden op basis van inter-subject correlaties, en dat hierbij zowel bottom-up als top-down aandacht een rol speelt.

Aandacht varieert onder invloed van expliciete aandachtinstructie, maar varieert zeker buiten het lab ook impliciet tussen en binnen personen als gevolg van variaties in persoonlijke eigenschappen en variaties in de aandacht over tijd. In **Hoofdstuk 4** onderzochten we of inter-subject correlaties in EEG verschillen in aandacht tussen personen reflecteren die voortkomen uit verschillen in persoonlijke eigenschappen, in deze studie food neophobia of de mate van angst om nieuw voedsel uit te proberen. We vonden inderdaad hogere inter-subject correlaties in EEG bij individuen met een hoger niveau van food neophobia bij het kijken naar een film over de oorsprong en productie van nieuw voedsel. In **Hoofdstuk 5** onderzochten we of variatie in aandacht over tijd binnen personen kon worden gevangen door inter-subject correlaties in EDA en hartslag. We berekenden inter-subject correlaties tijdens tien filmpjes die werden gepresenteerd in de loop van een nacht waarin de proefpersonen niet mochten slapen. Inter-subject correlaties in EDA voorspelden de prestatie op een opeenvolgende vigilantietaak. Dit geeft aan dat inter-subject correlaties variaties in algemene aandacht kunnen weerspiegelen.

II: fysiologische synchronie van het laboratorium naar het dagelijks leven

In deel II gaan we in op de onderzoeksvraag: “In hoeverre kan fysiologische synchronie gebruikt worden als hulpmiddel om aandacht te monitoren in het dagelijks leven?”.

In **Hoofdstuk 6** onderzochten we of inter-subject correlatie als maat voor aandacht ook bepaald kan worden met wearables in plaats van high-end lab-grade apparatuur. Inter-subject correlatie in EDA en hartslag gemeten met wearables bleek minstens net zo goed onderscheid te kunnen maken tussen proefpersonen met een verschillende aandachtinstructie als wanneer gemeten met nauwkeurige laboratorium apparatuur. Voor het monitoren van fysiologische synchronie zijn we dus niet afhankelijk van dure, immobiele laboratoriumapparatuur. De resultaten laten ook zien dat nauwkeurige signalen met een hoge sample frequentie die informatie bevatten over het hele frequentiespectrum niet essentieel zijn voor het monitoren van fysiologische synchronie. In **Hoofdstuk 7** doken we dieper in de voorwaarden in termen van groepsgrootte en opnameduur voor het succesvol monitoren van inter-subject correlaties in EDA en hartslag zoals gemeten met wearables. Zoals verwacht vergroot zowel de groepsgrootte als de opnameduur het percentage deelnemers met significante inter-subject correlaties. Het bleek dat de totale hoeveelheid gegevens (ofwel, de groepsgrootte maal de opnameduur) het percentage deelnemers met significante inter-subject correlaties bepaalt, waarbij het niet uitmaakt door welke verhouding van groepsgrootte en opnameduur deze hoeveelheid gegevens wordt bereikt.

Voor toepassing in real-life omstandigheden kunnen inter-subject correlaties gecombineerd worden met nieuwe machine learning technieken. In **Hoofdstuk 8** combineerden we inter-subject correlaties in EEG, EDA en hartslag met unsupervised learning algoritmes met als doel om individuen met dezelfde selectieve aandachtfocus te clusteren. We maakten gebruik van de dataset uit hoofdstuk 2 en probeerden de twee groepen individuen te clusteren die elk een verschillende aandachtinstructie hadden gekregen, zonder labels te gebruiken met betrekking tot aandachtinstructie voor elk van de individuen. Met EEG werden nauwkeurigheden tot 85% bereikt, hoger dan verwacht op basis van toeval, wat niet mogelijk bleek te zijn bij het gebruik van EDA of hartslag. Echter, het combineren van inter-subject correlaties in EEG, EDA en hartslag in een multimodale aanpak resulteerde in een maximale nauwkeurigheid van 96%. Bovendien bleken bij de multimodale aanpak de resultaten minder afhankelijk van het specifieke algoritme dat werd gebruikt.

In **Hoofdstuk 9** keken we naar een omgeving in het dagelijks leven, en onderzochten we inter-subject correlaties in EDA en hartslag in data die eerder was verzameld bij leerlingen in de klas in twee verschillende studies. We lieten zien dat de inter-subject correlaties hoger zijn bij leerlingen in dezelfde klas vergeleken met in verschillende klassen, maar alleen in één van de twee studies. In deze studie waren meer data beschikbaar met meer personen bij wie EDA en hartslag waren gemeten per klas. De resultaten waren robuuster voor hartslag dan voor EDA. Inter-subject correlaties kunnen dus gemeten worden in de klas, onder voorwaarde dat voldoende hoeveelheden data beschikbaar zijn.

In **Hoofdstuk 10** bespreken we de implicaties van onze bevindingen met betrekking tot de algemene doelstellingen van dit proefschrift. In deel I werd gevonden dat inter-subject correlaties meerdere soorten aandachtmodulaties weerspiegelen, beïnvloed door zowel sensorische bottom-up als top-down aandachtmechanismen. Dit proefschrift vult eerder werk aan door te laten zien dat niet alleen synchrone hersen-, maar ook synchrone lichaamsmaten aandacht-mechanismen weergeven, waarbij zij complementaire aspecten van aandacht weerspiegelen. Top-down aandachtfocus wordt het best weergegeven door synchronie in EEG metingen. Emotioneel opvallende gebeurtenissen worden het best weergegeven door synchronie in lichaamsmaten; EDA en hartslag. De bevindingen in deel II zijn veelbelovend voor het gebruik van fysiologische synchronie buiten het lab, omdat fysiologische synchronie met succes kan worden gemonitord met behulp van wearables in omstandigheden van het dagelijks leven en kan worden gecombineerd met nieuwe unsupervised-learning algoritmen. Wel moet rekening gehouden worden met de vereiste hoeveelheid gegevens die nodig is voor robuuste monitoring, en met de beperkte verklaarde variantie in aandacht. Om fysiologische synchronie verder te ontwikkelen als hulpmiddel om aandacht in real-life omstandigheden te monitoren, moet toekomstig werk zich richten op de toegepast-wetenschappelijke, methodologische en ethische vragen die nog onbeantwoord zijn. Toekomstig werk zou bijvoorbeeld kunnen onderzoeken in hoeverre gedeelde aandacht ten grondslag ligt aan het optreden van fysiologische synchronie in situaties met sociale interacties. Het zou nieuwe, robuuste multimodale maten van fysiologische synchronie kunnen ontwikkelen. Als laatste zou toekomstig werk de ethische aspecten van het gebruik van fysiologische synchronie als hulpmiddel om aandacht te monitoren in kaart kunnen brengen.

1.

General introduction

1.1. Research motivation

Attentional engagement is important in all settings where we humans have to process information. For instance, attentional engagement among students is essential for learning (Newmann et al., 1992; Carini et al., 2006). We specifically refer to 'attentional engagement', as it encompasses not only the ability to focus one's attention on presented information, but also active involvement in processing that information. Engagement is important because without active involvement information that is presented is not well processed and stored (Schmidt, 1992). Parra and colleagues first introduced the concept of attentional engagement as "emotionally laden attention" (Dmochowski et al., 2012). Attfield and colleagues defined the concept user-engagement as "the emotional, cognitive and behavioral connection that exists, at any point in time and possibly over time, between a user and a resource" (Attfield et al., 2011). We define attentional engagement as the emotional, cognitive and behavioral connection with information to which the attention is focused. Teachers continuously try to assess students' attentional engagement to monitor whether the presented information is taken in. Yet, teachers' assessments are limited to behavioral manifestations of student engagement, whereas cognitive and affective aspects of attentional engagement are hardly perceptible (Mandernach, 2015). Assessment is even more problematic in online educational environments, where teachers are seriously hampered to evaluate behavioral manifestations of student engagement (Means et al., 2009).

This example indicates the importance of additional measures of attentional engagement in the classroom. Such measures may also support individuals working together in a team or storytellers aiming to create the most engaging experience. In all such settings, it is desirable to known how the attentional engagement of individuals varies over time. External referees, such as teachers, may estimate attentional engagement, but they do not have access to all manifestations of attentional engagement such that their estimates are not always reliable. Alternatively, one can continuously ask individuals how engaged they are, but this disturbs the attentional engagement itself and is therefore undesirable as well. There thus is a need for alternative, implicit and continuous measures of attentional engagement. We build upon studies indicating that the similarity in the neural responses across individuals may act as an index of attentional engagement (Hasson et al., 2004; Dmochowski et al., 2012, 2014; Ki et al., 2016; Poulsen et al., 2017). In this thesis, we aim to exploit the similarity in neurophysiological responses across individuals as measure of attentional engagement. We refer to this similarity in physiological responses recorded from brain or body across individuals as physiological synchrony. Here and further on the term brain refers to measures reflecting central nervous system activ-

Figure 1-1. Concept of physiological synchrony as tool to provide feedback on the attentional engagement in the classroom to the teacher. Students may be equipped with wearables measuring neurophysiological signals. Physiological synchrony is established among the students, from which the level of attentional engagement can be calculated.

ity. The term body refers to measures reflecting autonomic nervous system activity. Figure 1-1 depicts an example of how physiological synchrony may be used as tool to provide feedback on the attentional engagement in a real-life setting, the classroom. Though physiological synchrony is said to reflect attentional engagement, it is unclear how modulations of attention affect physiological synchrony. Furthermore, it is unclear under which conditions physiological synchrony can be used as a tool to monitor attentional engagement in real-life settings. This thesis is therefore aimed at uncovering how attention may modulate the occurrence of physiological synchrony and to what extent physiological synchrony may be used as tool to monitor attention in real-life settings.

In this introduction, we first provide background information on attentional processing (Section 1.2) and describe how physiological measurements of brain and body, including physiological synchrony, can provide insight in this attentional processing (Section 1.3). We then state the difficulties of exploiting neurophysiological markers of attention in real-life and illustrate why physiological synchrony may be a suitable index of attentional engagement (Section 1.4). We then argue that currently it is unclear how attentional modulations affect physiological synchrony and introduce the research needed (Section 1.5). We also discuss the steps needed for it to be a tool in real-life settings (Section 1.6). We then outline how the next chapters of this thesis contribute to this thesis' aims (Section 1.7).

1.2. Attentional processing and the brain and body

"Everyone knows what attention is" is a popular statement by William James from the late 19th century (James, 1890). Though we indeed have a feeling for its meaning, attention is a complex construct that is actually difficult to grasp. What is clear is that attention has something to do with the allocation of our limited cognitive processing resources to certain parts of the information in the environment around us (Anderson, 1985).

Two distinct functions determine where we focus and sustain our attention on. Bottom-up attention refers to attentional guidance that is driven by external factors related to stimulus saliency. Think for instance of a flashing light alongside the road or the sound of a horn that automatically draws your attention. Top-down attention, on the other hand, refers to internal guidance of attention. Think for instance of the conversation at a party you focus your attention on while ignoring the conversations others are having at that party. A prevailing view in literature is that attentional selection is a process of biased competition (Desimone and Duncan, 1995; Kastner and Ungerleider, 2001), in which the allocation of attention is determined by a mix of bottom-up and top-down attentional processes that influence the relative importance of specific information around us. The relative importance of information is determined by the characteristics of that information, the characteristics of information competing for resources through bottom-up processes and by top-down control acting upon the information characteristics.

Attention is intertwined with emotion. Emotion can alter the relative importance of information thereby affecting attentional prioritization. One of the ways the relative importance is affected is through bottom-up attentional prioritization. Inherent characteristics of emotional stimuli increase the stimulus' salience and can thereby cause bottom-up attentional prioritization over other stimuli (Compton, 2003). Numerous studies have for example reported increased activity in cortical visual processing areas upon presentation of emotionally provocative images compared to neutral images (Lane et al., 1997; Lang et al., 1998; Simpson et al., 2000), indicating that processing of emotional stimuli is prioritized. Some researchers have argued that attention may also be modulated by emotion through top-down processes (Mohanty and Sussman, 2013). Top-down attentional prioritization can be caused by environmental context, past-experience or prior knowledge. For example, happy and threatening facial expressions capture attention when they are the target of search, but not when attending to them is in opposition of task-goals (Williams et al., 2005; Hahn and Gronlund, 2007). Top-down guidance is also active in the process of selective attention, when individuals must select between competing stimuli that differ in their emotional salience (Elliott and Dolan, 1998). Top-down modulation may thus alter the bottom-up attentional capture by emotional information.

A large part of our sensory environment consists of other individuals that also influence the relative importance of presented information and thus influence where we focus and sustain our attention on. As humans we tend to prioritize focusing our attentional resources at information that is also attended to by others. Already as babies we are especially interested in objects that are also attended by others (Bruner, 1985; Baron-Cohen, 1997). This similar focus of attention across individuals is referred to as shared attention. As for attention considered in isolation, in situations of shared attention attentional selection is a process of biased competition where bottom-up and top-down mechanisms affect where attention is focused and sustained on. When listening to a live presentation together with an audience you may for instance be distracted through bottom-up sensory processes by a bird whistling outside or your attention may be drawn to the presentation by an emotionally salient picture. During a garden party, top-down mechanisms allow you to sustain your attention on one conversation, while ignoring speakers not involved in the conversation, known as the cocktail-party effect (Arons, 1992; Marchegiani et al., 2011).

Although the brain is most closely involved in attentional processing, also the body is involved through the peripheral nervous system. The brain and body are inherently and dynamically coupled; our mental processing responds to and changes the state of the body (Critchley et al., 2013). In this regard the concept arousal is important to introduce. By arousal we refer to the physiological state of activation, that ranges from sleep or coma on the one end, to excitement or panic on the other end (Coull, 1998; Critchley et al., 2013). Whereas the brain is generally tapped into to monitor attention, the body is generally tapped into to monitor arousal. Arousal and attention are inherently and dynamically coupled; arousal is said to enhance attention to, and enhance processing of potentially relevant information (Critchley et al., 2013). Indeed, arousal can narrow the attentional focus (Easterbrook, 1959; Loftus et al., 1987), arousal enhances the attentional dwell time on stimuli and delays the disengagement from the stimulus (Fox et al., 2001) and arousal causes more efficient attentional processing (Hansen and Hansen, 1988). Arousal and attention also share a common neural substrate (Critchley, 2002), such that similar brain regions are active during attentional processing as during elicitation of physiological arousal.

1.3. Neurophysiological measures of attention

1.3.1. Individual neurophysiological measures

As long as researchers have been interested in conceptualizing attention they have been interested in measuring attention. As the brain plays a central role in attention it is also the brain that is most often “tapped into” to gain information on the attentional processing of individuals. Since Hans Berger captured the first human electroencephalogram (EEG) in the 1920s (Berger, 1929), EEG features are broadly used as index of brain activity. The EEG captures the electrical activity that originates from potentials of neurons in the cortex of the brain by electrodes placed on the scalp. As the signal is distorted by tissues and bone, it is mainly cortical activity close to the scalp that is captured by the EEG and not so much activity of deeper brain structures (Kandel et al., 2000). Through EEG one can capture the fast brain-responses that occur in initial attention selection processes. The EEG response that is measured directly after presentation of a stimulus is referred to as an event-related potential (ERP). The specific timing and form of the ERP can be informative of the attentional processes occurring in relation to the presentation of the stimulus. A well-known marker of attention is the P300, which refers to a positive deflection in the ERP recorded over the parietocentral cortex that occurs roughly 300 ms after stimulus onset (Smith et al., 1970). It is said to index the allocation of attention to relevant stimuli (Polich, 2007). Upon the presentation of an infrequent distractor tone in a series of frequent tones that draws attention through bottom-up mechanisms, a positive deflection in parietocentral regions is observed (Snyder and Hillyard, 1976). Commonly, the P300 is elicited in an oddball paradigm, where participants detect a presented target throughout a range of presented distractors (Picton, 1992). The P300 is amplified when participants are actively engaged in the task of detecting the target (Picton, 1992), indicating that also top-down attentional processes affect the P300. Furthermore, when participants attend to the auditory stimuli in one ear, but ignore the auditory stimuli in the other ear, an oddball only elicits a P300 if it occurs in the attended ear (Hillyard et al., 1973; Donald and Little, 1981). We discussed before that emotional relevance can contribute to relative stimulus salience. Indeed, also with respect to the P300, emotional pictures or sounds of either high or low valence elicit larger deflections in the event-related potentials than their neutral counterparts (Schupp et al., 2000; Thierry and Roberts, 2007).

In the EEG, not only ERPs can index attention. The power of the alpha oscillations, EEG activity in the frequency range of 8-12 Hz, can also index attention. Unlike ERPs, frequency domain metrics can provide information over longer periods of time. Alpha waves are thought to reflect inhibition of areas of the cortex not primarily in use (Palva and Palva, 2007). Alpha power is for instance increased during lapses in sustained attention (O’Connell et al., 2009) and is increased when individuals focus their attention inward and away from externally presented stimuli (Ray and Cole, 1985; Cooper et al., 2003). During viewing of TV commercials, alpha power was found to correlate negatively with recall of the commercial contents (Reeves et al., 1985), thereby showing predictive value on behavior reflecting attentional engagement.

An alternative mechanism to capture brain activity is functional magnetic resonance imaging (fMRI). It captures activation of brain activity by detecting changes in blood flow (Huettel et al., 2014), with the underlying reasoning that when an area is more active, the oxygenated blood flow to that region will also increase (Logothetis et al., 2001). fMRI is therefore broadly used to localize brain areas that are involved in attentional processes, such as those involved in visual attention (Kanwisher and Wojciulik, 2000) or auditory attention (Pugh et al., 1996). Additionally, with the use of fMRI a network of regions involved in the control of attention has been established (Pessoa et al., 2003). With the development of ultra-high field scanners of more than 7T, one can obtain fMRI with high spatial resolution such that localizations are more precise than previously (Goense et al., 2016). This higher spatial resolution comes at a cost of relatively low temporal resolution (Constable, 2012). The blood oxygenation level dependent (BOLD) response in captured with fMRI responds in the order of seconds, whereas EEG ERPs respond in the order of milliseconds. EEG is therefore much more suited to capture the fast responses corresponding to processes of attention selection.

In the previous section we introduced arousal and its close connection to attention. Measuring arousal may thus also provide insight in attentional processes. Arousal can be captured through the autonomic nervous system; a part of the peripheral nervous system that is predominantly responsible for the involuntary control of internal organs. The autonomic nervous system innervates organs through two main branches: the sympathetic nervous system, often considered to promote a fast fight-or-flight response, and the parasympathetic nervous system, often considered the rest-and-digest system that slowly dampens activation (Cacioppo et al., 2000). It is the sympathetic branch that is mainly involved in preparing the body for the rapid activation that is arousal (Cacioppo et al., 2000). Sympathetic activation is among others expressed through increased heart rate and sweating (Cacioppo et al., 2000; Boucsein, 2012).

Electrodermal activity (EDA), the conductivity of the skin that varies with sweat gland activation, is unlike most indices of autonomic nervous system activity uniquely innervated by the sympathetic branch of the autonomic nervous system (Boucsein, 2012). It therefore serves as a unique index of arousal, which in turn can reflect increased attentional processing. As for ERPs, increased EDA is for instance reported for emotional images or sounds that are attentionally prioritized through bottom-up mechanisms compared to their neutral counterparts (Lang et al., 1998; Bradley and Lang, 2000).

Heart rate can also serve as index of increased arousal (Cacioppo et al., 2000). Indeed, increased heart rate is also reported upon presentation of emotional images or sounds compared to their neutral counterparts (Lang et al., 1998; Bradley and Lang, 2000). Unlike EDA, however, the heart is also innervated by the parasympathetic branch of the autonomic nervous system. This parasympathetic branch can also index attention. Decreased parasympathetic activity is said to reflect focused attention (Cacioppo et al., 1978; Suess et al., 1994). Increased parasympathetic activity is said to predict relaxation and divided attention (Porges, 2001).

1.3.2. Physiological synchrony in brains and bodies as an index of attentional engagement

In the section above we focused on the responses of individual brains or bodies to index attentional engagement. However, there are also studies indicating that the similarity in the neural responses across individuals may act as an index of attentional engagement. Through analysis of the BOLD response in fMRI it was found that the brains of individuals “tick collectively” when they are all presented with the same movie (Hasson et al., 2004, 2008). More precisely, strong inter-subject correlations in the BOLD response were found, not only in visual and auditory areas, but also in higher order associating cortical areas. These inter-subject correlations were higher than expected based on chance level. With the use of EEG, inter-subject correlations in the brain were studied at much higher resolution in time. Also here, inter-subject correlations where higher than expected based on chance level (Poulsen et al., 2017). Further findings suggested that peaks in the inter-subject correlations can occur in correspondence with arousing moments in the film (Dmochowski et al., 2012; Poulsen et al., 2017), suggesting that inter-subject correlations are related to the variations in the level of engagement across the presentation of the film. This was further substantiated when inter-subject correlations in the EEG of a small group of individuals presented with a popular television series were found to predict viewership over the course of the episode worldwide (Dmochowski et al., 2014). There are some indications that inter-subject correlations in the EEG are also associated with behavioral outcomes. Individuals with higher inter-subject correlations during presented narratives were found to answer more questions about the content of these narratives correctly (Cohen and Parra, 2016; Cohen et al., 2018). This is an important finding, substantiating brain-to-brain synchrony as a measure of attentional engagement, that is actually predictive of metrics reflecting attentional performance. Figure 1-2 visually summarizes the concept of inter-subject correlations as measure of attention and indications that it may

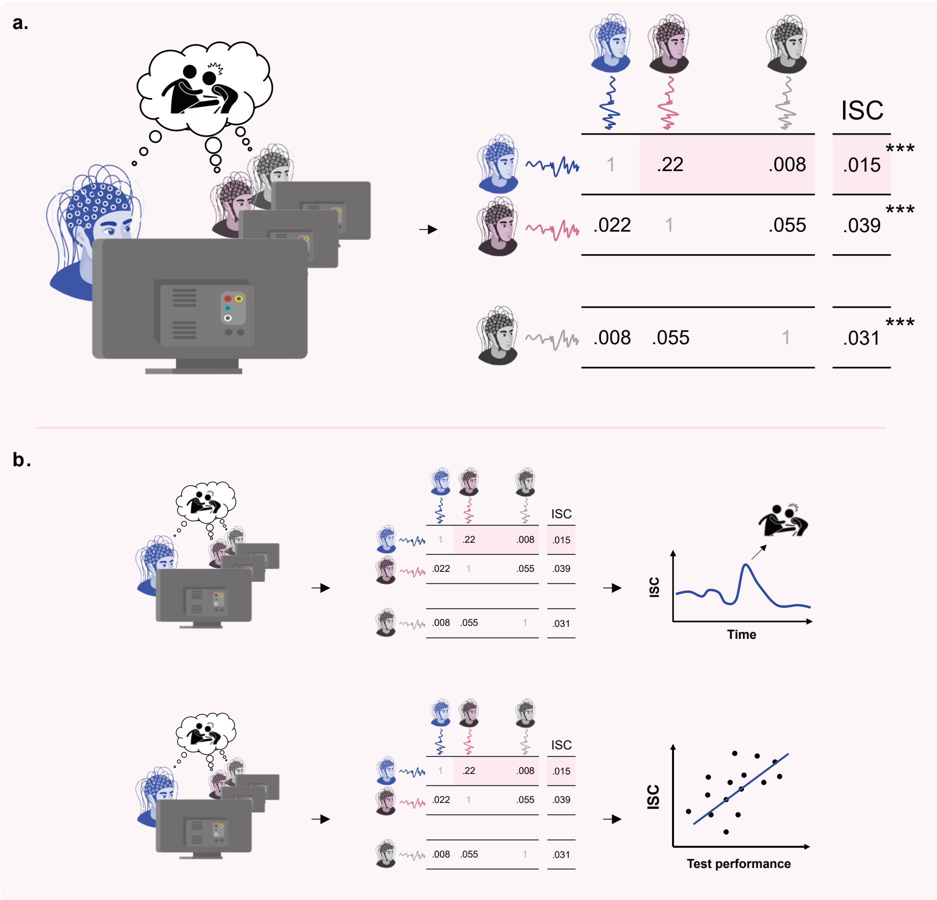

Figure 1-2. Visual depiction of the concept of inter-subject correlations (ISC) and indications that it may be suited as measure of attentional engagement. a. When multiple individuals are presented with the same narrative stimulus, one can compute inter-subject correlations by averaging over the correlations with the neurophysiological signals of all other individuals. Among individuals attending to the same stimulus, these inter-subject correlations are higher than expected based on chance. b. Indications inter-subject correlations may prove suitable as measure of attention. From top to bottom: inter-subject correlations appear in corresponding with arousing moments in the presented film (Dmochowski et al., 2012; Poulsen et al., 2017); individuals with higher inter-subject correlations during presented narratives also answer more questions about the narratives correctly.

be suited as measure of attentional engagement.

In parallel to the studies investigating similarity in brain responses upon the presentation of narrative stimuli, other researchers focused on the similarity in body activity of individuals in social interaction. Researchers uncovered that the physiological activity between two or more individuals can show similar dynamics in new and established relationships across a range of contexts (Palumbo et al., 2017). Among others, physiological synchrony appeared in therapist-patient dyads (Marci et al., 2007), in teammates (Elkins et al., 2009), in singers of a choir (Müller and Lindenberger, 2011), in parent-child dyads (Feldman et al., 2011; Woltering et al., 2015; Suveg et al., 2016) and even pairs of strangers meeting for the first time (Silver and Parente, 2004). The level of physiological synchrony has been reported to be predictive of relationship quality (Levenson and Gottman, 1983), psychotherapy success (Koole et al., 2020), empathy (Marci et al., 2007), team-performance (Elkins et al., 2009), collaborative learning (Malmberg et al., 2019) and more (Palumbo et al., 2017). Various factors have been proposed to underly physiological synchrony, such as shared attention among students in the classroom (Dikker et al., 2017), but also synchronized breathing among couples instructed with the abstract task to mirror each other's physiology (Ferrer and Helm, 2013), empathy among patient and therapist during a therapy session (Marci et al., 2007), shared arousal among performers of a fire-walking ritual and related spectators (Konvalinka et al., 2011; Mitkidis et al., 2015), shared metabolic demands through matched activity or behavior (Palumbo et al., 2017) and environmental influences (Strang et al., 2014).

Though attention is acknowledged to underly the occurrence of neural synchrony upon the presentation of a shared stimulus, for autonomic synchrony in settings of social interaction there is no unified conceptual model explaining the variety of findings. An increasing number of studies links synchronous physiological responses in some social-interactive setting to a psychological construct said to be important for the interaction. However, these studies do not explain why physiological synchrony occurs or why it differs between groups or conditions (Palumbo et al., 2017). Physiological synchrony has been observed in absence to each of the above-mentioned mechanisms and physiological synchrony does not consistently co-occur with the psychological constructs mentioned (Palumbo et al., 2017). There is a need for systematic, experimentally manipulated research to uncover the elements that contribute to the occurrence of physiological synchrony. Until then, interpretation of physiological synchrony results will be limited (Sbarra and Hazan, 2008).

We think that attention is an important element that contributes to physiological synchrony. There are some studies that specifically indicate that findings of physiological synchrony co-occur with shared attention. For instance, Woltering et al. (2015) reported that synchronous heart rate between mother and child related to observed levels of shared attention. Dikker et al. (2017) reported that synchrony in brain potentials increases among students in the classroom as they are more attentive to the classroom activity and thus increase their shared attention towards the externally presented information. However, we also see studies in which shared attention is not specifically reported as contributor of physiological synchrony. Instead, metrics such as psychotherapy success (Koole et al., 2020), empathy (Marci et al., 2007), team-performance (Elkins et al., 2009) or collaborative learning (Malmberg et al., 2019) are reported to be linked with physiological synchrony. Shared attention may be associated with these metrics and with the occurrence of physiological synchrony, such that shared attention may be underlying the reports of physiological synchrony. Take for instance reports of perceived empathy as factor underlying physiological synchrony in therapist-patient dyads (Marci et al., 2007; Palumbo et al., 2017). A simple shift in gaze, reducing the joint attention of therapist and patient, reduces both perceived empathy and physiological synchrony (Marci and Orr, 2006). Though physiological synchrony could thus be linked to perceived empathy, it may actually be shared attention that underlies both the perceived empathy and physiological synchrony. Similar reasoning can be made for collaborative learning. Physiological synchrony occurred among students during collaborative learning (Malmberg et al., 2019). The achievement of joint attention among students is found essential for successful collaborative learning (Barron, 2003). Though physiological synchrony could thus be linked to collaborative learning success, it may again be shared attention that underlies both the collaborative learning success and the findings of physiological synchrony.

1.4. Using neurophysiological measures of attention in real-life settings

Most of the indices of attention discussed above have for a large part been established and further used in lab-based settings. However, research using neurophysiological signals to assess mental state is shifting from lab to life. The research field originating with this shift is referred to as neuroergonomics, with the aim “to study the brain at work and in everyday life” (Parasuraman, 2003; Dehais et al., 2020). The need to move mental-state monitoring to real-life settings has been stressed many times, as results from lab-studies may not transfer to real-life settings (Brouwer et al., 2015b). Compared to lab research there are new challenges to be faced when moving to real-life settings (Dehais et al., 2020).

In real-world settings, physiological signals need to be recorded with unobtrusive, easy-to-use wearable sensors. Such wearable sensors for the monitoring of heart rate (Fuller et al., 2020), EDA (Tronstad et al., 2022) and EEG (Casson, 2019) are increasingly available, but these sensors often have lower signal quality than high-end laboratory equipment. In addition, monitoring in real-life settings often comes at the cost of less experimental control and potentially lower signal-to-noise ratio as increased metabolic demands due to increased motion also affect the data. In such settings it is difficult to separate changes in physiological activity originating from movement from activity originating from changes in mental state (Brouwer et al., 2018; Linssen et al., 2022). All these characteristics of real-life monitoring thus require more robust neurophysiological markers than lab-based settings.

A major challenge for the transition to real-life settings is that many of the indices of attention introduced above are not inherently suited for measurement in real-life settings. The links between the measurements and attention are often complex (Brouwer et al., 2015b) and suffer from large interindividual differences (Näpflin et al., 2007). A common approach to uncover the complex links is the use of supervised learning algorithms (e.g., Aliakbaryhosseinabadiet al., 2017; Hamadicharef et al., 2009; Liu et al., 2013). Such algorithms establish the complex relationships between neurophysiological features and the mental state of interest using labeled training data. This labeled training data usually is a large dataset of neurophysiological features where data are linked to associated labels. Think for instance of a dataset of EEG data where a set of datapoints from different EEG channels recorded at one point in time is labeled as either belonging to an attentive state or distracted state. The goal of the algorithm is learning a function that maps the feature input to the label output based on many examples of input-output pairs. In our example that means learning the function that links the input heart rate to the output being the attentional state. Using such algorithms is challenging already in lab environments, let alone in real-life settings. First, obtaining labeled training datasets is time-consuming, such that sufficient amounts of training or calibration data are hardly available in real-life settings (Brouwer et al., 2015b; Lotte et al., 2018; Brouwer, 2021). In addition, determining a 'ground-truth' mental state that can be used for labeling is very difficult (Brouwer et al., 2015b). Furthermore, even under conditions where reliable training data is available, humans require transparency and explainability if artificial decision-making algorithms are to be used (Fellous et al., 2019). Many supervised learning algorithms are still a black-box, such that decisions are not transparent. Additionally, classification by supervised learning approaches is often limited to a small number of discrete states. Attention, emotion and cognition, however, are of more continuous nature, such that discrete classification is only a rough estimation of the mental state of interest (Zehetleitner et al., 2012; Rosenberg et al., 2013). Last, generalization of trained models to other individuals and context, for instance using transfer learning or domain adaptation, is hard (Lotte et al., 2018).

Although neurophysiological measures, capturing activity of the brain or body, can thus provide insight in the attentional processing of individuals, applying such measures in real-life is difficult. We see physiological synchrony as potential means to circumvent the problems of unsupervised learning. Unlike individuals' physiological responses, the analysis of physiological synchrony is not dependent on labeled training data. It is the degree of similarity in responses across individuals that appears directly proportional with the attentional engagement towards the narrative. There thus is no need to train a model to uncover complex links between the physiological signal and the mental state of interest. This directly proportional relationship also appears to capture the continuous nature of attentional processing, such that one is not limited to two-classes as in a binary classification model.

1.5. Attentional modulations and physiological synchrony in brains and bodies

Though, as discussed, there are indications that physiological synchrony can reflect attentional engagement, up to now it remains unclear how the different attentional processes contribute to the occurrence of inter-subject correlations. If inter-subject correlations indeed reflect overall attentional engagement, is it low-level visual or auditive features attracting attention (Itti and Koch, 2001; Polich, 2007), bottom-up mechanisms related to emotional relevance (Lang et al., 1997) and top-down mechanisms related to attentional instructions (Polich, 2007) that all contribute to the occurrence of inter-subject correlations? If this is so, variations in any of the attentional mechanisms, in time or across individuals, should all result in variation in the inter-subject correlations. The findings of synchronous neural responses upon the presentation of a shared stimulus we presented up to now can mainly be explained by sensory, bottom-up processes of attention. Namely, by such sensory processes attention is especially drawn to arousing stimuli, such that inter-subject correlations are higher for structured movie clips than for unedited footage of people walking in front of an office building and much higher than movie clips scrambled in time (Hasson et al., 2004, 2008; Jääskeläinen et al., 2008; Wilson et al., 2008). Parts of stimuli that strongly attract attention through sensory processes as they are arousing also cooccur with high inter-subject correlations (e.g., Dmochowski et al., 2012; Poulsen et al., 2017). These moments that strongly attract attention also cause high viewership, such that during moments of high viewership when considering a large population worldwide also high inter-subject correlations were found (Dmochowski et al., 2014). We are aware of only one study manipulating attention separate from sensory bottom-up processes and studying the effects on inter-subject correlations. This study is visually depicted in Figure 1-3. Individuals watched narrative movie clips with the instruction to either focus attention on the movie clip or focusing attention inward on a mental arithmetic task while their EEG was monitored. Individuals showed higher inter-subject correlations with others when actively attending to the narrative then when

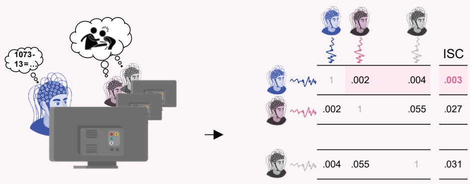

Figure 1-3. Depiction of the study by Ki and colleagues in which attention was manipulated by instructing individuals to either focus attention on the movie clip or focus attention inward on a mental arithmetic task while EEG was monitored. When focusing attention inward, inter-subject correlations with others decreased (Ki et al., 2016).

focusing attention inward on the mental arithmetic task during presentation of the stimulus (Ki et al., 2016). This finding indicates that top-down mechanisms of attentional instruction are involved in the occurrence of inter-subject correlations. This manipulation, however, is very coarse and separates between internally oriented versus externally directed attention. Perhaps also the sensory attention attracting processes are diminished when attention is focused inwards. All in all, the respective contributions of bottom-up sensory processes and top-down processes of attention to the occurrence of physiological synchrony has not been adequately monitored. For instance, it is not clear whether inter-subject correlations can also separate between multiple externally directed attentional foci or whether inter-subject correlations can capture when the attentional abilities of monitored individuals vary over time.

In investigating the respective contribution of bottom-up and top-down attentional processing especial attention should be paid to the comparison of synchrony in brains and bodies in their ability to reflect attentional engagement. Synchronous bodies and synchronous brains may both capture shared attention, but indications for this relation came from different settings for both modalities. In the paragraphs above we described that brain-to-brain synchrony may relate to shared attention based on findings that the cognitive processing of a shared stimulus induces synchronous response in the EEG, fMRI BOLD response or magnetoencephalogram (MEG) (Hasson et al., 2004, 2008; Dmochowski et al., 2012, 2014; Lankinen et al., 2014). On the other hand, body-to-body synchrony may relate to shared attention based on reports of physiological synchrony during social interaction that is predictive of constructs related to shared attention (Palumbo et al., 2017). There are a few exceptions in which body-to-body synchrony is established among individuals presented with the same narrative stimulus. (Golland et al., 2014) studied the dynamics of heart rate and EDA during the emotional cinematic film “Mystic River”. They found synchronous responses in EDA and heart rate driven by the emotional experience of the film. Similarly, (Bracken et al., 2014) found synchronous responses in EDA and heart rate in response to a 100-second emotional film. In both studies, the synchronous responses correlated with moments in the stimulus that were rated as highly emotional. These results indicate that autonomic activity can also synchronize among individuals presented with the same stimulus. It however, remains unclear how inter-subject correlations in the body, such as measured through EDA and heart rate, relate to inter-subject correlations in the brain, such as measured through EEG. The limited literature suggests that body-to-body synchrony is more affected by emotional processing (Bracken et al., 2014; Golland et al., 2014), whereas brain-to-brain synchrony is more affected by lower level factors affecting attentional processing, such as stimulus modality or stimulus saliency (Poulsen et al., 2017). Monitoring synchrony in brains and bodies simultaneously would allow investigation of how brain-to-brain and body-to-body synchrony compare in their reflection of attentional engagement. It could help increasing understanding of respective contributions of different attentional processes to physiological synchrony in each of these modalities. If synchronous bodily measures indeed reflect attentional engagement it could mean that in real-life settings there will not always be the need to monitor EEG and to take the precautions coming with its use. Directly monitoring brains and bodies would also allow multimodal analysis of physiological synchrony. Multimodal sensor fusion allows to exploit the particular strengths of multiple data sources and is a broader trend among neurotechnologies (Brouwer et al., 2015b; Fairclough and Lotte, 2020). Ultimately, multimodal analysis of physiological synchrony may thus result in a more sensitive index of attentional engagement, capturing both low level attentional processing, higher-order cognitive processing and emotional engagement.

Another question that remains about synchronous brains and bodies as potential measure of attentional processing is how the estimated levels of attention correspond to a momentary attentional state only valid during the narrative stimulus presentation or to a more long term pattern of attentional processing corresponding to personal trait. The finding that individuals with higher inter-subject correlations during presented narratives were found to answer more questions about the content of these narratives correctly (Cohen and Parra, 2016; Cohen et al., 2018) may be explained both by state and trait. It may be so that some individuals are more capable of attending to the stimulus and thus answer more questions correctly; a trait explanation, but it may also be so that some individuals were just more attentive at that specific moment in time; a state explanation. Indeed, there are some indications for physiological synchrony to reflect both state and trait. The findings that moments of high neural inter-subject correlations appear in correspondence with arousing moments in the film (Dmochowski et al., 2012; Poulsen et al., 2017) and that this variation in inter-subject correlations over time predicts viewership (Dmochowski et al., 2014) show that physiological synchrony can reflect a momentary attentional state. The observation that individuals with autism spectrum disorders, depression or first-episode psychosis show more varying neural patterns and thus reduced neural inter-subject correlations during naturalistic stimulus presentations than typically developing individuals (Hasson et al., 2009; Salmi et al., 2013; Guo et al., 2015; Mäntylä et al., 2018) shows that physiological synchrony is also affected by personal trait. This is supported by the finding that inter-subject correlations among students in a classroom correlated positively and significantly with self-reported levels of group affinity and empathy (Dikker et al., 2017). When considering physiological synchrony as tool to reflect a momentary attentional trait, as of yet it is not clear whether the attentional state captured during a narrative stimulus is an index of the general attentional processing capabilities at that moment in time. To answer this question it should be investigated whether physiological synchrony monitored during the presentation of a narrative covaries in time with another measure of attentional processing when variation in attentional processing are expected. Also for the influence of personal trait on physiological synchrony unclarities remain. Though it is shown that physiological synchrony is decreased among individuals with autism spectrum disorders, depression or first-episode psychosis, it is unclear how physiological synchrony is affected when the relevance of the presented information varies between individuals due to variations in personal trait.

1.6. Physiological synchrony as a tool in real-life settings

The approach of assessing physiological synchrony seems suited to capture attentional processing in real-life settings. First, by utilizing the physiological responses from multiple individuals there is no need for machine learning models and the training data required for such models. Second, besides in brain measures such as EEG, also in body measures such as EDA and heart rate physiological synchrony may have the potential to reflect attentional engagement. Such body measures can be more easily captured in real-life settings.

However, the pre-conditions for successful monitoring have not yet been established. Currently it is unclear what the minimum requirements for successful monitoring of physiological synchrony are in terms of sensor quality and sample size. There is hardly any research on the relation between synchrony in autonomic physiological measurements and attention, let alone using wearable devices or in real-life settings. Successful monitoring of physiological synchrony as measure of attentional engagement using wearable devices in real-life settings is yet to be demonstrated.

1.7. Research aim, approach and thesis outline

In this thesis we aim to uncover whether different types of attention modulation are captured by physiological synchrony and to what extent physiological synchrony may be used as tool to monitor attention in real-life settings. In part I, we study how different causes of varying attention affect physiological synchrony in brains and bodies. In part II we try to bridge the gap from lab to life, by determining minimally necessary conditions for successful monitoring of inter-subject correlations and exploring whether and to what extent information is lost when relying on wearable sensors rather than high-end equipment and recording in real life conditions .

I: attentional modulations and physiological synchrony in brains and bodies

In part I, we address the research question: "How do different manipulations of attention affect physiological synchrony in brains and bodies?" This part is focused on manipulations of attention and their potential impact on physiological synchrony in brains and bodies. Figure 1-4 provides a visual depiction of the research presented in part I.

We here focus on inter-subject correlations as measure of physiological synchrony, therewith quantifying synchrony as the simple linear instantaneous correlation in the signals between individuals, following previous work (Dmochowski et al., 2012, 2014; Cohen and Parra, 2016; Poulsen et al., 2017; Cohen et al., 2018). There are much more complex analyses to quantify synchrony. Most such analyses are aimed at capturing delayed or non-linear relationships that may be expected in settings where the experiences of individuals are not expected to occur simultaneously, such as in leader-follower behavior or an audience attending to a speaker (Ferrer and Helm, 2013; McAssey et al., 2013; Liu et al., 2016; Helm et al., 2018). In our settings, however, all individuals are presented with the same information at the same time. We therefore expect linear instantaneous correlations to capture physiological synchrony sufficiently. More complex non-linear analysis would also require more data to be collected to set the additional modeling parameters, such that inter-subject correlations are the more suitable approach (Pérez et al., 2021).

In our collection of studies we monitor inter-subject correlations in EEG, EDA and heart rate. Though synchronous activation across individuals upon presentation of the same narrative stimulus has also been found in fMRI (Hasson et al., 2004, 2008) and MEG (Lankinen et al., 2014), these modalities do not allow measurements in real-life settings.

We previously introduced how attention and emotion are entangled. Emotional information can attract attention and emotional information can be experienced differently when attended. It is therefore not in the scope of this thesis to disentangle attention from emotion in their respective association with physiological synchrony. Rather, we aim to identify how sensory, low-level, bottom-up mechanisms and higher-order cognitive, top-down mechanisms of attention together affect physiological synchrony and how the respective contribution of these processes may vary across brains and bodies.

In *Chapter 2* we work towards this aim by answering the research question 'How do inter-subject correlations in EEG, EDA and heart rate compare in their ability to reflect attentional engagement?' We investigate to what extent individuals show higher inter-subject correlations with individuals with the same selective attentional focus compared to individuals focusing on different stimulus aspects. This allows us to investigate the ability of inter-subject correlations in each of these modalities capture top-down selective attentional processes. Relating results to stimulus retention helps us understand the meaning of varying levels of physiological synchrony for attentional ability. We hypothesize that individuals show higher inter-subject correlations with other individuals with the same versus different selective attentional focus.

In *Chapter 3* we aim to study the respective influences of bottom-up sensory and top-down higher-level attentional processes on inter-subject correlations in EEG, EDA and heart rate. We aim to answer the research question: 'How well do inter-subject correlations in EEG, EDA and heart rate predict the occurrence of attentionally engaging moments in time and do results depend on the attentional instruction of participants, the type of stimuli and the physiological measure used?' We investigate to what extent these inter-subject correlations are responsive to the occurrence of the interspersed stimuli. As we instructed half of the participants to focus on these sounds and the other half not to, we can separate effects of higher-level top-down and sensory bottom-up attentional processing on the inter-subject correlations. In addition, this study design allows us to investigate how the sensitivity of inter-subject correlations in response to higher-order top-down and low-level sensory events vary across EEG, EDA and heart rate.

In *Chapter 4* and *Chapter* 5 we investigate to what extent inter-subject correlations can capture interpersonal and intrapersonal variations in attentional engagement. By doing so, we aim to investigate whether inter-subject correlations can capture changes in the momentary attentional state and whether inter-subject correlations can capture differences in attentional processing related to personal trait. In *Chapter 4* we aim to answer the research question:

'How are inter-subject correlations in EEG related to variations in the relevance of presented information across individuals?'

In *Chapter 5* we investigate to what extent inter-subject correlations in EDA and heart rate capture intrapersonal variations in attentional engagement induced by sleep deprivation verified by a response-time metric extracted from a low-level button press task. We aim to answer the research question: 'How do inter-subject correlations in EDA and heart rate covary in time with another metric of attentional processing?'

I: attentional modulations and physiological synchrony in brains and bodies

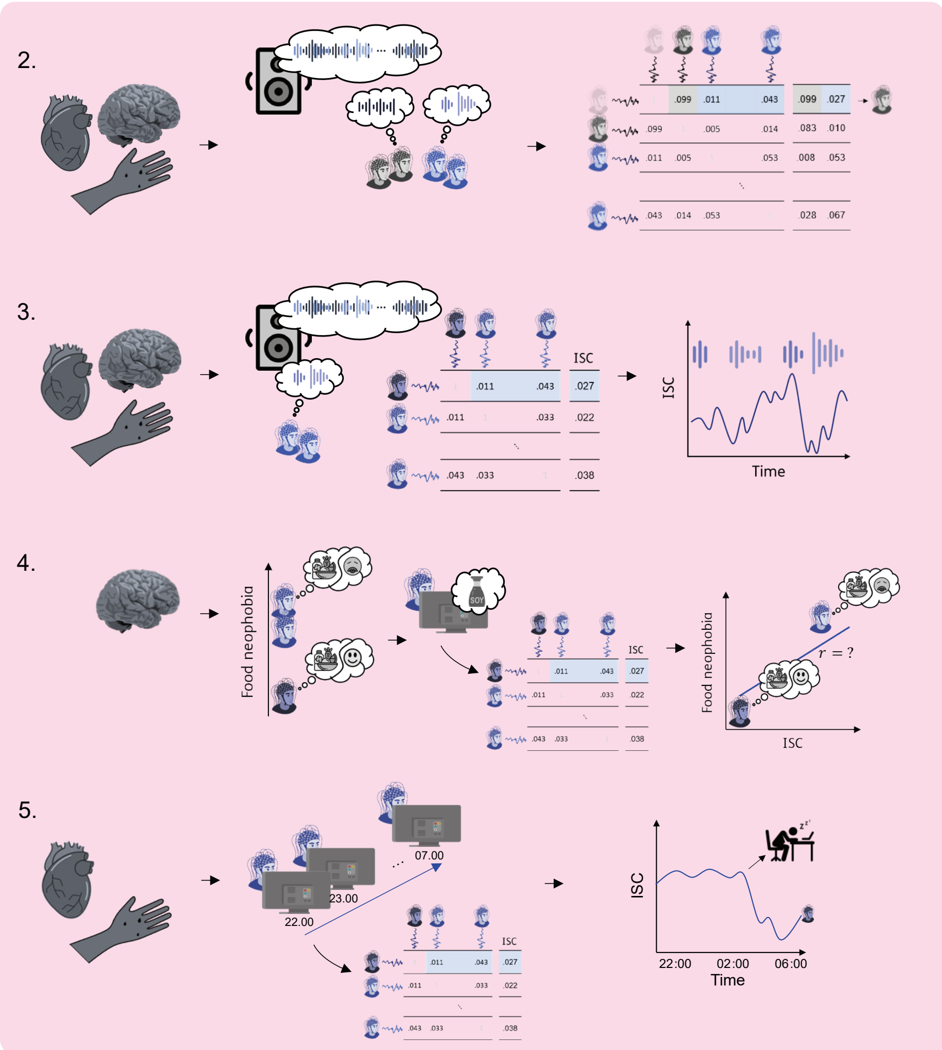

Figure 1-4. Graphical depiction of the research presented in part I.

II: physiological synchrony from lab to life

In part II of this thesis, we address the research question: "To what extent may physiological synchrony be used as tool to monitor attention in real-life settings?" We explore what the requirements are for monitoring physiological synchrony in ambulatory environments. In this part the use of wearable sensors for monitoring physiological synchrony is explored and specifically the conditions that are needed to monitor synchrony robustly. In addition, we move toward more applied methodologies by investigating how unsupervised learning algorithms can assist in clustering individuals sharing attentional focus. Figure 1-5 provides a visual depiction of the research presented in part II.

In *Chapter 6* we compare inter-subject correlations in EDA and heart rate from wearable and high-end laboratory-grade equipment. We aim to answer the research question: "How do inter-subject correlations in EDA and heart rate from wearable and high-end laboratory-grade equipment relate in their abilities to identify individuals with the same selective attentional focus?"

In *Chapter 7* we then further study the preconditions for successful monitoring of inter-subject correlations in EDA and heart rate recorded with wearable devices. We aim to answer the research questions: "How do recording length and group size affect the robustness of inter-subject correlations?" Specifically, we investigate how the recording length and group size affect the percentage of participants that show inter-subject correlations exceeding chance level. We expect that increasing recording length and group size both positively affects this percentage, but have no hypothesis recording the specific relation.

In *Chapter 8* we combine unsupervised learning algorithms with inter-subject correlations in EEG, EDA and heart rate. We aim to answer the research question: "How well can individuals with the same selective attentional focus be clustered together without using information on the attentional condition of any of the individuals by combining inter-subject correlations and unsupervised learning algorithms?"

In *Chapter 9* we move towards real-life settings, by monitoring inter-subject correlations in EDA and heart rate among students in the classroom. As a proof of concept, we aim to identify whether students show higher inter-subject correlations with others in the same compared to different classrooms. We answer the research question: "How well can inter-subject correlations in EDA and heart rate distinguish between individuals in the same versus in different classrooms?"

Finally, in *Chapter 10*, the general discussion, the implications of our findings regarding the overall aims of this thesis are discussed. We also discuss limitations of the current work and future directions.

II: physiological synchrony from lab to life

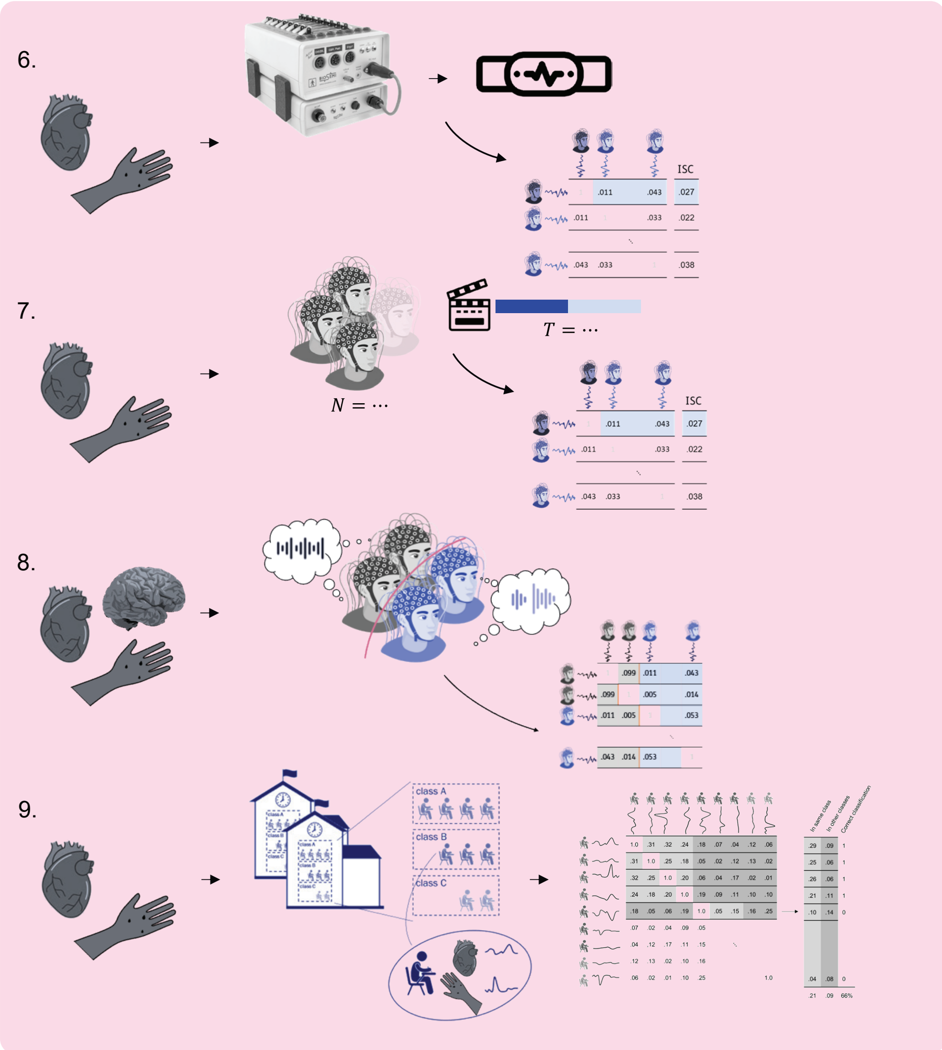

Figure 1-5. Graphical depiction of the research presented in part II.

Part I:

Attentional modulations and physiological synchrony in brains and bodies

2.

Physiological synchrony in EEG, electrodermal activity and heart rate reflects selective auditory attention

Stuldreher, I.V., Thammasan, N., van Erp, J.B.F., Brouwer, A.M. (2020). Physiological synchrony in EEG, electrodermal activity and heart rate reflects selective auditory attention. *Journal of Neural Engineering*, ***17***, 046028. doi: 10.1088/1741-2552/aba87d

Abstract

Objective: Concurrent changes in physiological signals across multiple listeners (physiological synchrony – PS), as caused by shared affective or cognitive processes, may be a suitable marker of selective attentional focus. We aimed to identify the selective attention of participants based on PS with individuals sharing attention with respect to different stimulus aspects.

Approach: We determined PS in electroencephalography (EEG), electrodermal activity (EDA) and heart rate of participants who were instructed to either attend to an audiobook or to interspersed auditory events such as affective sounds and beeps that attending participants needed to keep track of.

Main results: Even though all participants heard the exact same audio track, PS in EEG and EDA, but not in heart rate, was higher for participants when linked to participants with the same attentional instructions than when linked to participants instructed to focus on different stimulus aspects. Comparing PS in EEG between a participant and members from the same or the different attentional group allowed for the correct identification of the participant's attentional instruction in 96% of the cases. For both EDA and heart rate this was 73%. Even when only data was included coming from 'narrative only' time intervals, classification performance was above chance level for EEG and heart rate, though not for EDA. PS with respect to the attentional groups predicted performance on post-audio questions about the groups' stimulus content.

Significance: Our results show that selective attention can be monitored using PS, not only in EEG, but also in EDA and heart rate. These results are promising for real-world applications, where wearables measuring peripheral metrics like EDA and heart rate may be preferred over EEG sensors.

2.1. Introduction

Selective attentional engagement is critical for efficient and effective learning (Jiang and Chun, 2001). Sustaining attention to a single continuous stream of information is a constant challenge, especially when competing sensory stimuli are present. Individuals who suffer from learning disabilities in particular have troubles narrowing the focus of their attention (Richards et al., 1990). To assist students with learning disabilities or to evaluate learning materials, it would be helpful to continuously and implicitly measure selective attentional engagement. Such continuous and implicit measures of attention may be extracted from physiological signals, such as brain potentials as measured through the electroencephalogram (EEG), electrodermal activity (EDA) or heart rate. Rather than investigating responses for specific events and individual observers as is commonly done in research using physiological measures to monitor mental state, one may also determine the relationship between individuals' physiological measures. Interpersonal analyses of physiological synchrony (PS) as analyzed though inter-subject correlations (ISC) in brain signals were found to be a strong marker of shared attentional engagement to narrative movie or audio clips (Hasson et al., 2004, 2010; Hanson et al., 2009; Dmochowski et al., 2012). Note that we refer to the term physiological synchrony not only to cover synchrony in peripheral measures, such as EDA and heart rate, but also to cover synchrony in neural measures, such as EEG. Moments of high PS correlated strongly with general expressions of interest and attention (Dmochowski et al., 2014), supporting the validity of PS as a measure of attention. In addition, individuals with neural responses that were more synchronous to the group that was attending to a narrative stimulus, remembered more information about this stimulus (Cohen and Parra, 2016; Cohen et al., 2018). A first step toward real-time inference of engagement in the classroom was taken by (Poulsen et al., 2017), who demonstrated that shared attention to narrative stimuli may be quantified using PS in wearable EEG among students in a classroom. Other recent studies in the educational domain also found promising results regarding neural PS as measure of attentional engagement. PS in EEG among students reflected classroom engagement and social dynamics (Dikker et al., 2017). Further results suggested that the interaction between an instructor and a learner is reflected by the degree of PS in neural activity between the two (Zheng et al., 2018; Bevilacqua et al., 2019; Pan et al., 2020). In some cases the degree of PS between an instructor and a learner predicted learning outcomes (Zheng et al., 2018; Liu et al., 2019; Pan et al., 2020), although others did not find this relation (Bevilacqua et al., 2019). PS in brain activity has also been related to attentional engagement in other settings, such as in responses to political speeches (Schmälzle et al., 2014) or music (Madsen et al., 2019).

There is also a body of literature on synchrony in measures of the peripheral autonomic nervous system, such as heart rate and EDA, reviewed by (Palumbo et al., 2017). Rather than as indicators of shared attention, these have generally been interpreted as indicators of some form of connectedness between people or as indicators of shared affective and cognitive processes related to specific events in the world. Studies are conducted in a broad range of application areas, including psychotherapy (Koole et al., 2020), marital counseling (Wilson et al., 2018; Tourunen et al., 2020) and collaborative learning (Malmberg et al., 2019). PS in autonomic activity has for example been associated with relationship quality of romantic couples, empathy in therapist-patient dyads and team-performance of team-mates (Levenson and Gottman, 1983; Marci et al., 2007; Elkins et al., 2009). Findings in this literature may also have been driven by mechanisms of shared attention. Shared attention has been emphasized in models of social rapport during social interaction (Tickle-Degnen and Rosenthal, 1990).

In our view there are two gaps in current literature. First, it remains unclear how PS in central and peripheral modalities are related when capturing shared attentional engagement. In fact, in earlier work we did not find any studies concurrently monitoring PS in EEG and measures of autonomic nervous system activity (Stuldreher et al., 2019). For future real-world studies and applications, autonomic measures may be preferred over neural measures as they can more easily be monitored through wearable sensors that are broadly available (e.g., Garbarino et al., 2014). The second gap in current work is studying PS as a measure of attention during selective attention, i.e., under conditions where an individual has to focus on one type of stimulus when other stimuli are concurrently present. A specific, famous example of such a situation is the cocktail party problem (Cherry, 1953), where listeners are capable to selectively attend to one of several simultaneously speaking voices. Research has shown that EEG in relation to sound characteristics can indicate which speaker the participant attended to in such problems using single-trial analysis (Horton et al., 2014; O'Sullivan et al., 2015). Even though PS is not dependent on sound characteristics, it may thus be expected that PS for individuals attending to the same speaker will be stronger compared to situations in which different speakers are attended to. In addition, while PS has not been used to distinguish the focus of attention on two concurrently presented stimuli, it has been shown that PS in EEG distinguishes conditions in which individuals attend or do not attend to external stimuli (Ki et al., 2016; Cohen et al., 2018).

In the current work we try to fill the two abovementioned gaps. We compare PS across EEG, EDA and electrocardiographic inter-beat interval (IBI). We aim to determine selective auditory attention of individuals who are all presented with the same auditory stimulus and are all attending to it, but to different stimulus aspects. Reminiscent to a classroom setting where students hear the teacher talk as well as hearing other, potentially interesting sounds like the school bell or whispering students, we present our participants with an audiobook, interspersed with short auditory stimuli. Participants are instructed to attend either to the narrative of the audiobook (audiobook-attending – AA), or to the interspersed stimuli (stimulus-attending – SA). Unlike the popular cocktail party paradigm, the two stimulus streams used in the current paradigm are not homogeneous. We selected this custom design for two main reasons. First, the selected design roughly mimics the environment of a dynamic classroom, where a long, continuous lecture is interspersed with short, inconsistent distractors. During the continuous lecture of a teacher, some students focus continuously to the lecture. This group of students is represented by the AA group in our current design. Another group of students may focus their attention more to other environmental events, such as whispering students or cars driving by outside. This group is represented by the SA group in our current design. It can be argued that during realistic cocktail parties, listeners also rather filter one speaker out of a great variety of sounds rather than out of a homogeneous collection of voices. For this reason, we also chose to present the audiobook and the interspersed stimuli both to the left and right ear, rather than one stream of sound in each ear. Second, including multiple stimulus sets that intersperse the audiobook allowed us to investigate whether PS may occur more reliably during a specific type of stimulus than during other stimulus types.

We formulated the following research questions. First, is PS of participants higher when paired with participants that received the same selective attentional instructions (within-group) than with participants that received instructions to focus on the other stimulus aspects (between-group)? If this is indeed the case, our second research question is whether the selective attentional focus of a participant can be identified based on synchrony in physiological responses with participants that have known attentional instructions. While participants in the SA group are instructed to ignore the narrative, this is probably hard to do at times without concurrent short-stimuli. Our third research question therefore is: does zooming in on intervals with interspersed stimuli increase classification accuracy? We hypothesize that classification of the selective attentional focus is enhanced when zooming in on intervals with interspersed stimuli, whereas zooming in on intervals with data from ‘audiobook only’ intervals results in decreased classification performance. We also expect that results are different for different measures. Because mental workload mainly affects EEG (Hogervorst et al., 2014), we hypothesize the group-distinguishing capability of PS in EEG to work well during the beep counting task. As emotional stimuli have been strongly related to sympathetic nervous system activity as measured through

EDA (Bradley and Lang, 2000; Boucsein, 2012), we hypothesize the group-distinguishing capability of PS in EDA to work well during the affective sounds. Our fourth research question is: does PS of participants paired with participants attending to a stimulus aspect correlate with performance metrics reflective of paid attention? Based on earlier work relating the degree of synchrony with an attending group to stimulus retention (Cohen and Parra, 2016; Cohen et al., 2018), we hypothesize that this is the case. To get a grip of what drives possible effects of attentional instruction on PS, we also obtain physiological response traces locked to the onset of interspersed stimulus events. Our final research question is: do traces of EEG, EDA and IBI locked to the onset of the interspersed stimuli differ between the attentional groups? We hypothesize stronger deflections in EEG, EDA and IBI traces for participants attending to the interspersed stimuli than for participants attending to the narrative of the audiobook.

2.2. Methods

2.2.1. Participants

Before recruitment, the study was approved by TNO's Institutional Review Board. The approval is registered under the reference 2018–70. Twenty-seven participants (17 female), aged between 18 and 48 ($M = 31.6$, $SD = 9.8$) years, were recruited from the institute's participant pool. Prior to the experiment all participants signed informed consent, in accordance with the Declaration of Helsinki. After the experiment they received a small monetary reward for their time and travel costs. None of the participants indicated problems in hearing or attention. Participants were randomly assigned to one of the two experimental groups. Data of one participant were discarded due to failed physiological recordings, resulting in equal group size.

2.2.2. Materials

EEG, EDA and electrocardiogram (ECG) were recorded at 1024 Hz using an ActiveTwo Mk II system (BioSemi, Amsterdam, Netherlands). EEG was recorded with 32 active Ag/-AgCl electrodes, placed on the scalp according to the 10–20 system, together with a common mode sense active electrode and driven right leg passive electrode for referencing. The electrode impedance threshold was set at 20 kOhm. For EDA, two passive gelled Nihon Kohden electrodes were placed on the ventral side of the distal phalanges of the middle and index finger. For ECG, two active gelled Ag/-AgCl electrodes were placed at the right clavicle and lowest floating left rib. EDA and heart rate were also recorded using wearable systems (Movisens EdaMove 4 and Wahoo Tickr, respectively). These data are not discussed here.

2.2.3. Stimuli and design

Participants performed the experiment one by one. Each participant listened to the same audio file, composed of a 66 min audiobook (a Dutch thriller 'Zure koekjes', written by Corine Hartman) interspersed with other auditory stimuli. The 13 participants in the AA group were asked to focus on the narrative of the audiobook and ignore all other stimuli or instructions and the 13 participants in the SA group were asked to focus on the other stimuli, perform accompanying tasks and ignore the narrative. Figure 2-1(a) and (b) visualizes the experimental paradigm and participant instructions. The order of interspersed affective sounds and beeps was randomly determined, but was identical for each participant. Intervals between the end of one stimulus and the onset of the next stimulus varied between 35 and 55 s (M = 45, SD = 6.1 s). In the supplementary material (Tables 1–3) (available online at stacks.iop.org/JNE/17/046028/mmedia) the exact types and order of interspersed stimuli can be found. The short auditory stimuli were affective sounds, blocks of beeps, and the instruction to sing a song.

Affective sounds were taken from the second version of the International Affective Digitized Sounds (IADS) (Bradley and Lang, 2007a). The IADS is a collection of six second acoustic stimuli that have been normatively rated for emotion, based on valence, arousal and dominance. Examples of stimuli are the sound of a crying baby or a cheering sports crowd. We selected 12 neutral sounds (IADS number 246, 262, 373, 376, 382, 627, 698, 700, 708, 720, 723, 728), 12 pleasant sounds (110, 200, 201, 202, 311, 352, 353, 365, 366, 367, 415, 717) and 12 unpleasant sounds (115, 255, 260, 276, 277, 278, 279, 285, 286, 290, 292, 422) based on their normative ratings of valence and arousal.

Beeps were presented in blocks of 30 s, with every 2 s a 100-ms high (1 kHz) or low (250 Hz) pitched beep. SA participants were asked to separately count the number of high and low beeps presented in a block, as in (De Dieuleveult et al., 2018). This task was practiced with them beforehand. In total, 27 blocks of beeps were presented.

At the end of the audiobook, the instruction was presented to sing a song aloud after the subsequent auditory countdown reached 0. This instruction had to be followed by the SA group and was expected to induce stress and a strong increase in EDA and a strong decrease in IBI (Brouwer and Hogervorst, 2014). Physiological data obtained after this stimulus are discarded in further analysis as some participants started talking during or right after this stimulus. In total, we consider 3800 s of data in further analyses, out of which 1026 s involved concurrent presentation of the audiobook and interspersed stimuli.

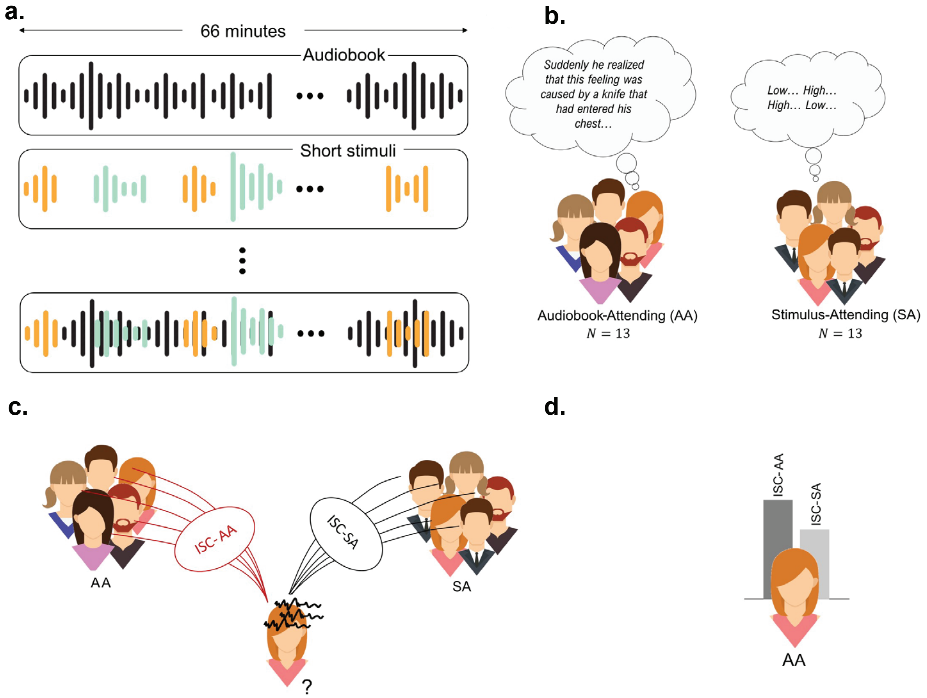

Figure 2-1. Overview of the experimental paradigm. a. The paradigm consists of a narrative auditory stimulus of 66 min that is interspersed with short auditory cognitive (depicted in green) and affective (in orange) stimuli. b. Half of the participants were instructed to focus their attention on the audiobook (AA group), while the other half of the participants were instructed to focus on the interspersed stimuli (SA group). c. For each participant, the inter-subject correlations (ISC) of her/his EEG, electrodermal activity and inter-beat interval with those of all other participants in the AA condition (ISC-AA) and SA condition (ISC-SA) are computed. d. If the physiological responses of a participant are more correlated with those of participants in the AA group, the participant is classified as a AA participant, if the responses are more correlated with those of participants in the SA group, the participant is classified as a SA participant.

After the experiment, all participants were asked to answer two questionnaires. In the first questionnaire, participants used a slider on a horizontal visual analogue scale running from 'not at all' to 'extremely' to rate their mental effort, distraction and emotion during the short emotional sounds, and the level of stress induced by the sing-a-song assignment. The second questionnaire was on the content of the stimuli: participants were asked to report as many emotional sounds as they could remember, they were asked to estimate the average number of beeps in a block, and they were asked questions about the content of the narrative. The questions and correct answers can be found in the supplementary material (Table 4).

2.2.4. Analysis

2.2.4.1. Pre-processing

Data processing was done using MATLAB 2019a software (Mathworks, Natick, MA, USA). EEG was pre-processed using EEGLAB v14.1.2 for MATLAB (Delorme and Makeig 2004). To remove potentials not reflecting sources of neural activity, like ocular or muscle-related artefacts, logistic infomax independent component analysis (Bell and Sejnowski, 1995) was performed. EEG was first downsampled to 256 Hz and high-pass filtered at 1 Hz. This relatively high cut-off frequency has shown to work better for independent component analysis compared to lower cut-off frequencies (Winkler et al., 2015). Data were then notch filtered at 50 Hz, using the standard finite-impulse-response filter implemented in EEGLAB function pop_eegfiltnew. Channels were re-referenced to the average channel value. Independent component analysis was performed and the Multiple Artifact Rejection Algorithm (Winkler et al., 2011) was used to classify artefactual independent components, i.e. components not reflecting sources of neural activity, but ocular or muscle-related activity. Components that were marked as artefactual were removed from the data. Then, samples whose squared amplitude magnitude exceeded the mean-squared amplitude of that channel by more than four standard deviations were marked as missing data ('NaN') in an iterative way with four repetitions to remove outliers. By doing so, 0.8% of data were marked as missing.

EDA was downsampled to 32 Hz. The fast changing phasic and slowly varying tonic components of the signal were extracted using Continuous Decomposition Analysis as implemented in the Ledalab toolbox for MATLAB (Benedek and Kaernbach, 2010). In further analyses we use the phasic component, as this component of the EDA signal is mainly related to responses to external stimuli.

ECG measurements were processed to acquire the inter-beat interval (IBI—inversely proportional to heart rate). After downsampling to 256 Hz, ECG was high-pass filtered at 0.5 Hz. R-peaks of the ECG signal were detected following (Pan and Tompkins, 1985), resulting in a semi-timeseries of consecutive IBIs. This IBI semi-time series was transformed into a timeseries by interpolating consecutive intervals at 32 Hz.

2.2.4.2. Computation of inter-subject correlations as a measure of physiological synchrony

For EEG, we computed ISC in the time-domain as a measure of PS. Rather than treating EEG signals separately, the ISC were evaluated in the correlated components of the EEG (Dmochowski et al., 2014, 2012). The goal of the correlated component analysis is to find underlying neural sources that are maximally correlated between participants, using linear combinations of electrodes. The technique is similar to the more familiar principle component analysis, differing in that projections capture maximal correlation between sets of data instead of maximal variance within a set of data. EEG data from each participant were projected on the component vectors. Participant-to-group ISC were then computed as the sum of correlations in the first three component projections. Correlations in higher order projections are usually discarded as they are close to chance level correlations (Ki et al., 2016).

For EDA and IBI, we also computed ISC in the time-domain as a measure of PS. We followed the approach of (Marci et al., 2007). Pearson correlations were calculated over successive, running 15 s windows at 1 s increments. The overall correlation between two responses was computed as the natural logarithm of the sum of all positive correlations divided by the sum of the absolute values of all negative correlations. Participant-to-group ISC were computed by averaging over all correlations with other participants in the group.

2.2.4.3. Identifying selective attention through comparing within-group and between-group physiological synchrony

To investigate whether within-group PS was higher than between-group PS, we computed for each participant the ISC with participants with the same attentional instructions (within-group) and the ISC with participants with other attentional instructions (between-group). For EEG, correlated component vectors were extracted from both the AA and SA groups. Data from each participant were then projected on both of these component vectors. Data from the to-be tested participant were excluded in the component extraction step of EEG. We then tested whether the ISC scores were normally distributed using the Shapiro-Wilk tests for both the AA and SA groups in EEG, EDA and IBI. If the null hypothesis of normally distributed data was not rejected, we conducted paired-sample t-tests to test for differences between within-group PS versus between-group PS, otherwise the non-parametric Wilcoxon signed rank test was used.

To examine how well PS can be used to identify whether an individual participant had been attending to the narrative of the audiobook or to the interspersed stimuli, we also classified each participant into the attentional condition that he or she showed more synchrony with, for EEG, EDA and IBI. Chance level classification performance was determined using surrogate data with 100 renditions of randomly shuffled attentional condition labels. For each shuffle the same procedure as above was followed. Two sample one-tailed t-tests were conducted to test whether classification performance was above chance level. Figure 2-1(c) and (d) visualizes the classification paradigm.

2.2.4.4. Influence of interspersed stimuli on the difference between within-group and between-group physiological synchrony

We hypothesized that differences between attentional groups are present during interspersed stimulus presentation, but not or to a lesser extent when the audiostream only contains the audiobook. Therefore, we zoomed in on intervals with concurrent audiobook and stimulus presentation, and, as a comparison, audiobook parts without interspersed stimuli. We computed within-group and between-group PS three extra times; when considering only physiological responses recorded during blocks of beeps; during presentation of affective sounds; and during parts of the audiobook without interspersed stimuli. For EEG, we extracted new correlated components in each of the three data selections before computing correlations in the projections. Procedures that followed were identical to those for the whole narrative stimulus; we used paired-sample t-tests or the non-parametric Wilcoxon signed rank test to test for differences between within-group and between-group PS and we classified the attentional condition of each participant as the condition of the attentional group he or she showed more synchrony with.

2.2.4.5. Behavioral performance and its association with physiological synchrony

To examine whether participants followed their attentional instructions, we tested if the performance metrics on the questionnaires about the content of the interspersed stimuli and narrative differed between groups using non-parametric Wilcoxon rank sum tests. We then tested whether higher PS with respect to an attentional group also results in higher performance on the post-audio questions about that group's stimulus content. Outliers in the performance metrics were first removed. Three participants were left out for this analysis because of outlying performance data. Two of these participants reported '395' and '110', respectively, to the number of beeps, while the correct answer was 15; one correctly identified 25 of the 36 IADS sounds. Values were then ranked based on relative performance across all participants: the participant performing best on a question received score 26, the worst performing participant received score 1. This was done for each of the three metrics of performance (correct questions of the narrative, number of reproduced affective sounds and absolute deviation from the correct number of beeps). The correlations between these performance scores with ISC toward the AA group and ISC toward the SA group were computed. We also tested whether a large difference between PS with respect to both attentional groups in a participant leads to a large difference between the performance metrics reflective of attention toward the AA and SA groups. To do so, for each participant ISC toward the AA group was subtracted from ISC toward the SA group. The score corresponding to narrative performance was subtracted from the average of the affective sound score and beeps score (e.g. for a participant with a score of 10 for the narrative, a score of 26 for the affective sounds and a score of 16 for the beeps this thus results in a score of $10-(26 + 16)/2 = -11$). We computed the correlations between the subtracted ISC metric and the subtracted performance-metric.

As an exploratory comparative analysis, we also computed correlations between the self-reported measures of mental effort, distraction and emotion with the performance metrics reflective of paid attention.

2.2.4.6. Short-stimulus response traces

To get an understanding of what drives possible effects of attentional instruction on PS, response traces were extracted for EEG, EDA and IBI in response to the beeps and affective sounds. EEG event-related potentials were obtained from the parietal site on the anterior-posterior midline of the scalp (Pz). We chose this location as responses here have been shown to reflect attentional, emotional and working memory processes (Polich and Kok, 1995; Polich, 2007; Hettich et al., 2016). Pre-processed EEG was cut in 1100 ms short stimulus-locked epochs (100 ms pre-stimulus onset, 1000 ms post-stimulus onset) and baseline corrected based on the average value of the 100 ms before stimulus onset. For the blocks of beeps, responses were locked to each beep in a block and then averaged over all beeps in that block. Grand-average potentials were obtained by averaging over all participants in each condition. Running independent-sample t-tests were conducted to test for significant between-group differences over time. Tests were adjusted for multiple comparisons by controlling the false discovery rate (FDR) using the Benjamini–Hochberg procedure (Benjamini and Hochberg, 1995). In this procedure, p-values are sorted and ranked. The smallest value gets rank 1, the largest rank N. All p-values are then multiplied by N and divided by their rank to obtain adjusted q values. The FDR threshold was set at $q = .05$. Phasic EDA and IBI were cut in 31 s stimulus-locked epochs (1 s pre-stimulus onset, 30 s post-stimulus onset) and baseline corrected based on the average value of the 1 s before stimulus onset. As for EEG, grand-average responses were obtained by averaging over all participants in each condition. Phasic EDA was standardized into z-scores - i.e. mean of zero, standard deviation of one - following (Ben-Shakhar, 1985). Running independent t-tests corrected for multiple comparison using FDR were conducted to test for significant between-group differences over time.

2.3. Results

2.3.1. Physiological synchrony as a measure of selective attention

Figure 2-2 shows the within-group and between-group ISC averaged across AA participants and SA participants in EEG, EDA and IBI. Within-group and between-group ISC of individual participants are plotted on top of the bars. Figure 2-2(a) shows ISC over the whole audiobook. Results for EEG are in line with our hypothesis. ISC are higher for participants when paired to participants from their own attentional group compared to participants from the other group. This is the case both for participants in the AA group ($t(12) = 4.72$, $p = 10^{-4}$) and SA group ($t(12) = 4.79$, $p = 10^{-4}$). EDA partly follows our hypothesis. Within-group PS is higher than between-group PS for SA participants ($t(12) = 4.07$, $p = .002$), but not for AA participants ($t(12) = 0.74$, $p = .476$). In IBI, no significant group-level effects were found: (AA: $t(12) = 2.17$, $p = .051$, SA: $W = 1.64$, $p = .110$). When assuming for each participant that she or he follows the attentional instruction as indicated by the group with whom she or he shows the highest averaged synchrony, classification accuracies are high and well above chance level, as shown in the first column of Table 2-1. Using this leave-one-participant-out paradigm, ISC in EEG correctly identifies the attentional condition of all but one of the participants. Using EDA and IBI, 73% of the participant are correctly identified. Figure 2-2(b)–(d) shows ISC averaged across AA participants and SA participants when paired with participants of the AA group or SA group during beep presentation (b), affective sound presentation (c) and when considering only audiobook parts without interspersed stimuli (d). The classification accuracies are shown in columns two to four of Table 2-1. During beep presentation, ISC-EEG are clearly higher for SA participants when paired with other SA participants than when paired with AA participants ($t(12) = 5.59$, $p = 10^{-4}$). AA participants do not synchronize more within-group than between-groups ($t(12) = 2.05$, $p = .062$). During affective sound presentation both groups have higher within-group PS than between-group PS (AA: $t(12) = 2.35$, $p = .037$; SA: $t(12) = 2.26$, $p = .043$). Overall, classification accuracy is lower rather than higher with respect to the whole audiobook, both during the beeps (88%) and especially during affective sounds (73%). When excluding experiment parts with interspersed stimulus presentation (audiobook only), AA participants clearly have stronger ISC with other participants attending to the narrative, than with participants not attending to the narrative ($t(12) = 5.05$, $p < .001$). For SA participants there is no significant difference between ISC with respect to both groups ($t(12) = 0.63$, $p = .541$). In EDA similar effects are found as in EEG. Figure 2-2 shows that again ISC during blocks of beeps are higher for SA participants when paired with participants in their own attentional group ($t(12) = 4.66$, $p < .001$), but this does not hold for AA participants ($t(12) = -1.52$, $p = .155$). During affective sound presentation both groups have higher within-group than between-group PS (AA: $t(12) = 2.40$, $p = .034$; SA: $t(12) = 2.34$, $p = .038$). Compared to the whole stimulus, classification accuracy drops (69%) during beeps, but remains constant (73%) during affective sounds. When considering audiobook parts without interspersed stimulus presenta-

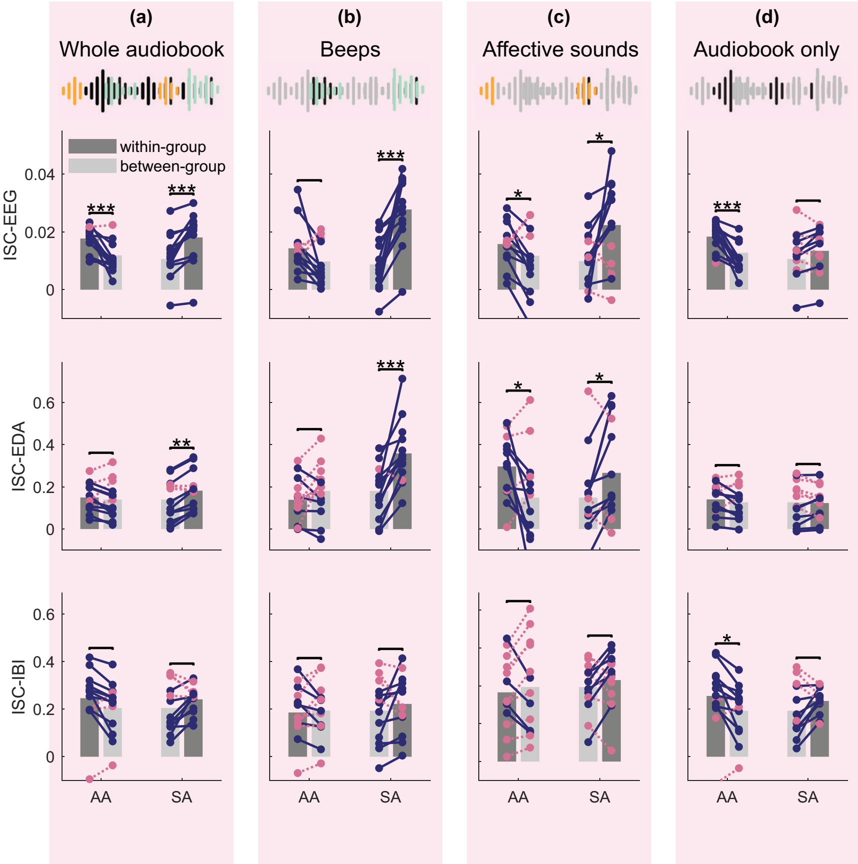

Figure 2-2. Within-group and between-group inter-subject correlations (ISC) for audiobook-attending participants (AA, left bars) and stimulus-attending participants (SA, right bars) for EEG, electrodermal activity (EDA) and inter-beat interval (IBI). a. shows ISC computed over the whole audiobook, b. when considering only parts with concurrent beep presentation, c. when considering only parts with concurrent affective sounds and d. when considering only audiobook parts without interspersed stimuli. Connected dots display participant-to-group ISC of each participant, where blue lines indicate participants for whom within-group ISC are higher than between-group ISC and pink dotted lines indicate individuals for whom between-group ISC are higher than within-group ISC. Paired sample t-test were conducted to test whether within-group correlations were higher than between-group correlations (* $p < .05$, ** $p < .01$, *** $p < .005$).

tion, no significant differences are found (AA: $t(12) = 1.15$, $p = .273$; SA: $t(12) = -0.34$, $p = .738$). Classification accuracy is not significantly higher than chance (62%) for narrative only. Results for IBI differ from the other measures. During beep presentation IBI ISC are not higher within-group than between-groups for both AA ($t(12) = -0.36$, $p = .725$) and SA groups ($t(12) = 0.09$, $p = .927$). Also during affective sound presentation there are no higher ISC within-group than between-groups (AA: $t(12) = -0.76$, $p = .460$; SA: $t(12) = 1.06$, $p = .310$). Table 2-1 shows that these re-

Table 2-1. The percentage of participants of which the attentional condition is correctly identified using inter-subject-correlations in EEG, electrodermal activity (EDA) and inter-beat interval (IBI) considering all four time intervals. In brackets the mean and standard deviation chance level classification performance is shown. Bold text depict classification accuracies significantly higher than this chance level distribution. p-values are shown in the table.

	Whole audiobook	Beeps	Affective sounds	Audiobook only
EEG	**96 (49 ± 11)** **p < .001**	**88 (52 ± 13)** **p < .001**	**73 (50 ± 13)** **p = .037**	**73 (50 ± 10)** **p = .010**
EDA	**73 (50 ± 10)** **p = .009**	**69 (50 ± 10)** **p = .032**	**73 (49 ± 10)** **p = .009**	62 (49 ± 11) p = .115
IBI	**73 (50 ± 11)** **p = .009**	58 (52 ± 9) p = .266	42 (49 ± 10) p = .742	**73 (50 ± 10)** **p = .009**

sults are reflected in classification accuracies. Classification accuracies are not higher than chance level for beeps (58%) and affective sounds (42%). For the audiobook parts without interspersed stimuli, PS is higher within-group than between groups for AA participants ($t(12) = 2.64$, $p = .022$), but not for SA participants ($W = 1.57$, $p = .116$). Classification accuracy is identical to performance considering the whole stimulus (73%).

2.3.2. Correlations between physiological synchrony and performance measures indicative of attentional focus

The results on the post-audio stimulus-content questionnaire confirmed that participants followed their attentional instructions. SA participants remembered more affective sounds (Mdn = 6) than AA participants (Mdn = 4) ($W = 2.68$, $p = .007$) and more accurately estimated the number of beeps in the counting task than AA participants, with significantly smaller estimation error for SA participants (Mdn = 1) than AA participants (Mdn = 10), ($W = 2.82$, $p = .005$). AA participants recalled the narrative of the audiobook more accurately. They answered more questions about the narrative correctly (Mdn = 6) than SA participants (Mdn = 3), ($Z = 2.68$, $p = .007$). Strong attentional focus, following the instruction to attend either to the narrative or to the short stimuli, can be expected to result in high performance on respectively the AA or the SA questionnaires and high ISC toward the AA and SA group. To investigate whether ISC were predictive of performance on the questionnaires, we computed correlations of the directional synchrony measures ISC-AA and ISC-SA with the questionnaire performance measures. Table 2-2 shows the correlation coefficients r and corresponding p values for the different combinations. In the grey cells we hypothesized posi-

Table 2-2. Correlation coefficients (r) and corresponding p-values between inter-subject correlations (ISC) with the audiobook-attending group or stimulus-attending group and the number of correctly answered narrative questions, the number of reproduced affective sounds, estimated average number of beeps, for EEG, electrodermal activity (EDA) and inter-beat interval (IBI). Additionally, correlations between the difference of ISC toward the Audiobook Attending group and ISC toward the Stimulus Attending group with the difference of the performance metrics are shown. Bold text depicts significant correlations. Italics depicts hypothesized positive correlations.

		ISC toward Audiobook Attending group	ISC toward Stimulus Attending group
Ranked performance of number of correctly answered narrative questions	EEG	***r = .50, p = .010***	r = -.01, p = .978
	EDA	*r = -.16, p = .462*	r = -.28, p = .168
	IBI	***r = .59, p = .002***	r = .08, p = .697
Ranked performance of number of reproduced affective sounds	EEG	r = -.02, p = .912	***r = .61, p = .001***
	EDA	r = .00, p = .985	*r = .16, p = .453*
	IBI	r = -.09, p = .686	*r = .08, p = .715*
Ranked performance of estimated average number of beeps	EEG	r = -.11, p = .615	*r = .38, p = .071*
	EDA	r = -.19, p = .387	*r = -.06, p = .770*
	IBI	r = -.04, p = .858	*r = .18, p = .402*
		Difference between ISC toward Audiobook Attending group and ISC toward Stimulus Attending group	
Ranked performance difference	EEG	***r = .65, p = .001***	
	EDA	*r = .34, p = .128*	
	IBI	***r = .61, p = .003***	

tive correlations: attending to short stimuli would result in both high ISC with respect to the SA group and high performance on the questions about the affective sounds and beeps; attending to the narrative would result in both high ISC with respect to the AA group and high performance on the questions about the narrative. The significant correlations are shown in bold. For EEG, results are in line with our hypothesis. ISC with respect to the SA group strongly correlates with the number of reproduced affective sounds, whereas ISC with respect to the AA group strongly correlates with the number of correctly answered narrative questions. In IBI, ISC with respect to the AA group is also significantly correlated with the number of correctly answered narrative questions. Correlations were not significant for the other combinations in the grey cells. However, correlations tend to be positive in the grey cells, and negative in the other cells as

Table 2-3. Correlations coefficients (r) and corresponding p-values between self-reported measures of distraction and mental effort with performance metrics reflective of paid attention. In all cells with text we hypothesized correlations. In the case of self-reported distraction, these were expected to be negative, in the case of mental effort these correlations could be either negative or positive. Cells with significant correlations are depicted in bold.

	Audiobook Attending	**Stimulus Attending**	
	Ranked performance of number of correctly answered narrative questions	**Ranked performance of number of reproduced affective sounds**	**Ranked performance of number of estimate average number of beeps**
Distraction by the other stream of audio	r = -.22, p = .465	r = -.20, p = .529	r = -.11, p = .738
Mental effort during the experiment	**r = -.56, p = .045**	r = -.18, p = .586	r = .25, p = .465
Distraction by blocks of beeps	**r = -.66, p = .015**	x	x
Distraction by affective sounds	r = .00, p = .999	x	x
Distraction by the audiobook	x	r = -.21, p = .512	r = -.12, p = .735
Mental effort during beep counting	x	x	r = .25, p = .461

might be expected when attending to the one aspect (narrative or short stimuli) decreases performance on questions about the other aspect (short stimuli or narrative). Some participants might be able to attend to all of the stimulus content while others might not even be able to attend to their own stimulus content. We tested whether the difference between the directional synchrony measures was also predictive of the difference between the performance metrics. Table 2-2 therefore also shows correlations between ISC-SA minus ISC-AA with the difference in performance on SA and AA questionnaires. Correlations are in line with our hypothesis, with strong positive significant correlations for EEG and IBI.

Table 2-3 shows that none of the self-reported measures of distraction and mental effort predicted performance on post-stimulus questions for SA participants. For AA participants, reported overall mental effort predicted the number of correctly answered narrative questions ($r = -0.56$, $p = .045$), where a high reported mental effort was associated with low performance. Reported distraction by the beep blocks correlated negatively with performance ($r = -0.66$, $p = .015$), while no significant correlations were found between self-reported distraction and distraction by the affective sounds with performance on the narrative questions.

2.3.3. Stimulus-locked response traces

In the first section of the Results, the effect of interspersed stimulus presentation on ISC was presented. In this section we further focus on epochs with interspersed stimuli to investigate their effect on the physiological responses. Figure 2-3 shows midline parietal (Pz) event-related potentials, time-locked to interspersed stimulus onset (respectively beeps and affective sounds). Independent-sample running t-tests corrected for multiple comparisons revealed significant between-group differences ($q < .05$) in response toward the beeps, with larger deflections in SA participants than AA participants. In response to affective sounds, no between-group differences in responses were found. Figure 2-3 also shows response traces for standardized phasic EDA and IBI. Although for EDA, on average responses of SA participants seem to show larger deflections than those of AA participants, statistical tests do not reveal significant between group differences in response to any of the stimuli. For IBI, response traces are very similar and no significant between-group differences were found.

2.4. Discussion

2.4.1. Summary of findings

In the current study we determined physiological synchrony (PS) through inter-subject correlations (ISC) in EEG, EDA and IBI to determine the selective attentional focus of individuals who were all presented with the same auditory stimulus and were all attending to it, but were attending to different stimulus aspects. PS in all three modalities was associated with selective attention. EEG and EDA responses of participants were more synchronized with those of participants sharing attentional focus than with those of participants attending to other stimulus aspects, but for IBI no significant effects were found. Using the correlations of an individual's EEG with the two groups of differently attending individuals as a predictor of attentional instruction resulted in a classification accuracy of 96%. For EDA and IBI, accuracies of 73% were reached. All of the classification accuracies are well above chance level. Even when only data was included coming from 'audiobook only' intervals, classification performance was above chance level for EEG and IBI, although not for EDA. The level of synchrony toward the groups also correlated with post-stimulus performance metrics reflective of paid attention, reinforcing the validity of PS as measure of attention and suggesting PS as a suitable predictor of performance. The results are framed in terms of a broader picture in the following sections.

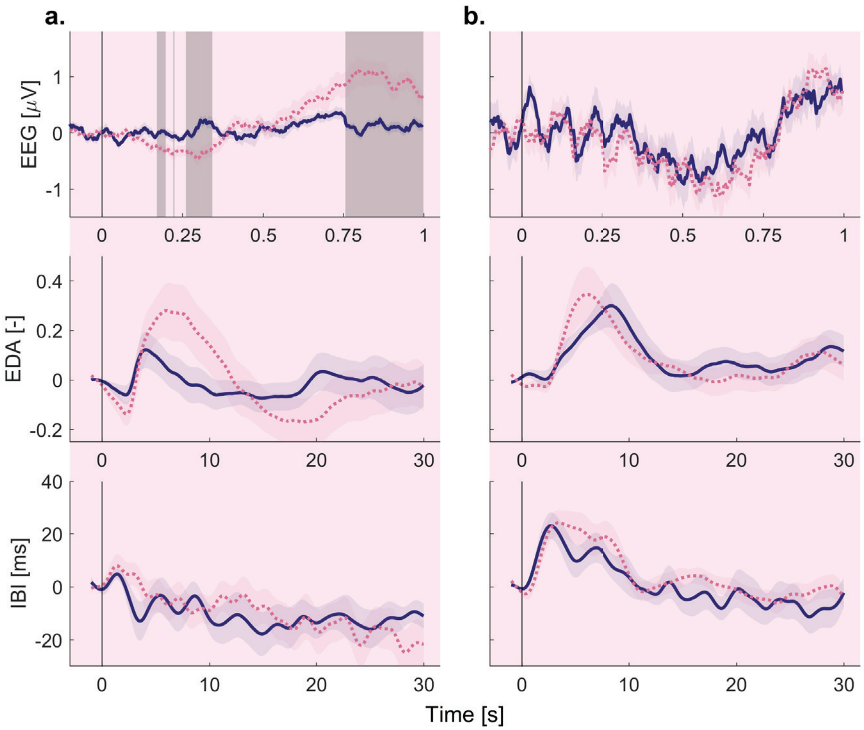

Figure 2-3. Midline parietal event-related EEG potentials, standardized phasic electrodermal responses (EDA) and inter-beat interval (IBI) time-locked to stimulus onset of a. the beeps in each block and b. affective sounds, averaged over Audiobook-Attending participants (blue, solid line) and Stimulus-Attending participants (pink, dotted line). The standard error of the mean across participants in each group is depicted in shaded areas around the grand average potentials. Significant between-group differences ($q < .05$, corrected for multiple comparisons using false discovery rate) are depicted with gray areas in the potential plots.

2.4.2. Physiological synchrony as measure of selective attention

This is not the first study associating PS in EEG with attentional engagement to naturalistic stimuli, but our study differs from previous studies in several important aspects. Rather than relating PS in EEG to shared attentional engagement toward a single stream of information or distinguishing between attentive and inattentive conditions (Cohen et al., 2018; Dmochowski et al., 2014, 2012; Ki et al., 2016), we here show that we can also distinguish between two different selective outward auditory attentional conditions with 96% accuracy.

EDA and IBI performed quite well in distinguishing between groups too. To our best knowledge, this is the first time that PS in EDA or IBI was shown to be modulated by attentional focus only. The promising performance of these measures is important from a user perspective, as EDA and IBI can be more easily monitored in ambulatory environments than EEG. We must note that effects for IBI are not as strong as for EEG and EDA. Classification accuracies for the whole stimulus and audiobook-only-parts were at least as high or higher for IBI than for EDA. However, when considering the whole stimulus, ISC values in IBI were not significantly different for between- and within-attentional groups. As can be seen in the IBI panel in Figure 2-2(a), IBI ISC is higher within than between attentional groups for the same number of participants as in EDA, presented in the panel above it. However, because the size of this difference in IBI is relatively variable across participants, the statistical test did not produce a significant effect for IBI while it did for EDA. This may be explained by the fact that the relation between IBI and mental processing seems less straightforward, and more person dependent than for EDA. Whereas EDA has consistently shown a positive relation with arousal (Boucsein, 2012), IBI shows a more complex relation with arousal, as both heart rate accelerations (e.g. Brouwer and Hogervorst, 2014) and heart rate decelerations (e.g. Brouwer et al., 2015a) have been reported. The reason for this is probably that arousal can be associated with the body being prepared for action (the defense reflex) or with a concentrated, focused state (the orienting reflex), that have been associated with heart rate accelerations and decelerations, respectively (Graham and Clifton, 1966). As increased physiological arousal has been associated with increased emotional and attentional engagement (Boucsein, 2012; Critchley, 2002), a more complex relation with arousal may result in attenuated PS in IBI for some participants.

2.4.3. Influence of interspersed stimuli on the identification of selective attention

We hypothesized that high classification performance would be driven by moments in the audiobook with concurrent stimulus presentation. We expected that in the large parts of the experiment where only the audiobook was played it was probably hard for SA participants to ignore the narrative. This would result in similar physiological activation across all participants. Our results suggest otherwise. When considering parts of the audiobook where no stimuli were interspersed, classification accuracies were still above chance level for EEG and IBI, although not for EDA. EEG and IBI of AA participants were also found to synchronize significantly more with other AA participants than with SA participants, revealing the difference in shared attentional focus between participant groups, also during audiobook only. This result does not mean that PS was not influenced by the interspersed stimuli, although we did not find our hypothesized effect that classification accuracies would be higher when considering only data with concurrent stimulus presentation. This may partly be due to the specific chosen interspersed stimuli durations and presentation frequency. Nonetheless, during presentation of beeps and affective sounds, EEG and EDA of SA participants were much more strongly synchronized with the signals of other SA participants than with those of AA participants. Figure 2-2 suggests that this effect is more pronounced than when considering the entire experiment.

PS results were different for different modalities. Because mental workload mainly affects EEG, and because it is expected to respond in a similar way across participants to a well-timed, attended stimuli (Hogervorst et al., 2014), we hypothesized the group distinguishing capability of ISC in EEG to work well during the beep counting task. As emotional stimuli have been strongly related to sympathetic nervous system activity as measured through EDA (Boucsein, 2012; Bradley and Lang, 2000), we hypothesized the group-distinguishing capability of ISC in EDA to work well during the affective sounds. Indeed, we found no strong drop in classification performance for EEG during the beeps as compared to the whole stimulus and no strong drop for EDA during the affective sounds. In addition to this, ISC in IBI identified the selective attention relatively well during audiobook parts without interspersed stimuli. These findings support a multimodal approach that can exploit the particular strength of each neural and peripheral measure. Also note that that the attentional condition of all participants was correctly classified by at least one of the three physiological measures (see the identification of the selective attention for each participant and each physiological measure in Table 5 in the supplementary material).

PS in different modalities are not only expected to differ in reflecting selective attention because they are associated with different types of mental activity, but also because they unfold on different timescales. Whereas EEG unfolds in the range of milliseconds, response latencies of the peripheral physiological measures are two orders of magnitude larger. Especially when interested in fusion data from all three sensors into a single index of multimodal PS, the issue of timescales has to be resolved in future work.

2.4.4. Behavioral performance and its association with physiological synchrony

We hypothesized that more synchronized physiological responses with respect to an attentional group would lead to better performance on the accompanying post-stimulus questionnaires. Participants with high PS elicit physiological activity that is similar to that of their peers and they are therefore thought to be more engaged with the stimulus (Cohen et al., 2017; Dmochowski et al., 2014). For EEG, this has indeed been found to result in correlations with performance on immediate and delayed memory retention questions (Cohen et al., 2018; Cohen and Parra, 2016). Following our hypothesis, ISC in EEG strongly correlated with performance on questionnaires reflective of paid attention. The degree of synchrony with respect to AA participants predicted performance on questions about the narrative, whereas the degree of synchrony with respect to SA participants predicted performance on questions about the short stimuli. Also in IBI the degree of synchrony with the SA group predicted short-stimulus retention. For both EEG and IBI, we also found that the degree to which participants synchronize more with one of the attentional groups significantly correlates with the degree to which they score better on that groups' retention questions than on questions reflective of the other group's content. This is important when monitoring selective attentional engagement. Rather than only being able to distinguish overall attentive individuals (generally high PS toward both attentional groups) from overall inattentive individuals (generally low PS toward both attentional groups), this finding enables the identification of well-focused individuals, that attend well to specific information, while shutting-off other information (high PS toward one attentional group, low PS toward the other attentional group). These differences are found to be meaningful in terms of performance. Simply asking participants how distracted they were by other stimulus aspects or how much mental effort they invested during the experiment was not as informative of performance on post-stimulus questions as measures of PS. Performance of AA participants on the questions about the narrative of the audiobook was predicted by the degree of invested mental effort and the degree of distraction by blocks of beeps. However, these results were inconsistent with other self-reported measures - performance on narrative questions was not predicted by the degree of distraction by all interspersed stimuli. Furthermore, for SA participants none of the self-reported metrics of mental state predicted performance resulting in an incomplete view.

2.4.5. Interspersed stimulus response traces

To obtain an understanding of what drives the found effects of attentional instruction on PS, we locked the physiological response traces locked to the onset of the interspersed stimuli. We hypothesized larger deflections for SA participants than for AA participants. For the blocks of beeps, this hypothesis was confirmed; event-related responses in parietal EEG in response to beeps were significantly more deflected for SA participants than AA participants and responses of phasic EDA show a similar although non-significant effect. However, responses to affective sounds were indistinguishable between groups, with deflections for both attentional groups. Our stimuli, beeps and affective sounds, differed with respect to their capacity to draw attention. The blocks of beeps mainly attract attention through top-down mechanisms related to task instructions, whereas the affective sounds also attract attention through bottom-up mechanisms related to salience or emotional relevance (Öhman et al., 2001). The affective sounds could thus be expected to attract attention of all participants and therefore to induce responses in physiological measures of all participants. This may have resulted in responses to affective sounds that were indistinguishable between groups.

2.4.6. Processes underlying physiological synchrony

Our findings, together with those of others who found PS in electroencephalographic and hemodynamic cortical responses as a function of attentional instruction, suggest that neural correlates of cognitive processes are reliable and reproducible (Hasson et al., 2004, 2010; Furman et al., 2007; Jääskeläinen et al., 2008; Wilson et al., 2008). It is not yet clear which underlying processes are reflected in cortical synchronization. The inter-subject synchronization has been associated with a broad range of higher-level processes, such as memory encoding, emotional processing and stimulus preference (Furman et al., 2007; Hasson et al., 2008; Jääskeläinen et al., 2008; Wilson et al., 2008; Dumas et al., 2010; Nummenmaa et al., 2012; Dmochowski et al., 2014; Ki et al., 2016). The similarity of scalp topographies of the cortical correlated components across sensory modalities indicates that the fundamental processes underlying cortical ISC are low-level and supramodal (Cohen and Parra, 2016; Ki et al., 2016). Our findings of synchrony in peripheral measures suggest that both systems are to some degree influenced by the same high-level processes. Research has shown that sympathetic autonomic activity is indeed influenced by higher subcortical and cortical brain areas implicated in high-level processes of attention, emotion and motivation (Kaada, 1951; Neafsey, 1991). Some of these brain areas were found to covariate with synchronization in EEG of participants sharing attention to a narrative visual stimulus (Dmochowski et al., 2014). It may be the case that activation of the autonomic measures is induced through mechanisms of arousal, as increased attention has been shown to be associated with heightened arousal (Critchley, 2002). However, future research is needed to unravel the underlying mechanisms of PS in cortical and autonomic measures.

Nonetheless, determining how strongly physiological measures synchronize across individuals is a valuable way to monitor attentional or emotional engagement. The simplicity of the current analysis may make this a valuable approach compared to other ways to determine emotional or attentional engagement using physiological variables. A common approach for the assessment of attention or engagement in this field is based on supervised learning algorithms, where a machine learning model is trained to predict attentional engagement (Liu et al., 2013; Aliakbaryhosseinabadi et al., 2017) or emotional engagement (Bailenson et al., 2008) from a feature set of physiological variables. These approaches require labeled training data, i.e. a set of physiological responses that are labeled with the degree of attentional or emotional engagement. Not only is this time-consuming, it is also very difficult to determine a 'ground truth' mental state than can be used for data labeling (Brouwer et al., 2015b). Determining the degree of PS does not depend on labeled training data. This is especially valuable when there is limited information about events in the world, as is the case in real-world environments like classrooms, where it is difficult to obtain a set of labeled training data.

2.4.7. Future work

While the current study and analyses produced interesting findings, there are a number of topics we have in mind in order to improve and add to our current results. Firstly, we will investigate ways of combining PS in the three modalities into one multimodal measure of PS.

Furthermore, in the current work PS in EDA and IBI was computed using simple Pearson correlations in moving windows. While this method is computationally inexpensive and easily adaptable for online use, limitations of the method include oversampling as a result of overlapping windows as well as potentially spurious correlations as a result of not controlling for autocorrelation (Levenson and Gottman, 1983). While such correlations would not explain the difference between selective attentional conditions, they could influence overall correlation levels. Future research could investigate whether other methods of synchrony assessment would result in similar findings. Synchrony assessment would not even have to be limited to the time domain, but could also include frequency domain metrics, such as wavelet coherence in IBI (Quer et al., 2016) or one of the many coherence metrics in neural measures (Babiloni and Astolfi, 2014). An innovative method of synchrony assessment for ECG was presented by (Verdiere et al., 2020) who analyzed concurrent ECG peaks and found this to be a relatively effective method to detect concurrent, high workload in teams. Future work could also compare the currently obtained results with other methods of attention monitoring. For example, (Ki et al., 2016) showed that not only EEG ISC but also alpha power could distinguish naturally attending participants form inward focused participants, be it with a weaker modulation.

Finally, we want to suggest future work to focus on more unsupervised mechanisms identifying groups with different attentional focus. Unsupervised clustering techniques may be applied to this dataset. As we encourage other researchers to test other synchrony metrics or classification paradigms, the MATLAB scripts and physiological data reproducing the results in this study are publicly available on https://github.com/ivostuldreher/physiologicalsynchrony-selective-attention.

2.5. Conclusion

In this study we monitored EEG, EDA and IBI responses and assessed physiological synchrony within and between groups, that were either instructed to focus attention on an audiobook or on interspersed auditory stimuli. We showed that PS in neural and autonomic measures reflects selective attentional engagement. Out of the complete set of measures, EEG showed the best results, with strong group-level differences and correct identification of the selective attentional focus in 96% of the cases. PS in EDA and IBI also showed good results, with significant group level differences in EDA and classification accuracies of 73%. Even when only data was included coming from 'audiobook only' time intervals, classification performance was above chance level for EEG and IBI, though not for EDA. The level of synchrony toward the groups also predicted performance on post-stimulus questions reflective of paid attention. Our results support that synchrony in physiological responses with others reflects selective attentional engagement. To our best knowledge this is the first time PS has been monitored in neural and autonomic measures concurrently. The relatively high classification accuracies with the use of PS in EDA and IBI are convenient from a user perspective and should enable researchers to monitor PS in autonomic measures in situations where intrusive neural measurements are not suited. However, as each modality performed relatively good in specific stimulus conditions, we also have the ambition to combine the physiological measures into a multimodal index of PS. Work in this area may lead to applications for evaluating educational material or provide feedback to educators or other types of presenters in real time.

Acknowledgments

We thank Ana Borovac for all the help while performing the experiment. This work was supported by The Netherlands Organisation for Scientific Research (NWA Startimpuls 400.17.602).

2

3.

Physiological synchrony in EEG, electrodermal activity and heart rate detects attentionally relevant events in time

Stuldreher, I.V., Thammasan, N., van Erp, J.B.F., Brouwer, A.M. (2020). Physiological synchrony in EEG, electrodermal activity and heart rate detects attentionally relevant events in time. *Frontiers in Neuroscience*, ***14***, 575521. doi: 10.3389/1741-2552/aba87d

Abstract

Interpersonal physiological synchrony (PS), or the similarity of physiological signals between individuals over time, may be used to detect attentionally engaging moments in time. We here investigated whether PS in the electroencephalogram (EEG), electrodermal activity (EDA), heart rate and a multimodal metric signals the occurrence of attentionally relevant events in time in two groups of participants. Both groups were presented with the same auditory stimulus, but were instructed to attend either to the narrative of an audiobook (audiobook-attending: AA group) or to interspersed emotional sounds and beeps (stimulus-attending: SA group). We hypothesized that emotional sounds could be detected in both groups as they are expected to draw attention involuntarily, in a bottom-up fashion. Indeed, we found this to be the case for PS in EDA or the multimodal metric. Beeps, that are expected to be only relevant due to specific "top-down" attentional instructions, could indeed only be detected using PS among SA participants, for EDA, EEG and the multimodal metric. We further hypothesized that moments in the audiobook accompanied by high PS in either EEG, EDA, heart rate or the multimodal metric for AA participants would be rated as more engaging by an independent group of participants compared to moments corresponding to low PS. This hypothesis was not supported. Our results show that PS can support the detection of attentionally engaging events over time. Currently, the relation between PS and engagement is only established for well-defined, interspersed stimuli, whereas the relation between PS and a more abstract self-reported metric of engagement over time has not been established. As the relation between PS and engagement is dependent on event type and physiological measure, we suggest to choose a measure matching with the stimulus of interest. When the stimulus type is unknown, a multimodal metric is most robust.

3.1. Introduction

Knowing what events in the external environment people attend to, and how their shared attentional engagement to events varies over time, can be useful in a range of settings, from evaluating educational or entertaining material, to real time adjustment of important instructions. Unlike explicit measures, such as questionnaires in which people are asked to specify their attentional engagement, physiological signals can provide continuous and implicit information on mental state (Zander and Kothe, 2011). However, the link between mental state and physiological measures [e.g., electroencephalography (EEG), electrodermal activity (EDA) or heart rate] is not straightforward (Brouwer et al., 2015b). A popular approach to uncover the complex links between physiology and mental state is the use of supervised learning algorithms. These algorithms predict mental state based on a set of features extracted from physiological variables (Hamadicharef et al., 2009; Hussain et al., 2011; Fleureau et al., 2012; Liu et al., 2013; Aliakbaryhosseinabadi et al., 2017). A disadvantage of these types of analyses is the need for labeled training data, i.e., a set of physiological data that are labeled with a known value for the mental state of interest. Not only is it time consuming to obtain such a labeled dataset, it is also very difficult to determine the 'ground truth' mental state than can be used for data labeling (Brouwer et al., 2015b). A second drawback of these supervised learning approaches is that classification is often limited to a small number of discrete states. Attentional, emotional or cognitive state, however, cannot realistically be represented by a small number of discrete states, but are naturally of more continuous nature (Zehetleitner et al., 2012; Rosenberg et al., 2013).

For monitoring attentional engagement, an approach that may be suited to circumvent both of the abovementioned problems is to monitor the physiological synchrony (PS) between individuals. PS is the degree to which physiological measures of multiple people uniformly change. Studies exploring PS in functional magnetic resonance imaging data have revealed strong voxel-wise inter-subject correlations across participants exposed to a common narrative stimulus (Hasson et al., 2004, 2010; Hanson et al., 2009). In the faster EEG signals, similar results were found (Dmochowski et al., 2012, 2014). The fast-changing EEG enabled the computation of a continuous measure of PS in time and suggested that moments of high PS corresponded with emotionally arousing scenes of the movie clips (Poulsen et al., 2017). For instance, high PS was found when scenes were viewed that involved the threat of a gun. Dmochowski et al. (2014) further showed that moment-to-moment variation in the PS predicted the general expressions of interest and attention of the public as indicated by number of tweets during a popular television series. Davidesco et al. (2019) found that PS over time indicated what specific information was retained by students in a lecture. Namely, PS was higher in lecture parts that provided answers for questions that students answered incorrectly in the pre-test and correctly in the delayed post-test than for questions where students' answers did not change. The relationship between neural PS and attentional engagement was also found to be less complex than most traditional physiological metrics. Neural PS was found to be directly proportional to attentional engagement, as strong correlations were found between PS and performance on questionnaires reflective of paid attention (Cohen and Parra, 2016; Cohen et al., 2018; Stuldreher et al., 2020b). This directly proportional relationship may thus be used to circumvent supervised learning approaches and the problems that come with such approaches, such as the dependency on labeled training data.

In the current work, we aim to employ the relation between PS and attentional engagement to detect the occurrence of attentionally relevant events in time. Rather than limiting the analyses to EEG, we also include PS measures of peripheral nervous system activity (EDA and heart rate), and quantify their comparative sensitivity of detecting relevant events. Up to recently, PS in peripheral physiological measures has been studied mainly as a metric of some form of affective connectedness between individuals (reviewed by Palumbo et al., 2017). Examples include peripheral PS in therapist-patient dyads as a measure of psychotherapy success (Koole et al., 2020), in couples in marital counseling as a measure of therapy outcome (Tourunen et al., 2020) and as measure of collaborative learning (Malmberg et al., 2019). Positive results found in these contexts may (partly) be driven by shared attentional engagement to external events, as connectedness between people may be strongly associated with mutual attentiveness (Tickle-Degnen and Rosenthal, 1990). Recently, it was found that PS in EDA and heart rate can indeed reflect shared attention toward narrative stimuli (Stuldreher et al., 2020b; Pérez et al., 2021).

The advantage of peripheral physiological measures over EEG is that they can be recorded more easily and less obtrusively. In addition, EEG and peripheral measures may complement each other since they likely reflect different mental processes. EEG is, for example, sensitive to selective attention (Polich, 2007), whereas EDA and heart rate are sensitive to (emotional) arousal (Cacioppo et al., 2000; Boucsein, 2012).

As of yet, it is unknown whether PS in EEG, EDA and heart rate can be used to detect relevant moments in time. For EDA and heart rate, time-resolved dynamics of PS have not been investigated at all in the context of attentional engagement. For EEG, time-resolved dynamics have been explored (see for instance Dmochowski et al., 2014; Poulsen et al., 2017), but this has not been done systematically, using a-priori known cognitively or emotionally engaging stimuli for which detection performance can be evaluated. We here evaluate whether PS in EEG, EDA and heart rate can be used to detect cognitively or emotionally relevant moments in time. Our goal is not to compare detection performance directly between the different types of stimuli, but to evaluate PS for a range of events differing in terms of total duration, sound onset, mental processes addressed and more. Just as in real-world conditions, some event may capture attention in a bottom-up fashion, related to salience or emotional relevance, whereas others may only capture attention due to top-down mechanisms related to task instruction (Lang, 1995; Öhman et al., 2001; Schupp et al., 2003). We invited participants to come to our lab and listen to an audiobook that was interspersed with short auditory events, that we expected to induce emotional and cognitive load. We divided the participants in two equal-sized groups. Participants in the audiobook-attending group (AA) were instructed to focus their attention on the audiobook and ignore the interspersed stimuli. Participants in the stimulus-attending group (SA) were instructed to focus their attention on the interspersed stimuli and ignore the audiobook. In a previous paper on this experiment (Stuldreher et al., 2020b), we showed that PS can be used to correctly classify a listener as being instructed to attend to the audiobook or to the sounds. In the current paper, we use PS among individuals in the same group to predict the occurrence of interspersed stimuli over time, for each of the three physiological measures. In addition, we investigated if the PS across AA participants was predictive of the occurrence of engaging moments in the book. We aimed to answer the following research questions:

Does PS in EEG and EDA, heart rate and a multimodal metric predict the occurrence of attentionally engaging moments in time? And does this depend on the attentional instruction, type of stimuli and physiological measure?

We expect that interspersed stimulus detection performance of PS measures depends on combinations of the attentional group (AA or SA), the interspersed stimulus type (emotional sounds or beeps) and the physiological measure (EEG and EDA, heart rate or the multimodal metric). We hypothesized the following;

(1) Attentional instruction and stimulus type: (a) for the SA group, detection performance based on PS is above chance for all interspersed stimuli. (b) For the AA group, detection performance based on PS is above chance for emotional sounds, since these attract attention through bottom-up mechanisms related to salience or emotional relevance (Lang, 1995; Öhman et al., 2001; Schupp et al., 2003) irrespective of task instruction. (c) For the AA group, detection performance based on PS is not above chance for beeps, as these are expected to mainly attract attention through top-down mechanisms related to task-instructions.

(2) Physiological measure and stimulus type (a) PS based on peripheral signals (EDA and heart rate) performs better on the detection of emotional sounds than on beeps, because they primarily reflect emotional state (Cacioppo et al., 2000; Boucsein, 2012). (b) PS based on EEG performs better on the detection of beeps than on the detection of emotional sounds, because they primarily reflect top-down selective attention or mental effort (Hogervorst et al., 2014).

(3) Combining physiological measures: combining the physiological measures into a single multimodal metric of PS would result in relatively high detection accuracies when disregarding the differences between stimulus types.

While for the SA group, the timing of short stimuli serve as “ground truth” relevant events to compare to the moments of high PS, we do not know a priori what constitutes relevant events or engaging moments in the audiobook. We therefore investigate ratings of post-hoc determined moments of high and low PS in the audiobook by an independent group of participants. We hypothesized that;

(4) Events in audiobook: moments of the audiobook that were associated with high PS in the AA group are rated as more engaging than moments of the audiobook that were associated with low PS.

3.2. Materials and methods

3.2.1. Participants

Twenty-seven participants (17 female), between 18 and 48 years old, with an average of 31.6 years and a standard deviation of 9.8 years, were recruited through the institute’s participant pool. Before performing the study, approval was obtained from the TNO Institutional Review Board (IRB). The approval is registered under the reference 2018–70. Prior to the experiment all participants signed informed consent, in accordance with the Declaration of Helsinki. After signing, all participants were randomly assigned to either the AA group or the SA group. After the experiment they received a small monetary reward for their time and traveling costs. None of the participants indicated problems in hearing or attention. Data of one participant were discarded due to failed physiological recordings, resulting in two equal-sized groups.

3.2.2. Materials

EEG, EDA, and electrocardiogram (ECG) were recorded at 1024 Hz using an ActiveTwo Mk II system (BioSemi, Amsterdam, Netherlands). EEG was recorded with 32 active Ag-AgCl electrodes, placed on the scalp according to the 10–20 system, together with a common mode sense active electrode and driven right leg passive electrode for referencing. The electrode impedance threshold was maintained below 20 kOhm. For EDA, two passive gelled Nihon Kohden electrodes were placed on the ventral side of the distal phalanges of the middle and index finger. For ECG, two active gelled Ag-AgCl electrodes were placed at the right clavicle and lowest floating left rib. EDA and heart rate were also recorded using wearable systems (Movisens EdaMove 4 and Wahoo Tickr, respectively). These data are discussed elsewhere (Borovac et al., 2020; Van Beers et al., 2020).

3.2.3. Stimuli and design

Participants performed the experiment one by one. Each participant was presented with the exact same audio file, composed of a 66 min audiobook (a Dutch thriller "Zure koekjes," written by Corine Hartman) interspersed with other short, auditory stimuli. Half of the participants were asked to focus on the narrative of the audiobook and ignore all other stimuli or instructions (AA group); and half of the participants were asked to focus on the short, interspersed stimuli and perform accompanying tasks, and ignore the narrative (SA group). The auditory stimuli were 36 emotional sounds, 27 blocks of beeps that SA participants had to keep track of, and an auditory instruction to sing a song. The order of sounds and beeps was randomly determined but was identical for each participant. Inter-stimulus intervals varied between 35 and 55 s, with an average of 45 s and a standard deviation of 6.1 s. We selected these stimuli to evaluate PS for a broad range of events, differing in e.g., audio profile and expected effect on mental processes as a function of task instructions.

Emotional sounds were taken from the second version of the International Affective Digitized Sounds (IADS) (Bradley and Lang, 2007a). The IADS is a collection of 6-s acoustic stimuli that have been normatively rated for valence (positive or negative affect), arousal and dominance. Examples of stimuli are the sound of a crying baby or a cheering sports crowd. We selected 12 neutral sounds (IADS number 246, 262, 373, 376, 382, 627, 698, 700, 708, 720, 723, 728), 12 pleasant sounds (110, 200, 201, 202, 311, 352, 353, 365, 366, 367, 415, 717) and 12 unpleasant sounds (115, 255, 260, 276, 277, 278, 279, 285, 286, 290, 292, 422) based on their normative ratings of valence and arousal. We expected these sounds to attract attention of all participants, even those instructed to ignore the interspersed sounds.

Beeps were presented in blocks of 30 s, with every 2 s a 100 ms high (1 kHz) or low (250 Hz) pitched beep. SA participants needed to separately count the number of high and low beeps presented in a block, as in (De Dieuleveult et al., 2018). This task was practiced with them beforehand. In total, 27 blocks of beeps were presented. We expected these sounds to only attract attention of participants clearly instructed to keep track of them.

Toward the end of the audiobook, the instruction was presented to sing a song aloud after a subsequent auditory countdown reached 0. This instruction had to be followed by the SA group and was expected to induce stress and a strong increase in EDA and heart rate (Brouwer and Hogervorst, 2014). For the current analyses, data following the onset of this stimulus were discarded, because some participants started singing before the counter reached 0. This prohibited analysis of the data in terms of mental processes due to confounding movement effects and artifacts in the data recording.

In total, we consider 3,800 s of data in further analyses, out of which 1,026 s involved concurrent presentation of the audiobook and interspersed stimuli.

3.2.4. Analysis

3.2.4.1. Pre-processing

Data processing was done using MATLAB 2019a software (Mathworks, Natick, MA, United States). For EEG pre-processing we also used EEGLAB v14.1.2 for MATLAB (Delorme and Makeig, 2004). To remove potentials not reflecting sources of neural activity, but ocular or muscle-related artifacts, logistic infomax independent component analysis (ICA) (Bell and Sejnowski, 1995) was performed. EEG was first down sampled to 256 Hz and high-pass filtered at 1 Hz. This relatively high cut-off frequency has shown to work better for ICA compared to lower cut-off frequencies (Winkler et al., 2015). Data were then notch filtered at 50 Hz, using the standard FIR-filter implemented in EEGLAB function pop_eegfiltnew. ICA was performed and the Multiple Artifact Rejection Algorithm (MARA) (Winkler et al., 2011) was used to identify artifactual independent components, i.e., components not reflecting sources of neural activity, but ocular or muscle-related artifacts. These components were removed from re-referenced, but uncleaned data. In these data, samples whose squared amplitude magnitude exceeded the mean-squared amplitude of that channel by more than four standard deviations were marked as missing data ("NaN") in an iterative way with four repetitions. By doing so, 0.82 % of data were marked as missing.

EDA was downsampled to 32 Hz. The fast changing phasic and slowly varying tonic components of the signal were extracted using Continuous Decomposition Analysis as implemented in the Ledalab toolbox for MATLAB (Benedek and Kaernbach, 2010). In the further analyses we use the phasic component of the signal as this component of the EDA signal is mainly related to responses to external stimuli.

ECG measurements were processed to acquire the inter-beat interval (IBI – inversely proportional to heart rate). After downsampling to 256 Hz, ECG was high-pass filtered at 0.5 Hz. Peaks were detected following Pan and Tompkins

(1985). The IBI semi-time series was transformed into a timeseries by interpolating consecutive intervals and resampling at 32 Hz.

3.2.4.2. Computation of inter-subject correlations as measure of physiological synchrony

We computed PS by measuring the inter-subject correlations of the neurophysiological signals. For EEG, rather than treating the signals from the 32 channels separately, we evaluated the inter-subject correlations in the correlated components of the EEG (Dmochowski et al., 2012, 2014). The goal of the correlated component analysis is to find underlying neural sources that are maximally correlated between participants, based on linear combinations of electrodes. Components were extracted separately from the AA group and SA group. EEG data from each participant were projected on the component vectors. Participant-to-group inter-subject correlations were then computed as the sum of correlations in the first three component projections, following (Dmochowski et al., 2012, 2014; Cohen and Parra, 2016; Ki et al., 2016; Cohen et al., 2018). Even though we used fewer participants in each attentional group than earlier work on auditory PS (e.g., Cohen and Parra, 2016; Ki et al., 2016), scalp projections of the components were very similar to those obtained in these earlier works, and our EEG PS values were in a similar range of 0.01 to 0.04. For the computation of time-resolved inter-subject correlations, correlations were computed in running 5 s windows at 1 s increments.

Inter-subject correlations in EDA and IBI were computed following (Marci et al., 2007). We computed Pearson correlations over successive, running 15 s windows at 1 s increments as measure of time-resolved inter-subject correlations. Participant-to-group correlations were computed by averaging over all correlations with all other participants in a group.

3.2.4.3. Physiological synchrony for the detection of interspersed stimuli

We designed a paradigm to detect relevant events using gradually increasing thresholds to capture the gradual nature of attentional engagement. Figure 3-1 provides a visual explanation of our detection paradigm. Consider the EEG, EDA and IBI response traces that were recorded during the experiment. The timestamps of the data recordings can be separated in moments where interspersed stimuli were presented and where an event detection would thus be considered correct (True) and moments where no interspersed stimuli were presented and where an event detection would be considered incorrect (False). Rather than using the raw physiological responses, the detection paradigm is based on the PS between the participants as a function of time. Now let us define a threshold t. The moments in time where the synchrony is higher than t are marked as an event (Positive) and the moments in time where the syn-

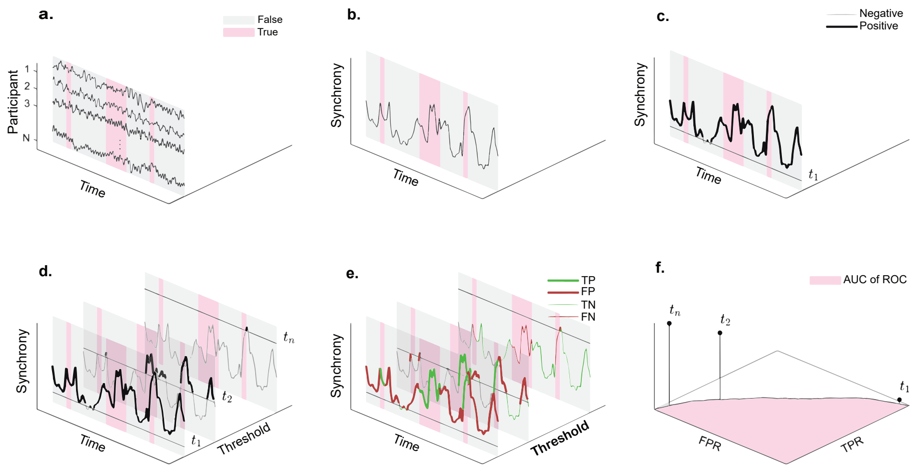

Figure 3-1. Illustration of the event detection paradigm. a. Consider the physiological response traces recorded from N participants who were presented with the same external stimuli at the same time. The timestamps of the data recordings can be separated in moments where events were presented and where an event detection would thus be correct (True) and moments where no events where presented and where an event detection would thus be incorrect (False). b. Rather than considering the raw physiological responses, the detection paradigm is based on the PS between the participants. c. Now let us define a threshold t. The moments in time where the synchrony is higher than t are marked as an event (Positive) and the moments in time where the synchrony is lower than t are marked as a non-event (Negative). d. Rather than using a single value for t, we consider a gradually changing threshold t_0 to t_n, so that at t_0 all data are marked as Positive and at t_n all data are marked as Negative. e. For each iteration t, we can now define the true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) and thus compute the true-positive rate (TPR) and false-positive rate (FPR). f. Plotting the FPR versus the TPR – both as a function of t – results in the receiver operating curve (ROC). Detection performance is defined as the standard metric of area under the ROC (AUC of ROC).

chrony is lower than t are marked as a non-event (Negative). Rather than using a single value for t, we consider a gradually changing threshold t, ranging from the minimum inter-subject correlation value to the maximum inter-subject correlation value. For each iteration of t, we can now define the true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN). Using this, the true-positive rate or sensitivity (TPR) is then computed as,

$$TPR = \frac{TP}{TP + FN}$$

and the false-positive rate (FPR) or specificity as,

$$FPR = \frac{FP}{FP + TN}$$

Plotting TPR against FPR provides the receiver operating curve (ROC). Detection performance was assessed using the standard metric of the area under the

ROC (AUC of ROC).

Chance level performance was assessed using permutations with randomized stimulus timing. In each permutation, the timing of all interspersed stimuli was randomized between the start and the end of the experiment. The same procedure as above was then applied to obtain the AUC of ROC metric of performance with random stimuli. This procedure was performed on 1000 renditions of such randomized data.

The above-mentioned procedure was repeated 2 × 3 × 4 times, namely for:

(1) Two attentional groups; considering PS between AA participants and PS between SA participants.

(2) Three stimulus types; considering as events (True) either blocks of beeps, emotional sounds, or both of these.

(3) Four physiological measures in which PS is computed; EEG, EDA, heart rate and a multimodal metric that is composed of PS in EEG, EDA and heart rate. To compose this multimodal metric, the PS in EEG, EDA and heart rate were each z-scored. The multimodal PS value at each timestamp was then computed as the average of the z-scored PS values in EEG, EDA and heart rate at that timestamp, for all timestamps ranging from zero to the end of the experiment.

In each condition, one-tailed two-sample t-tests were conducted to test whether detection performance was higher than chance level performance.

3.2.4.4. Correspondence between physiological synchrony and reported engagement with the audiobook

While for the SA group, the timing of short stimuli served as "ground truth" relevant events to compare to the moments of high PS, we did not know a priori what constituted relevant, engaging moments in the audiobook. To systematically examine whether moments of high PS were associated with moments of high relevance in the audiobook, we performed a follow-up test in which a second cohort of participants judged clips of the audiobook that were found to be associated with either high or low PS. We recruited 29 participants through the Prolific online experiment environment. All participants signed informed consent before participating. The participants received a small monetary reward for the invested time. We only included participants who indicated to be fluent in Dutch.

We selected clips based on continuous signals of PS among AA participants. We detected the positive and negative peaks in the signals using the 'findpeaks' function in MATLAB. For each measure (EEG, EDA, heart rate and the multimodal metric), the six peaks with highest positive peak-amplitude and six peaks with largest negative peak-amplitude were selected. For each detected peak, we created a 10 s sound clip, that was composed of the 10 s of audio before the detected peak. For four measures, this thus resulted in a total of 48 clips. Clips associated with peaks that were within 10 s of each other were considered to be overlapping, and were merged into one clip by using only the latest of the two clips in time. This resulted in a total of 38 clips that were presented to the participants.

The procedure of the online test was similar to the initial experiment. The participants were first presented with the same audiostream that was presented to the initial cohort of participants. The participants were instructed as participants from the AA group, i.e., to focus their attention on the narrative of the audiobook and ignore any interspersed stimuli as much as possible. After listening to the book, the participants were asked the same questions about the content of the narrative as participants in the initial cohort. We then presented the participants with the sound clips, each of them directly followed by a rating scale. Participants were instructed to rate the preceding clip using an 11-point Likert scale, ranging from 0 to 10. The lower the score, the more the participant's experience corresponded to the words on the left side of the scale (Dutch: 'verveeld', 'kalm', 'ontspannen'; Translated to English: 'bored', 'calm', 'relaxed'). The higher the score, the more the participant's experience corresponded to the words on the right side of the scale (Dutch: 'geïnteresseerd', 'geboeid', 'emotioneel', 'intens'; Translated to English: 'interested', 'fascinated', 'emotional', 'intense'). Using these words, we intended to capture mental states that are expected to be associated with perceiving relevant events, such as engagement, attention and arousal.

For each modality, we tested whether audio clips corresponding to a positive peak in PS were rated as more 'engaging' than audio clips corresponding to a negative peak in PS, using a Wilcoxon signed-rank test. Participants who answered less than three out of ten questions correctly on the questionnaire about the content of the audiobook were considered as not having participated seriously (AA participants in the main experiment answered 5.8 ± 2.0 questions correctly). This concerned three participants. Removing their data left us with data of 26 participants.

3.3. Results

3.3.1. Detection of interspersed stimuli using physiological synchrony

Figure 3-2 and Table 3-1 show our measure of interspersed stimuli detection performance, the AUC of ROC as described in the methods. It is presented separately for AA and SA participants; in EEG, EDA, heart rate, and the multimodal metric; and for blocks of beeps, emotional sounds or both of these stimuli together as to-be identified events. Figure 3-2 and Table 3-1 also show the mean and standard deviation AUC of ROC of permutations with randomized event timing as a chance level baseline. Detection performance was largely in line with hypotheses 1 - 3. For the AA group, we found that, as expected, only the occurrence of emotional sounds could be predicted, using PS in EDA ($p < .001$) or the multimodal metric ($p = .003$). For the SA group, occurrences of beep blocks could be detected well above chance level by PS in EEG, EDA and the multimodal metric ($p < .001$, $p = .002$ and $p < .001$, respectively). The occurrence of emotional sounds could be detected significantly better than chance using PS in EDA, heart rate and the multimodal combination ($p = .043$, $p = .023$ and $p=.011$, respectively). When stimuli were not differentiated according to stimulus type, detection performance was well above chance level for PS in EEG ($p < .001$), EDA ($p < .001$) and the multimodal metric ($p < .001$), but not for PS in heart rate.

Table 3-1. AUC of ROC metric of stimulus detection performance using PS in EEG, EDA, IBI and a multi-modal combination of the three (MM).

	AA				SA			
	EEG	**EDA**	**IBI**	**MM**	**EEG**	**EDA**	**IBI**	**MM**
Beep blocks	.506 (.500 ± .026) p = .410	.470 (.500 ± .042) p = .769	.463 (.500 ± .042) p = .813	.468 (.500 ± .040) p = .790	.642 (.499 ± .030) p < .001	.627 (.501 ± .044) p = .002	.475 (.498 ± .042) p = .708	.635 (.499 ± .039) p < .001
Emotional sounds	.540 (.502 ± .039) p = .168	.658 (.501 ± .048) p < .001	.576 (.500 ± .047) p = .055	.629 (.501 ± .046) p = .003	.505 (.498 ± .040) p = .433	.580 (.500 ± .046) p = .043	.592 (.500 ± .045) p = .023	.604 (.500 ± .045) p = .011
Both stimuli	.517 (.501 ± .023) p = .238	.522 (.500 ± .036) p = .273	.492 (.500 ± .036) p = .596	.512 (.501 ± .034) p = .370	.621 (.499 ± .026) p < .001	.631 (.501 ± .036) p < .001	.507 (.499 ± .036) p = .415	.644 (.499 ± .033) p < .001

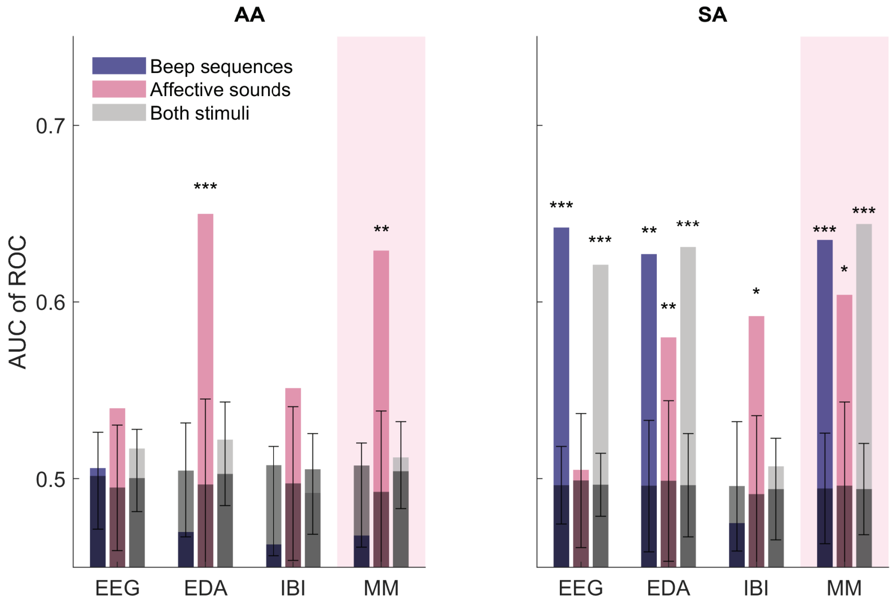

Figure 3-2. AUC of ROC metric of stimulus detection performance using PS in EEG, EDA, IBI and a multi-modal combination of the three (MM). Performance is shown for the AA and SA groups, when considering only beep blocks or emotional sounds as true positives and when considering both types of stimuli as true positives. In addition, the mean and standard deviation chance level detection performance based on 1,000 renditions with randomized stimulus timing is shown with test results comparing detection performance to chance level (*p < .05,**p < .01, ***p < .001). Note that for the AA group, the emotional sounds but not the beeps are expected to draw (bottom-up) attention, i.e., for the AA group we expect high AUC for emotional sounds only. For the SA group, both beep sequences and emotional sounds are relevant and expected to draw attention.

3.3.2. Correspondence between physiological synchrony and reported engagement

Figure 3-3 shows engagement ratings of audio clips corresponding to positive peaks and ratings of audio clips corresponding to negative peaks for PS in EEG, EDA, heart rate and the multimodal metric. Results did not follow our hypothesis that audio clips corresponding to positive peaks were rated as more engaging than audio clips corresponding to negative peaks. In fact, in EEG and EDA the opposite effect was found (Wilcoxon test statistic: $W = -3.06$, $p = .002$; $W = -3.44$, $p < .001$, respectively). In heart rate and the multimodal metric no significant difference between ratings corresponding to either positive or negative peaks was found ($W = 0.70$, $p = .486$, $W = 0.87$, $p = .385$).

3.4. Discussion

In the sections below, the hypotheses as stated in the Introduction are dis-

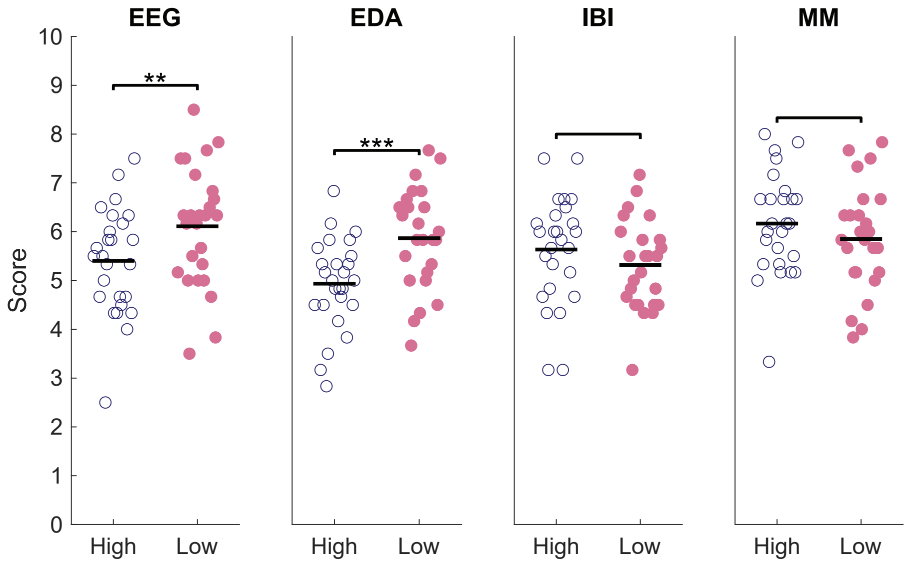

Figure 3-3. Self-reported engagement scores for audio clips corresponding to moments in the audiobook with high PS (blue, open markers) or low PS (pink, closed markers) in EEG, EDA, IBI and the multimodal metric (MM) (**$p < .01$, ***$p < .001$).

cussed separately.

3.4.1. Hypothesis 1: attentional instruction and stimulus type

We hypothesized that interspersed stimulus detection performance would depend on the attentional group (AA or SA) and the interspersed stimulus type (emotional sounds or beeps), due to bottom-up and top-down mechanisms of attention (Lang, 1995; Öhman et al., 2001; Schupp et al., 2003). For the SA group, we hypothesized that detection performance based on PS would be above chance for all interspersed stimuli, whereas for the AA group we hypothesized that detection performance would be above chance only for emotional sounds, but not for beeps. Results were largely in line with this hypothesis: for the AA group only the emotional sounds were detected with an accuracy above chance level, whereas for the SA group both stimulus types could be detected with above chance level accuracy. Note again that detection performance cannot be directly compared between the different stimulus conditions. There were differences in detection performance between the used physiological measures, these are discussed in the next section.

3.4.2. Hypothesis 2: physiological measure and stimulus type

Besides the dependency of detection performance on the attentional group and interspersed stimulus type, we expected that detection performance would depend on the used physiological measure and stimulus type. As hypothesized, EEG worked best for the detection of blocks of beeps, while it did not work well for the detection of emotional sounds. We also found EDA to perform well for detecting blocks of beeps. For the detection of emotional sounds, we hypothesized that the peripheral measures (EDA and heart rate) would perform well relative to EEG. Indeed, for both attentional groups, PS in EDA and heart rate perform relatively well for the detection of emotional sounds. Detection performance was significantly above chance using EDA in both groups and using heart rate in the SA group, and near significance ($p = .055$) using heart rate in the AA group, whereas detection performance using PS in EEG was far from significant for emotional sounds in both groups. We think that the observed EEG PS differences between the types of stimuli are the result of both the difference in mental processing (top-down, effortful attention for beeps, versus bottom-up, affective processing for emotional sounds) and low level stimulus features. The beep blocks consisted of precisely-timed, repeated beep occurrences, with constant sound levels, while our emotional stimuli consisted of sounds with irregular sound profiles. The positive results obtained with peripheral measures provide further insight in the mechanisms underlying PS. Whereas previous findings on peripheral PS have been viewed in terms of social relation (Palumbo et al., 2017), we here show that PS in peripheral measures can also be explained by shared attentional engagement. It may be the case that shared attention also underlies results found in contexts of social relation.

3.4.3. Hypothesis 3: combining physiological measures

As the three used physiological measures vary with respect to their ability to reflect different mental states, we hypothesized that combining the physiological measures into a single multimodal metric of PS would result in relatively high detection performance when differences between stimulus types are disregarded. Indeed, the multimodal metric performs best when considering both emotional sounds and beeps as relevant events. Detection accuracies are slightly higher than for the best performing unimodal measure when considering emotional sounds or both types of stimuli but not when considering blocks of beeps. We expect that sensor fusion is not beneficial when variables are highly correlated (Hogervorst et al., 2014) – for example for physiological variables all reflecting mental effort – but that sensor fusion can benefit from tasks involving emotional processing besides effortful attentional processing. Besides potentially higher detection performance, the main advantage of multimodal PS seems to be the robustness regarding different types of stimuli, i.e., detection performance varies less between different types of stimuli than for single physiological metrics. In previous work similar effects have been found. For example, when using sensor fusion on machine learning models to distinguish between 13 emotional states, maximum performance was not higher for the multimodal metric, but performance was more robust across the range of emotional states (Verma and Tiwary, 2014). In the end, although adding sensors does not lead to much higher performance compared to the most suitable unimodal recording, a multimodal approach seems to enable detection of relevant events when it is unknown what the best measure for certain stimulus types is. Also note that it is not always known whether certain stimuli will induce mostly effortful cognitive or emotional processing; in many practical cases such processes can co-occur and vary between individuals.

3.4.4. Hypothesis 4: events in audiobook

We hypothesized that audio clips corresponding to moments of highest PS would be post-hoc scored as more engaging than audio clips corresponding to moments of lowest PS, but our findings indicated rather the opposite. These findings may be caused by a mismatch between our index of PS and the rating scale of experienced engagement. Post-hoc qualitative analysis of the selected audio-clips revealed that part of the audio clips corresponding to very high PS in EEG coincided with short-term moments of tension or engagement, as expressed through keywords (e.g., swear words) and salient intonation (e.g., a phrase spoken in a very indignant manner). This is in line with earlier work, where moments of high PS in EEG were found to correspond to moments in video clips marked by a high level of short-term suspense, tension or surprise, such as the sight of a gun (Poulsen et al., 2017). Indeed, emotional images and sounds that are rated as highly arousing induce responses in peripheral and central physiological measures (Bradley and Lang, 2000, 2007b), which in term may lead to strong PS. In our used audiobook, the keywords that may have driven the particularly high PS contained relatively little important information about the narrative of the story. It seems that this could have been the aspect rated by participants using our engagement scale, leading to a mismatch between self-reported engagement and PS. However, this speculation would need to be investigated further, preferably without having to rely on varying engagement judgements after the fact, but for instance with systematic sentiment analysis (Wollmer et al., 2013). It is important to further specify what types of attentional engagement can and cannot be captured by PS and how that is dependent on the psychological measure used. Attentional engagement to well-timed events will be better reflected in PS than attentional engagement to less well-timed event on a more abstract level.

3.4.5. Limitations

It should be noted that the stimulus detection performance when not taking stimulus type into consideration ('both stimuli') were mainly driven by detection performance of beep blocks. These beep blocks were interspersed for a total of 810 s, whereas the emotional sounds were only interspersed for a total of 216 s. The stimuli also differed on other aspects. For instance, the beep blocks consisted of precisely timed beeps with immediate stimulus onset equal across trials, whereas the emotional sounds all differed in sound profile. For the AUC of ROC metrics when considering detection of both types of stimuli, beep blocks thus influence the performance metric more than the emotional sounds. While this can be seen as a limitation, this is exemplary for real life situations, where one is interested in detecting relevant, attentionally engaging events, without further specifying or knowing the different types of stimuli, and the proportion of in which they occur; i.e., in such a situation, the 'both stimuli' situation is the default.

We must also note that our simple multimodal approach is certainly not the optimal approach to combine data. In particular, we expect that detection performance can be enhanced by compensating for differences in response latencies across measures. To illustrate the difference in response latency, in response to the same set of emotional sounds, response peak latency ranges from a few 100 ms for EEG to multiple seconds in EDA heart rate (Bradley and Lang, 2000; Hettich et al., 2016). In this paper we simply averaged over response traces in a point-wise fashion, meaning that response-induced peaks may be spread out and their amplitude reduced. Our current results should therefore be interpreted as a first confirmation that multimodal sensor fusion can be of added value, but we expect that other approaches can greatly enhance performance. In future work we would like to explore other methods for the combination of physiological measures into a multimodal metric of PS.

3.4.6. Conclusion

We determined PS in EEG and EDA, heart rate and a multimodal fusion of these three sensors in two groups of participants, that were instructed to attend either to the narrative of an audiobook or to interspersed auditory events. We found that PS could detect the relevant interspersed stimuli with accuracies well above chance level, but also found that moments in the audiobook corresponding to high PS were not rated as more engaging than moments corresponding to low PS. Our results support the notion that PS can be valuable when interested in the course of attentional engagement over time. Currently the relation between PS and engagement is only established for well-defined, interspersed emotional or effortful cognitive stimuli, whereas the relation between PS and a more abstract self-reported metric of engagement is not yet established. We further note that obtained results vary between the used physiological measures. Interesting from a user perspective, EDA worked best overall. These results should enable researchers to monitor PS in situations where intrusive EEG measurements are not suited. However, we also note that the optimal physiological metric may be dependent on the goal of a study and suggest to choose a measure matching with the stimulus of interest. EEG works especially well for well-timed effortful cognitive stimuli, heart rate works especially well for emotional stimuli and EDA works quite well on both types of stimuli. When the stimulus type is unknown, a multimodal metric may work best as it seems most robust across a broad range of stimuli.

Acknowledgments

We thank Ana Borovac for her help with performing the experiment.

4.

EEG measures of attention toward food-related stimuli vary with food neophobia

Stuldreher, I.V., Kaneko, D., Hiraguchi, H., van Erp, J.B.F., Brouwer, A.M. (2023). EEG measures of attention toward food-related stimuli vary with food neophobia. *Food Quality and Preference*, ***106***, 104805. doi: 10.1016/j.foodqual.2022.104805

Abstract

Humans differ strongly in their willingness to try novel foods. Hesitance to try new foods is referred to as food neophobia. Understanding food neophobia is important, as it can be a significant barrier to adopt a healthy, balanced or plant-based diet. We here use electroencephalogram (EEG) recordings to obtain insight in the early attentional processes towards food stimuli as a function of food neophobia. 43 Dutch participants completed the food neophobia scale after which they were presented with pictures of familiar and unfamiliar foods and a 15-minute movie about the origin and production of an unfamiliar food. We extracted two EEG-based metrics of attention: the late positive potential (LPP) amplitude in response to the food pictures, and inter-subject correlations (ISC-EEG) during the movie. The latter is a novel metric, based on similarities in EEG over time between individuals who are presented with the same stimulus, and suitable for examining attention towards continuous stimuli such as movies. Additionally, participants were asked to taste familiar and unfamiliar soups, and they were asked to rate the pictures and soups for valence and arousal. ISC-EEG and the LPP amplitude increased and sip size decreased with food neophobia, not only for unfamiliar food pictures, but also for familiar food pictures. Self-reported emotional experience was affected by food neophobia for unfamiliar food pictures or soups, but not for familiar ones. We conclude that food neophobia is associated with increased attentional processing and immediate implicit behavior, for all food stimuli and not only for unfamiliar food stimuli. This indicates that all food-related stimuli are of high importance to food neophobic individuals and that self-reported emotion does not capture the entire experience of food. The results also indicates that, unlike the name suggest, food neophobia does not only affect processing of novel foods, but of any food regardless of familiarity.

4.1. Introduction

Humans differ strongly in their willingness to try novel foods. This willingness can be captured through the food neophobia scale (Pliner and Hobden, 1992). Individuals that score high on the food neophobia scale are generally hesitant to try or buy new foods (Arvola et al., 1999; Tuorila et al., 2001; Bäckström et al., 2004; Schickenberg et al., 2008; Henriques et al., 2009; Chung et al., 2012; Siegrist et al., 2013), including ethnic, unfamiliar foods (Choe and Cho, 2011; Filippo D'Antuono and Bignami, 2012). Food neophobia can be a significant barrier for a healthy, balanced or sustainable diet and can thereby lead to disordered eating patterns (Falciglia et al., 2000; Eertmans et al., 2005; Knaapila et al., 2011; Jaeger et al., 2017). Understanding food neophobia is important to remove the barrier and threat eating disorders.

Rozin and Fallon, 1980 proposed three main reasons for rejection of novel foods: dislike of its sensory characteristics, a fear of negative consequences of eating it, and disgust arising from the food's origin. Food neophobia may be part of wider cross-modal avoidance behavior of novel stimuli (Pliner and Hobden, 1992; Raudenbush and Frank, 1999). An increasing number of studies demonstrate a role of anxiety mediation in food neophobia, with studies showing significant correlations between measures of food neophobia and anxiety (Pliner and Hobden, 1992; Galloway et al., 2003). Anxiety traits can cause attentional biases toward threatening pictures compared to nonthreatening pictures (Li et al., 2005; Holmes et al., 2008; Berggren and Eimer, 2021). In a similar fashion, one could expect that food neophobia is associated with attentional biases toward novel foods. Indeed, there are indications that food neophobia results in changes in attention. For instance, children with high food neophobia showed a greater attentional bias as measured with a visual probe task towards pictures of unfamiliar fruits and vegetables compared to familiar fruits and vegetables than children with low food neophobia (Maratos and Staples, 2015).

We know of only one study that examined effects of food neophobia on neurophysiological measures. (Raudenbush and Capiola, 2012) found that heart rate, electrodermal activity and respiration were significantly higher in individuals with high food neophobia compared to controls when they were presented with a variety of food pictures. This indicated that food phobic individuals experience heightened arousal upon the presentation of food, and heightened arousal often co-occurs with heightened attention.

While the studies above suggest a role for attention in food neophobia, we are not aware of studies using electroencephalogram (EEG) to examine this. EEG is of special interest, because not only is it more directly related to attention than other, peripheral, physiological measures as mentioned above, it also offers a view on the very early stages of the attentional process (Polich, 2007), before any behavioral response has taken place.

In the current work we evaluate the association between food neophobia and attention toward food-related stimuli. We explore two EEG-based measures of attention: event-related potentials (ERP) upon presentation of food pictures and physiological synchrony as captured through inter-subject correlations in EEG (ISC-EEG) over the course of a narrative movie clip about ethnic food.

ERPs are extracted from the EEG in response to repeated presentation of stimuli such as pictures or sounds. Positive potentials obtained from the electrodes over the parietal cortex are often interpreted as reflecting differences in the allocation of attention (Näätänen, 1988). The P3 component, a positive deflection at roughly 300ms after stimulus onset, is known to be larger in response to oddball stimuli or to stimuli that are instructed to be attended (Polich, 2007). In response to affective pictures, similar, positive deflections over the parietal cortex have been reported. These positive deflections are often referred to as the Late Positive Potential (LPP), as the peak of this ERP occurs somewhat later than the traditional P3 in response to simpler stimuli. The slight timing difference is assumed to be due to the relatively high information load in affective pictures (Bradley et al., 2007). Still, the LPP can be interpreted as reflecting increased attention for motivationally relevant stimuli (Lang et al., 1997). Indeed, larger amplitudes in the LPP are observed for emotionally engaging than for neutral stimuli, where highly arousing pictures result in the largest amplitudes (Bradley et al., 2007).

ERPs have been studied in the context of food before. Participants with binge eating disorder showed greater LPP in response to high-caloric food pictures than control participants (Svaldi et al., 2010). Similarly, high external eaters – people with the tendency to eat when exposed to food-related cues – showed larger P3 amplitudes in response to food pictures than low external eaters (Nijs et al., 2009).

A recent measure of attention is based on the similarity of the EEG across individuals presented with the same stimulus, as assessed through ISC-EEG. A major advantage of this measure is that it enables the use of more naturalistic and dynamic stimuli than traditional controlled and repeated picture presentation. Parts of narrative stimuli that attract attention, such as engaging scenes of a popular television series, or emotionally salient sounds, result in heightened ISC-EEG (Dmochowski et al., 2014; Stuldreher et al., 2020a). Studies have also shown that individuals with higher ISC-EEG have better recall of the narrative, which further substantiates the association between ISC-EEG and attentional engagement (Cohen and Parra, 2016; Stuldreher et al., 2020b).

As for ERPs, top-down guided selective attention affects ISC-EEG. Individuals showed significantly higher ISC-EEG when instructed to focus their attention on the stimulus than when instructed to focus their attention inward upon presentation of the stimulus (Cohen et al., 2018). Furthermore, two groups of individuals with different selective attentional instructions to the same stimulus showed distinct patterns of ISC-EEG, where individuals show higher ISC-EEG with others in the same selective attentional condition than with individuals in the other condition (Stuldreher et al., 2020b).

In the current study, we investigate how attentional processes in response to food-related stimuli vary with food neophobia, using EEG ERPs in response to familiar and unfamiliar food pictures, and ISC-EEG in response to a movie about unfamiliar food. Dutch and Japanese food serve as familiar and unfamiliar food types for the Dutch participants (cf. (Kaneko et al., 2021)).

Besides EEG measures, we also examine a behavioral measure and explicit self-reports that reflect food experience after the initial attentional stage. To obtain self-reported experience we use the EmojiGrid, a 2D pictorial scale specifically designed to rate experience elicited by food-related stimuli (Toet et al., 2018; Kaneko et al., 2019b). As a behavioral measure, we record sip size upon tasting a familiar and unfamiliar soups. Sip size has shown to distinguish between high and low-valence drinks with more discriminative power than self-reports, or neuro- and psychophysiological response measures (Kaneko et al., 2019a). These three types of measures reflect different levels of emotional processing (Kaneko et al., 2018). Where EEG is an implicit measure that is capable of reflecting unconscious and early processes, and self-report is an explicit measure that can only reflect conscious experience, sip size can be considered as an implicit measure that is somewhere in between.

In sum, we here investigate how attentional processes in response to food-related stimuli, as indicated by EEG, vary with food neophobia. To align these results to explicit judgements, and implicit initial behavior, we additionally examine the association between food neophobia and rated food experience as well as sip size, in the same participants, and partly for the same stimuli that were used to elicit EEG responses. We hypothesize:

1) Food neophobia is positively correlated with attention toward unfamiliar but not toward familiar food stimuli, as captured by LPP ERP amplitude and ISC-EEG.

2) Food neophobia is negatively correlated with rated pleasantness of unfamil-

iar but not of familiar foods, as captured by EmojiGrid reports.

3) Food neophobia is negatively correlated with sip size of unfamiliar but not of familiar soups.

4.2. Materials and methods

4.2.1. Participants

53 participants were recruited through the institutes participant pool. Before performing the study, ethical approval was obtained from the TNO Institutional Review Board (IRB). The approval is registered under reference 2020-117. All participants signed informed consent before participating in the experiment, in accordance with the Declaration of Helsinki. After successful participation, participants received a small monetary compensation for their time and traveling costs.

Due to processing errors in the experiment script, the time-synchronization of the EEG with the stimuli could not be guaranteed for 10 participants. From 53 participants, data of only 43 participants (19 female) are therefore used in further analysis. Their age ranged from 19 to 64 years (M = 46.6, SD = 15.3) and Body Mass Index ranged from 18.4 to 46.4 (M = 25.6, SD = 5.2). The average time since their last meal consumption was 2.75 hours (SD = 2.4).

4.2.2. Materials

EEG was recorded at 1024 Hz using an ActiveTwo Mk II system (BioSemi, Amsterdam, Netherlands) with 32 active Ag/-AgCl electrodes, placed on the scalp according to the 10–20 system, together with a common mode sense active electrode and driven right leg passive electrode for referencing. The electrode impedance threshold was set at 20 kOhm.

4.2.3. Stimuli and design

The experiment consisted of three phases: a first picture phase, a movie phase and a second picture phase. The experimental procedure is depicted in Figure 4-1.

In each of the two picture phases, participants were presented with 80 images of food from the Cross Cultural Food Images Database (CROCUFID, (Toet et al., 2019)) on a computer screen in a randomized order. The presented images were of four different categories: unfamiliar Japanese food, familiar Dutch food, palatable food (i.e., universal food, such as fruits) and unpalatable food (i.e., molded food and food that was beleaguered by insects or snails). The latter two image categories represented 'ground truth' pleasant/neutral and unpleasant food and served to allow checking for sensitivity of the different measures and

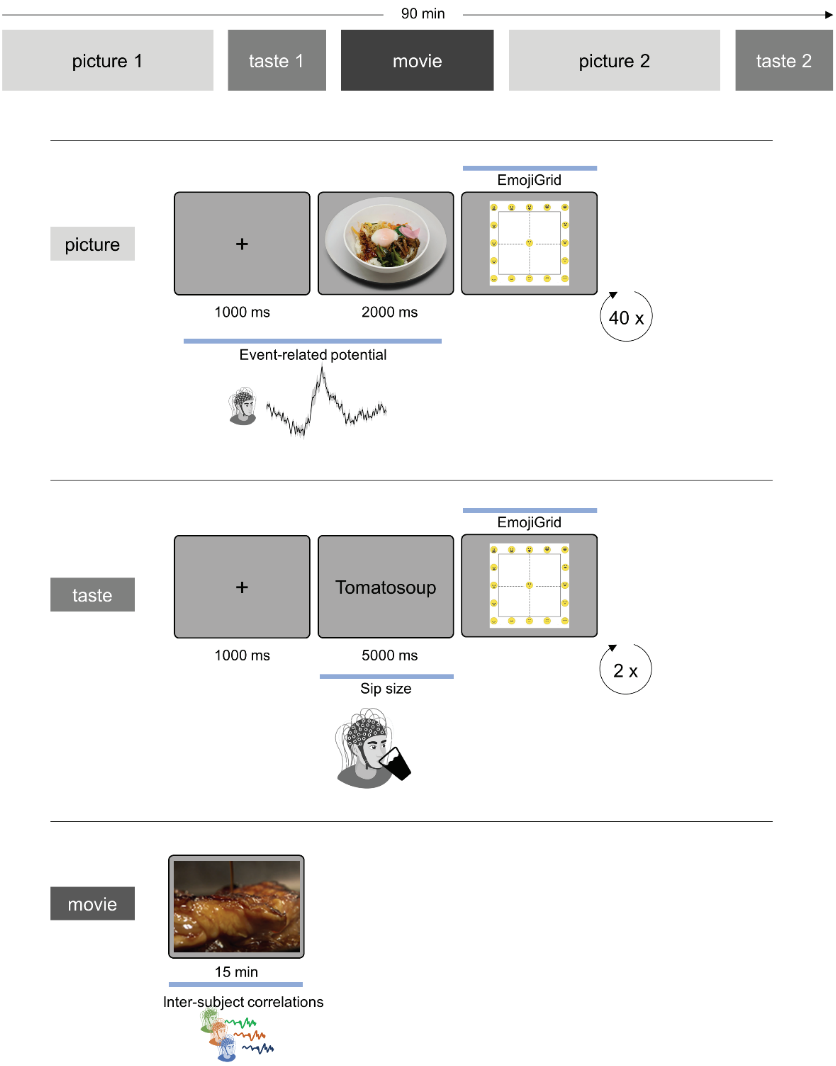

Figure 4-1. Overview of the experimental procedure. The experiment consisted of a picture phase and tasting phase, followed by a movie phase and then followed by another picture and tasting phase.

are not further discussed in the present manuscript. Twenty pictures of each category were presented per picture phase. For easier recognition a symbol in the right bottom corner of the food image displayed whether the depicted food is Dutch, Japanese or universal (palatable and unpalatable). An example of the Japanese and Dutch categories with symbols is displayed in Figure 4-2. Each

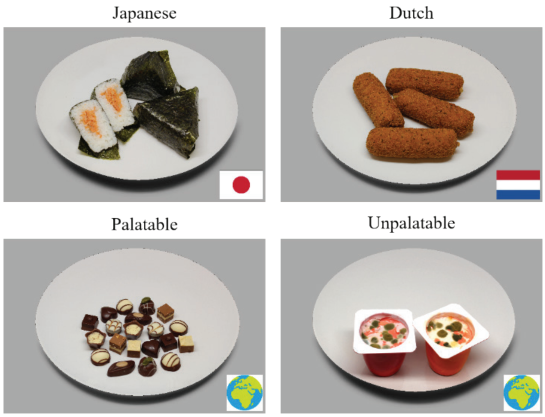

Figure 4-2. Examples of CROCUFID pictures in Japanese and Dutch food categories including the symbols indicating either a Japanese or Dutch food picture.

image was presented for two seconds preceded by a fixation cross displayed for half a second. Immediately after viewing each picture, participants were prompted to rate their emotion using the EmojiGrid (Toet et al., 2018). The EmojiGrid is a graphical and language-independent self-reporting tool to measure the emotional dimensions of valence (x-axis) and arousal (y-axis) in a food-related context. At the end of each picture phase a Dutch and a Japanese soup were presented in counterbalanced order to the participants to taste. The four soups were vegetable soup, tomato soup (familiar Dutch soups) and sumashi soup, Miso soup (unfamiliar Japanese soups). After tasting each soup, food-evoked emotions were rated using the EmojiGrid. The amount of soup consumed was recorded as an implicit measure of positive emotion.

Following the first picture phase, participants were presented with a 15-minute movie about the origin and production of Japanese Kikkoman soy sauce. Prior to the movie, participants were instructed differently based on a social pressure condition they were assigned to. Participants in the control group were told that their EEG sensors would be calibrated and meanwhile they could watch this movie. Participants in the social pressure condition were told a story aimed to increase social pressure to report liking of Japanese food after watching the movie.

After the movie phase the second picture phase started. The entire experimental procedure, including fitting the EEG electrodes, took about 60 minutes.

The social pressure intervention did not result in any between-group difference (Sabu et al., 2022). For the purpose of the current study, we collapse over the two social-pressure related instructions.

4.2.4. Self-report measures

Before the presentation of food pictures, participants answered a set of questionnaires. They provided a number of descriptives. In addition, participants filled out the food neophobia scale (Pliner and Hobden, 1992). This questionnaire consists of ten statements, for which a rating on a 7-point Likert scale, ranging from ’strongly disagree’ to ’strongly agree’, can be given. The outcome – a score from 10 to 70 – indicates the willingness to try novel foods. High scores indicate high food neophobia, meaning unwillingness to try new foods, while low scores indicate enthusiasm to try novel food.

4.2.5. Analysis

4.2.5.1. Pre-processing

EEG was processed using MATLAB 2021a (Mathworks, Natick, MA, United States) and the EEGLAB v14.1.2 toolbox for MATLAB (Delorme and Makeig, 2004). Potentials reflecting ocular or muscle-related artifacts were removed using logistic infomax independent component analysis (ICA; Bell and Sejnowski, 1995). Before doing so, EEG was down-sampled to 256 Hz and high-pass filtered with a 1 Hz passband frequency using the standard FIR-filter implemented in the EEGLAB function pop_eegfiltnew. The 1 Hz high-pass cut-off frequency was chosen as it has shown to work better for ICA compared to lower cut-off frequencies (Winkler et al., 2015). 50 Hz line noise was then removed using the cleanline plugin as implemented in PREP pipeline (Bigdely-Shamlo et al., 2015). Further artifacts were then removed using the clean_rawdata plugin for EEGLAB. Removed channels were interpolated and an average channel reference was applied. For ERP analysis, continuous data was epoched from -500 to 2000 ms with respect to the onset of pictures. For ISC analysis, the continuous EEG over the course of the movie was used. ICA was performed on either the epoched or continuous data and ICLabel was used to identify and remove artifactual independent components reflecting ocular or muscle-related artifacts (Pion-Tonachini et al., 2019).

4.2.5.2. Event-related potentials in response to the pictures

For each individual, the event-related potentials in response to unfamiliar and familiar food pictures were extracted. First, the epoched data were lowpass filtered with a passband at 30 Hz using pop_eegfiltnew. Then, data were baselined by extracting the average value of the signal from 200 ms to 0 ms before stimulus onset. Data of each participant were averaged over all pictures from the same picture category, aggregating over the picture phases before and after the movie. For each picture category and each participant, the LPP amplitude was extracted by finding the maximum of the averaged ERP at the midline-parietal site (electrode Pz) from 500 to 1000 ms after picture onset.

4.2.5.3. ISC-EEG in response to the movie

We computed physiological synchrony by computing ISC-EEG. ISC-EEG was evaluated in the correlated components of the data, following earlier work (Dmochowski et al., 2012, 2014; Stuldreher et al., 2020b). The goal of the correlated component analysis is to find underlying neural sources that are maximally correlated between participants, using linear combinations of electrodes. The components were extracted based on all datasets, after which EEG data from each participant were projected on the component vectors. Participant-to-group ISC-EEG was then computed as the sum of correlations in the first three component projections. The first three components were used as correlations in higher order projections are usually close to chance level (Ki et al., 2016).

4.2.5.4. Association with food neophobia

Results were analyzed in relation with the food neophobia score. We performed correlation analyses to highlight the continuous spectrum of the food neophobia scale, following (Jaeger et al., 2021). We investigated correlations with all the implicit and explicit measures included in the study, i.e. LPP amplitude, ISC-EEG, sip size and self-reported valence and arousal. These analysis were performed using Pearson correlations as implemented in the 'corr' function of MATLAB 2021a (Mathworks, Natick, MA, United States).

4.3. Results

4.3.1. Event-related potentials in response to food pictures

We investigated the relation between LPP ERP amplitude and food neophobia score. Figure 4-3 shows the average event-related potentials separately for each picture category, food neophobic and food neophiliac participants (participants scoring above and below the median food neophobia score, respectively – note that this is not an absolute categorization, but relative to scores in our specific sample of participants). The figure suggests higher LPP amplitudes for neophobic compared to neophiliac participants for all picture categories. In line with this, Figure 4-4 shows significant correlations between LPP amplitude in response to unfamiliar and familiar food pictures and food neophobia ($r = 0.36$, $p = .017$; $r = 0.36$, $p = .018$).

4.3.2. ISC-EEG during the movie

We investigated the relation between ISC-EEG and the food neophobia score. Figure 4-5 shows that food neophobia and ISC-EEG are significantly correlated ($r = 0.44$, $p = .003$) where higher food neophobia scores are associated with higher ISC-EEG.

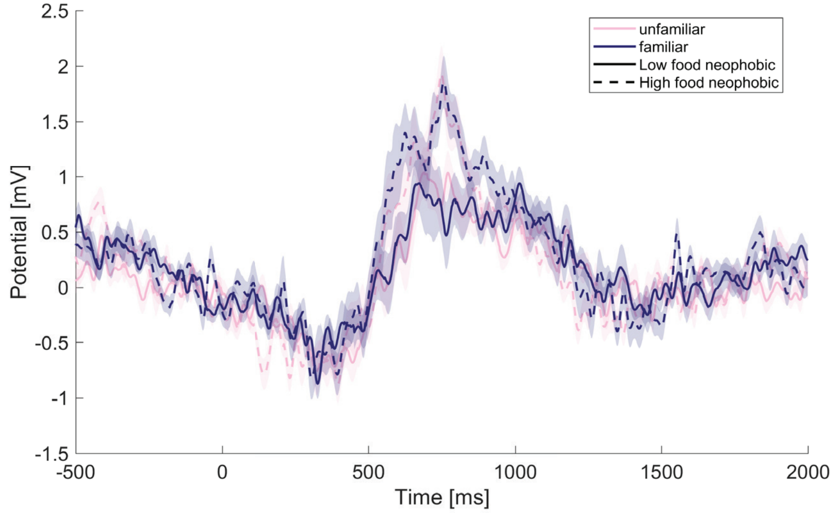

Figure 4-3. Event-related potentials averaged over familiar and unfamiliar picture types and averaged over neophobic and neophiliac participants. Shaded areas depict standard error of the mean.

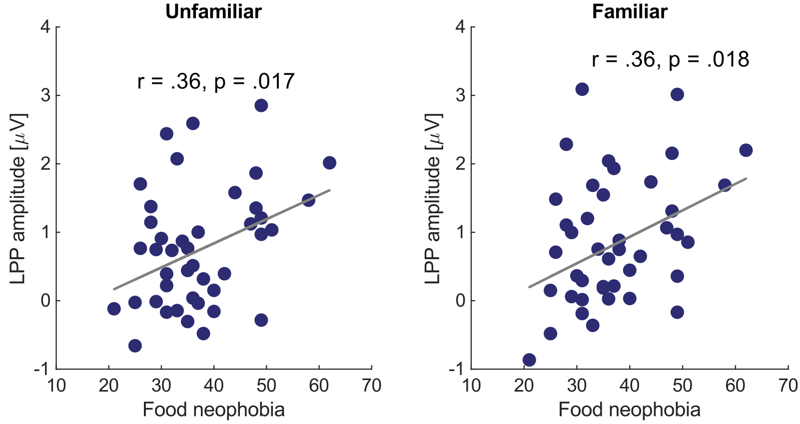

Figure 4-4. Correlations between food neophobia and LPP amplitude in response to unfamiliar and familiar food pictures. Each data point represents a participant.

4.3.3. EmojiGrid after image presentation

Figure 4-6 shows that self-reported valence in response to unfamiliar food pictures is strongly correlated with food neophobia ($r = -0.53$, $p < .001$) where high food neophobia is associated with low valence scores. As shown in the same figure, this correlation is not significant for any of the familiar food pictures. Although to a lesser degree than for valence, self-reported arousal in response to

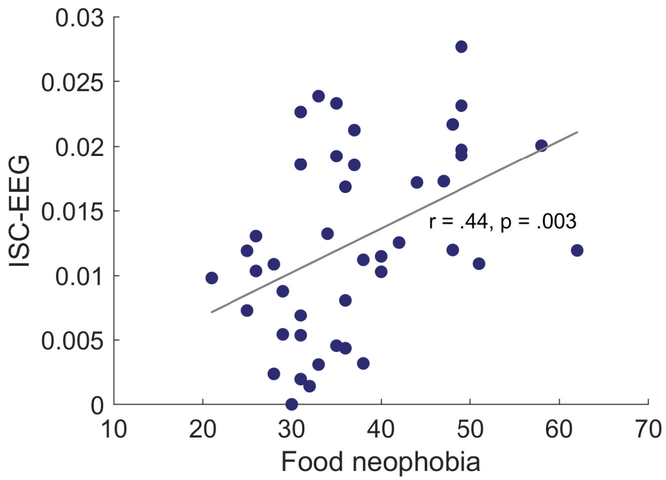

Figure 4-5. Correlation between food neophobia and ISC-EEG during presentation of the movie on the origin and production of an unfamiliar food. Each data point represents a participant.

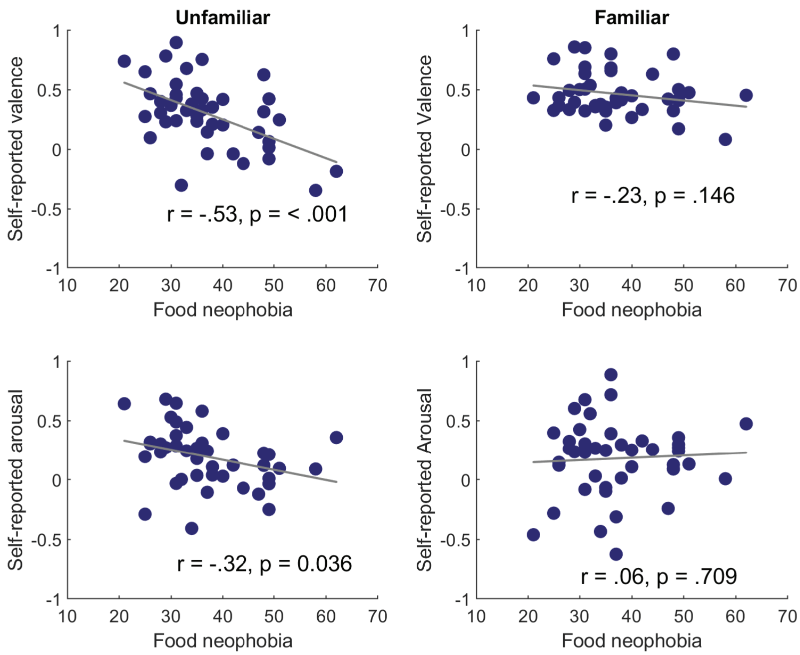

Figure 4-6. Correlations between food neophobia and self-reported valence (top) and arousal (bottom) for unfamiliar (left) and familiar (right) food pictures. Each data point represents a participant.

unfamiliar food pictures is also significantly correlated with food neophobia (r = -0.32, p = .036), where high food neophobia is associated with low arousal scores. Again, no significant correlations are found for familiar food pictures.

4.3.4. Sip size

Figure 4-7 shows significant correlations between food neophobia and average sip size for all soups (unfamiliar sumashi: r = -0.47, p = .002; unfamiliar miso: r = -0.44, p = .004; familiar vegetable: r = -0.15, p = .336; familiar tomato: r = -0.33, p = .033), where food neophobia is associated with smaller sip size, except the familiar vegetable soup.

4.3.5. EmojiGrid after soup tasting

Figure 4-8 shows a significant negative correlation between food neophobia and self-reported valence after tasting the unfamiliar miso soup (r = -0.31, p = .047) and a near significant negative correlation after tasting the unfamiliar sumashi soup (r = -0.28, p = .069). No significant correlations with self-reported valence are found after tasting familiar soups. No significant correlations between food neophobia and self-reported arousal are found for any of the soups.

4

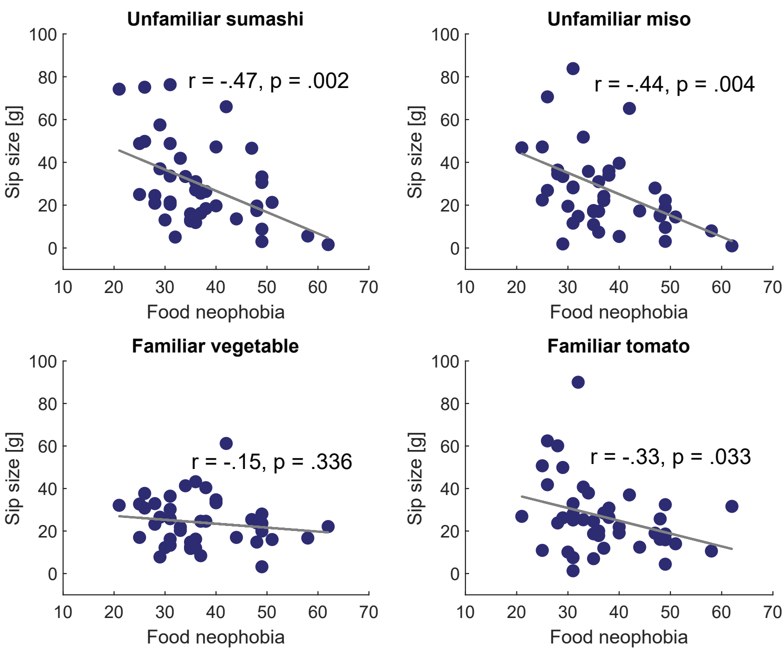

Figure 4-7. Correlation between food neophobia and sip size for unfamiliar sumashi and miso soups and familiar vegetable and tomato soups and. Each data point represents a participant.

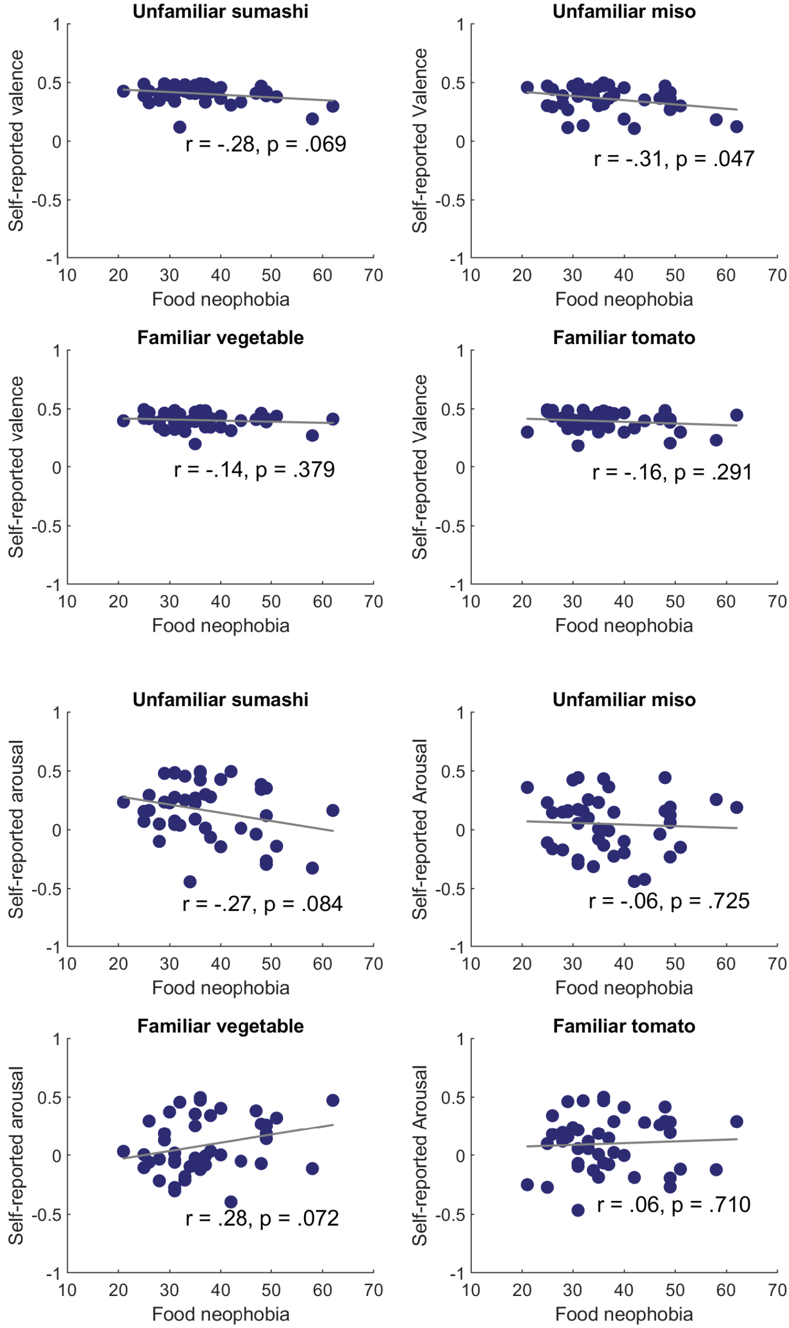

Figure 4-8. Correlations between food neophobia and self-reported valence (top) and arousal (bottom) for familiar tomato and vegetable soups and unfamiliar miso and sumashi soups. Each data point represents a participant.

4.4. Discussion

In the current work we investigated how food neophobia is associated with attention toward food-related stimuli as elucidated by two EEG measures: ERPs and ISC-EEG. We also examined the association between food neophobia and self-reported emotional experience, as well as the association between food neophobia and an implicit behavioral response (sip size). Below, the results will first be discussed for each of the hypotheses, followed by overall conclusions.

4.4.1. Hypothesis 1: LPP ERP amplitude

As hypothesized, the LPP amplitude in response to unfamiliar food pictures correlated with food neophobia, where participants with higher food neophobia showed higher LPP amplitude. However, nearly identical results were found for familiar food pictures. A higher LPP amplitude indicates increased attention to, and facilitated processing of motivationally relevant stimuli (Schupp et al., 2000; Schupp et al., 2003). It appears that, unlike its name suggests, food neophobia does not only have an effect on the attentional processing of novel foods, but on any food-related stimulus.

Comparable observations have occurred in other phobic populations. Individuals with social anxiety do not only show increased LPPs to threatening faces, but also to faces overall, regardless of valence (Mühlberger et al., 2009). The authors conclude that faces in general are important stimuli for socially anxious people (Moser et al., 2008). Similarly, our results indicate that not only novel foods, but all food-related stimuli are of high importance to food neophobic individuals. Previous work showed higher arousal in food neophobic individuals than food neophiliacs when presented with successively presented pictures of a range of food types, as indicated by increased heart rate, EDA and respiration rate (Raudenbush and Capiola, 2012). We now show specifically for different types of food images that this increased arousal is likely associated with increased attention to food pictures in general.The current study did not include non-food stimuli. Including such stimuli would have allowed us to examine the specificity of the relation between food neophobia and responses to food, rather than to stimuli more in general. As discussed in the introduction, food neophobia may be part of wider avoidance behavior of novel stimuli (Pliner and Hobden, 1992; B. Raudenbush and Frank, 1999), food neophobia and anxiety are positively associated (Galloway et al., 2003; Pliner and Hobden, 1992) and anxiety traits can cause attentional biases toward threatening pictures compared to nonthreatening pictures (Berggren and Eimer, 2021; Holmes et al., 2008; Li et al., 2005). Therefore, we would expect that higher LPP amplitudes and stronger ISC-EEG in individuals scoring high on food neophobia might be found for other, non-food stimuli as well.

4.4.2. Hypothesis 2: ISC-EEG

As for the LPP amplitude, ISC-EEG showed the hypothesized positive correlation with food neophobia. As higher ISC-EEG indicate higher levels of attention toward the presented stimulus (Stuldreher et al., 2020b), this result indicates that participants with higher food neophobia also show higher levels of attention directed to a naturalistic, narrative stimulus about the origin of an unfamiliar food product.

Several authors have argued that ISC-EEG is likely driven for a large part by consecutive ERPs occurring from attentionally relevant or emotionally salient events in the stimulus (Poulsen et al., 2017; Stuldreher et al., 2020a). Our current findings that both ERP LPP amplitude and ISC-EEG correlate with food neophobia are in line with this reasoning. Participants with high food neophobia likely showed higher and more consistent ERPs in response to food-related events throughout the movie, leading to higher overall ISC-EEG.

Compared to ERPs, requiring the controlled presentation of successive stimuli, assessing ISC-EEG enables the use of much more naturalistic and continuous stimuli, such as video clips (Cohen et al., 2018; Dmochowski et al., 2014) or audiobooks (Stuldreher et al.., 2020b). Up to now, ISC-EEG was shown to be modulated by explicit instructions to focus attention on specific stimulus aspects (Ki et al., 2016; Stuldreher et al., 2020b). To the best of our knowledge, we here show for the first time that ISC-EEG during such naturalistic stimuli is modulated by personal trait. This is an important development for the use of ISC-EEG as implicit measure of attentional processing in natural environments, where one usually does not aim to capture variations in attention due to explicit instructions, but due to natural variations, for instance related to personal trait.

4.4.3. Self-reported experienced emotion for images

To examine how the early attentional processes culminate in food experience we also obtained self-reported emotional experience. The more pronounced allocation of attentional resources of high food neophobic individuals to all food related stimuli did not result in different reported emotional experience for all food related stimuli.

Whereas food neophobia was positively associated with LPP amplitude for all types of food pictures, food neophobia was only associated with self-reported experienced emotion for unfamiliar food pictures, not for any of the other picture types. Specifically, individuals with higher food neophobia reported a lower valence after being presented with pictures of unfamiliar food, in line with previous reports (Brouwer et al., 2021). These individuals also reported lower arousal, as can be expected for food images scores that are shifted from pleasant to neutral, given the U-shaped relation between reports of valence and arousal (Kaneko et al., 2019a; Kuppens et al., 2013).

In contrast to the short-term attention allocation as captured by the ERPs, self-reported food-evoked emotions only reflected food neophobia after presentation of unfamiliar food pictures. Though individuals with higher food neophobia thus allocate more attentional resources to any food stimulus, only for the pictures of the unfamiliar foods they report different food-evoked emotions than individuals with lower food neophobia.

Note that the pictures of food were presented with either a Japanese or Dutch flag indicating the origin of the food. It might be that these flags also primed the expectations of participants.

4.4.4. Sip size and self-reported experienced emotion for soups

Similar to LPP amplitude, sip size reflected food neophobia for both familiar (tomato) and unfamiliar (miso and sumashi) soups - food neophobic individuals tend to take smaller sips of soups. However, this association appeared stronger for the unfamiliar soups, and did not reach significance for the familiar vegetable soup.

This association between behavior and food neophobia is in line with a study by (Raudenbush et al., 2003). They reported that severe food neophobic and neophiliac individuals differed in salivary response when presented with food items, regardless of the familiarity with the food item. For participants with average food neophobia, salivary response was dependent on the familiarity with the presented food items.

Just as with picture presentation, self-reported emotional experience of the soups only varies with food neophobia for the unfamiliar food stimulus. Food neophobia scores correlated negatively with self-reported valence for unfamiliar soups, but not for familiar soups.

These results indicate that implicit measures, taken before further evaluation processes took place and outside awareness, can be more sensitive to detect differences in food experience. (Kaneko et al., 2019a) also found that sip size distinguished between ground truth disliked and liked drinks better than self-reported experience.

4.4.5. Conclusions

Using different types of measures and stimuli, we here examined and described how early attentional, behavioral and emotional processing of familiar and unfamiliar food stimuli vary with food neophobia. Food neophobia seems to affect the short term attentional processing of all food-related stimuli, regardless of familiarity. Individuals with high food neophobia allocate more attentional resources for the processing of food-related stimuli than individuals with low food neophobia. When tasting, sip size was also found to covary with food neophobia for both familiar and unfamiliar stimuli, even though the association appeared stronger for unfamiliar stimuli. The differences in attentional resource allocation between individuals with varying food neophobia did not result in different emotional experience for all stimuli. After initial attention allocation, the presented stimulus is identified and evaluated in more detail. The result of this appraisal that is accessible for conscious awareness, i.e. the self-reported food experience, only covaried with food neophobia after presentation of unfamiliar food stimuli, both for viewing pictures and for tasting soups.

Taken together, these results indicate that cognitive, self-reported instruments to assess food experience do not capture the entire food experience. Based on those results, one would conclude that food neophobia only affects the experiences of novel food, whereas the implicit behavioral and attentional processing are affected for all food types, regardless of familiarity. One can therefore also argue that the food neophobia scale, unlike the name suggests, does not capture the fear of novel foods, but the fear of food experience in general.

All in all, our work revealed that there is a more profound difference between food neophobic and food neophiliac individuals than only the appreciation of novel food.

Acknowledgments

The authors would like to thank Jasper van Beers and Priya Sabu for their help with setting up and running the experiment.

5.

Physiological synchrony in electrodermal activity predicts decreased vigilant attention

Stuldreher, I.V., Maasland, E., Bottenheft, C., van Erp, J.B.F., Brouwer, A.M. (2023). Physiological synchrony in electrodermal activity predicts decreased vigilant attention induced by sleep deprivation. *Frontiers in Neuroergonomics*, ***4***, 1199347. doi: 10.3389/fnrgo.2023.1199347

Abstract

Introduction: When multiple individuals are presented with narrative movie or audio clips, their electrodermal activity (EDA) and heart rate show significant similarities. Higher levels of such inter-subject physiological synchrony are related with higher levels of attention towards the narrative, as for instance expressed by more correctly answered questions about the narrative. We here investigate whether physiological synchrony in EDA and heart rate during watching of movie clips predicts performance on a subsequent vigilant attention task among participants exposed to a night of total sleep deprivation.

Methods: We recorded EDA and heart rate of 54 participants during a night of total sleep deprivation. Every hour from 22:00 to 07:00 participants watched a 10-minute movie clip during which we computed inter-subject physiological synchrony. Afterwards, they answered questions about the movie and performed the psychomotor vigilance task (PVT) to capture attentional performance.

Results: We replicated findings that inter-subject correlations in EDA and heart rate predicted the number of correct answers on questions about the movie clips. Furthermore, we found that inter-subject correlations in EDA, but not in heart rate, predicted PVT performance. Individuals' mean EDA and heart rate also predicted their PVT performance. For EDA, inter-subject correlations explained more variance of PVT performance than individuals' mean EDA.

Discussion: Together, these findings confirm the association between physiological synchrony and attention. Physiological synchrony in EDA does not only capture the attentional processing during the time that it is determined, but also proves valuable for capturing more general changes in the attentional state of monitored individuals.

5.1. Introduction

When individuals are engaged in social interaction, their physiological signals can align (Palumbo et al., 2017). This similarity in physiological activity across individuals is referred to as physiological synchrony. Physiological synchrony has been shown for diverse groups of socially interacting individuals, such as parent-child dyads (Feldman et al., 2011), therapist-patient dyads (Marci et al., 2007), teammates (Elkins et al., 2009) and pairs of strangers meeting for the first time (Silver and Parente, 2004).

We think that findings of physiological synchrony may be at least partially explained by shared attention. Therefore, we argue that physiological synchrony can be a valuable tool to study shared attention in groups of individuals. In recent years physiological synchrony was indeed shown to reflect shared attention. When individuals attend to a joint narrative stimulus, the inter-subject correlations in the electroencephalogram (EEG), electrodermal activity (EDA) and heart rate are higher than one would expect based on chance (Poulsen et al., 2017; Stuldreher et al., 2023a). Altering the attentional focus of individuals was found to affect inter-subject correlations. Individuals showed decreased physiological synchrony when instructed to focus attention inward on a mental arithmetic task instead of the joint stimulus (Ki et al., 2016; Pérez et al., 2021). Individuals instructed to focus only on specific stimulus parts showed higher inter-subject correlations with individuals instructed to focus on the same instead of on different stimulus parts (Stuldreher et al., 2020b). Individuals with higher participant-to-group inter-subject correlations in EEG or heart rate better recalled the narrative (Cohen and Parra, 2016; Stuldreher et al., 2020b). In sum, physiological synchrony can be altered by interventions on attention and has been found to be associated with performance on a task that reflects how well a presented narrative was attended to.

There are also indications that inter-subject correlations in EEG and functional magnetic resonance imaging can capture interpersonal variations in attentional processing that relate to personality traits. For instance, we found that food neophobia, the hesitance to try new foods, was positively correlated with inter-subject correlations in EEG during a movie about a foreign food (Stuldreher et al., 2023b). That is, individuals who scored higher on the food neophobia scale showed less inter-subject correlations in EEG. In addition, individuals with autism spectrum disorders, depression or first-episode psychosis, are known to show more varying neural patterns and thus reduced neural inter-subject correlations during naturalistic stimuli than typically developing individuals (Hasson et al., 2009; Salmi et al., 2013; Guo et al., 2015; Mäntylä et al., 2018).

The indications that variations in attentional processing within an individual

5

can also be captured by inter-subject correlations are limited. It is established that inter-subject correlations in EEG decrease when viewing the same stimulus for the second time (Dmochowski et al., 2012; Ki et al., 2016), consistent with participants being less interested in the stimulus upon a second viewing. We may expect that inter-subject correlations as established during a short narrative can capture the momentary attentional state of an individual, and therefore predict performance on subsequent attentional tasks. However, inter-subject correlations have never been related to performance on attentional tasks separate from the narrative. As of yet, it is not clear whether inter-subject correlations monitored during a narrative presentation can capture variations in attentional processing abilities as also reflected in another task.

In the current work, we manipulate the attentional abilities within individuals over time by exposing them to a night of total sleep deprivation. During the night, we monitor inter-subject correlations in EDA and heart rate during the presentation of movie clips. Although work on inter-subject correlation as measure of attention originated in measures of brain activity such as functional magnetic resonance imaging (e.g. Hasson et al., 2004, 2008) and EEG (e.g. Dmochowski et al., 2012), we and others found that inter-subject correlations in body measures such as EDA and heart rate reflect attention as well (Stuldreher et al., 2020b, 2020a; Pérez et al., 2021; Madsen and Parra, 2022). This also holds when using measurements from wearables (Van Beers et al., 2020).

Sleep deprivation is known to have strong detrimental effects on attention, working memory and decision making (Alhola and Polo-Kantola, 2007; Pilcher et al., 2007). It especially impacts cognitive functioning in long, simple and monotonous tasks requiring reaction speed or vigilance (Alhola and Polo-Kantola, 2007; Lim and Dinges, 2008; Hudson et al., 2020). Pilcher et al., 2007 provide a reasoning for the especially strong effects of sleep deprivation on undemanding tasks through their model of attentional control. Undemanding cognitive tasks, like vigilance tasks, require more internal control over one's attention, since there are less external stimuli to keep one engaged. Sleep deprived individuals are thought to have difficulty exerting this internal control over their attention. We investigate whether inter-subject correlations in EDA and heart rate can predict performance on a task demanding vigilant attention, i.e., a task that is expected to be strongly affected by sleep deprivation. A positive result in this study would indicate that physiological synchrony may indeed capture the varying attentional abilities of individuals and would extend the predictive value of physiological synchrony on attentional performance to moments beyond the time that physiological synchrony was determined. Instead of investigating whether inter-subject correlations vary between discrete attentional conditions, this study allows us to explore the predictive value of inter-subject correlations on differences in attention that are of more continuous nature.

We compare the predictive value of inter-subject correlations in EDA and heart rate to the predictive value of individuals' mean EDA and heart rate. EDA and heart rate on individual level are known to be affected by sleep deprivation, and have been associated with cognitive performance during a sleep deprived night before (Miró et al., 2002; Posada-Quintero et al., 2017). Directly comparing the predictive value of inter-subject correlations with the individual physiological features allows us to display potential added value of inter-subject analyses.

In sum, we aim to answer the following main research question:

1. How well does physiological synchrony in EDA and heart rate predict later vigilant attention during a night of sleep deprivation?

Additionally, we aim to answer the following secondary research question:

2. How does the predictive value of physiological synchrony relate to predictive value of individual's mean physiological activity?

5.2. Materials and methods

5.2.1. Participants

101 Dutch speaking volunteers took part in a two-day study on the effects of sleep deprivation on cognitive performance. They were randomly assigned to the sleep deprivation condition (N = 54) or control condition (N = 48). This study describes results obtained from participants in sleep deprivation condition during the night. These 54 participants (29 female) were between 18 and 55 years old (M = 29.4, SD = 11.9). The study was approved by the Medical Research Ethics Committee (MREC) Brabant (reference number P2045, approval number NL74961.028.20). All participants provided written informed consent before partaking in the experiment. Exclusion criteria were: smoking, drug use in the last three months, signs of flue or viral infection in the last ten days, pregnant, history of psychiatric illness, including sleep disorders, autoimmune disease and/or hyperactive thyroid and people with known heart, kidney or liver disease or neurological complaints.

5.2.2. Physiological measurements

Participants' EDA and heart rate were recorded throughout the night. EDA was recorded at 32 Hz with an EdaMove 4 (Movisens GmbH, Karlsruhe, Germany), that was recording signals from the palmar surface of the non-dominant hand using two solid gelled Ag/AgCl electrodes (MTG IMIELLA electrode, MTG Medizintechnik, Lugau, Germany, W55 SG, textured fleece electrodes, 55 mm diameter). Heart rate was recorded at 1 Hz and with 1 bpm resolution using a Tickr chest-strap (Wahoo Fitness, Atlanta, GA, USA) that was coupled to an Android smartphone (Samsung Galaxy A41, Android version 10) via Bluetooth. Data were received, processed and saved with the use of the Wahoo Fitness Workout Application (version 1.40.0.56).

5.2.3. Stimuli

5.2.3.1. Movie clips

Over the course of the night, participants were presented with ten 10-minute movie clips (M = 9:56, min. = 9:04, max. = 10:58) every hour from 22:00 to 07:00. The movie clips were selected from the Dutch YouTube channels NPO3 and KORT! and featured short, moderately engaging stories. In a previous study using six of these movies we showed that they were effective in eliciting significant inter-subject correlations in EDA and heart rate (Stuldreher et al., 2023a). The order of presentation was the same for each participant. This presentation order, movie duration and URL for each movie can be found in Table 5-1. Directly after each movie, participants answered ten multiple-choice questions (four answers of which one correct). The questions concerned both general and specific details of the movie. The questions and answer options can be found in the Supplementary Table 1.

5.2.3.2. Psychomotor vigilance task (PVT)

After every sequence of watching a movie and answering the questions, participants performed the psychomotor vigilance task (PVT). The experimental software that presented the PVT was custom-made for this experiment using Python (version 3.8) and the compatible PsychoPy toolbox (Peirce et al., 2019).

Table 5-1. Details and order of the presented movie clips

Order	Name	Duration (min)	URL
1	Chauffeur	09:45	https://www.youtube.com/watch?v=jaFmvyH7dW8
2	El Mourabbi	09:04	https://www.youtube.com/watch?v=X9bJou2gKxo
3	De Chinese Muur	09:50	https://www.youtube.com/watch?v=yjGFuhPy3Qo
4	One of the boys	10:58	https://www.youtube.com/watch?v=PsGAuhgQ97k
5	Samual	09:45	https://www.youtube.com/watch?v=VUseoqCVnj4
6	Turn it around	09:26	https://www.youtube.com/watch?v=beC7IpQpTz4
7	En route	10:05	https://www.youtube.com/watch?v=M6ebApnH_XE
8	Mowgli en Fidel	10:03	https://www.youtube.com/watch?v=MocrSQW_r_M
9	Heen en weer dag	10:25	https://www.youtube.com/watch?v=yPueHzj9STE
10	Gutmensch	10:02	https://www.youtube.com/watch?v=P7MABwTYa58

The PVT was designed to be maximally sensitive to the effects of sleep deprivation by following recommendations from (Basner and Dinges, (2011). The course of the PVT was as follows. After pressing the spacebar to initialize the task, a black square (9.5 degrees of visual angle) surrounded by a red border centered in the middle of a grey screen was shown. For each trial, a random interval between 2-10 seconds was selected, after which a counter appeared in the center of the black square. The counter consisted of three yellow digits displaying the milliseconds since the onset of the counter. The participants' task was to press spacebar as soon as possible after appearance of the counter. As soon as they did, the reaction time was displayed for two seconds. If participants failed to press the spacebar within 3 seconds of the counter appearing or if participants pressed the spacebar without a counter being present, a yellow 'X' appeared in the black square instead of the digits. This error message stayed present for 2 seconds, after which the black square was cleared again and a new trial started. The PVT lasted 10 minutes, and the amount of trials varied based on the reaction times and misses of each participant.

5.2.4. Procedure

Data from participants was collected as part of a larger study over the course of seven experimental sessions conducted in 2021 between March and June. Each experimental session consisted of two morning sessions that are discussed in more detail in (Bottenheft et al., 2023), and the night spent at the research institute between the two mornings in which the measurements for this study were performed. In each experimental session on average 7.7 participants (SD = 1.48) participated. Participants joined a training session two weeks before the experimental session. During the training session, participants were familiarized with tasks they would have to perform during the course of the experiment. Of relevance for the current study is that the participants familiarized themselves with the PVT for five minutes. Additionally, they were taught how to self-apply the Tickr chest strap and EdaMove 4 (a safety measure taken to reduce the risk of spreading the COVID-19 virus). Furthermore, participants received instructions not to consume any caffeinated substance after 18:00 on the day of their experimental session and not to consume any alcohol in the 24 hours prior to the experiment.

On the day of the night session, participants came to the research institute at 21:00. Figure 5-1 describes the full procedure of the night session. Participants were led to the cafeteria of the research institute, where they spent most of the time during the night. Participants were instructed to apply the Wahoo Tickr and start a new measuring session in the Wahoo Fitness Workout Tracker application on the smart phone they had received for the duration of the study.

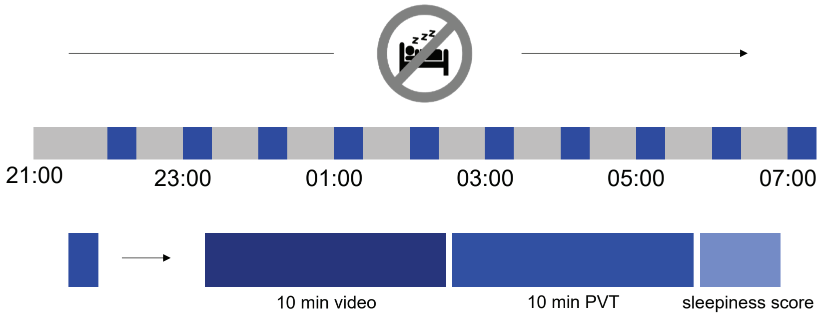

Figure 5-1. The experiment consisted of 10 blocks of a 10 minute movie clip, followed by questions about the movie, a 10 minute vigilance task and the Stanford sleepiness scale. These blocks started each hour from 22:00 to 07:00 o'clock.

Participants also applied the EdaMove 4 and corresponding electrodes. A researcher checked whether the data recording of both devices was running. After this procedure, participants were given a 30 minute rest period. During this and other resting periods, participants were allowed to move around freely and talk to one another or entertain themselves with a book, game or electronic device they brought from home. They were not allowed to exercise. If a participant fell asleep the researcher woke the participant.

During the subsequent time of the experimental session, participants followed a standard procedure every hour, depicted in Figure 5-1. The procedure consisted of the following elements: watching a 10-minute movie clip and answering ten questions about the presented movie, performing a 10-minute PVT and reporting their sleepiness on the Stanford Sleepiness Scale (SSS). After they had completed the procedure (about 30 minutes), participants were allowed to rest until the initiation of the next cycle. The first cycle was initiated at 22:00, and every subsequent cycle was initiated on the following hour. The last cycle was performed at 07:00. At 23:30, 01:30 and 04:30, participants received a snack. They received 150 grams of full fat quark, 150 grams of fresh vegetables with hummus or 30 grams of walnuts and could pick which snack they wanted at which point in time. They were allowed to drink water and theine-free teas.

5.2.5. Analysis

Data were analyzed using MATLAB R2021a (Mathworks, Natick, MA, USA). The data and MATLAB scripts used are available online at https://osf.io/69u8h/.

To be able to directly compare the predictive results of inter-subject correlations in EDA and heart rate, we used only data of participants for which EDA and heart rate recordings were available. Heart rate data of six participants were unavailable due to lost data recordings. EDA data of an additional two participants were lost due to sensor failure. For 10 additional participants, EDA or heart rate data during part of the recording blocks were lost due to detached electrodes or disconnected Bluetooth connection. In total 21 recording blocks were lost across these participants. In sum, the following analyses are conducted using 46 participants, with in total 439 recording blocks.

5.2.5.1. PVT analysis

PVT response times were processed to obtain the lapse probability. A PVT response was considered valid if given between 100 ms and 3000 ms after stimulus onsets. Responses within 100 ms of stimulus onset and responses without a stimulus being present were considered false alarms. Responses that occurred more than 500 ms after stimulus onset were considered lapses (Basner and Dinges, 2011). The lapse probability was computed by dividing the amount of lapses with the amount of correct responses excluding lapses. Invalid responses are thus not considered in the lapse probability.

5.2.5.2. EDA and heart rate

EDA was resampled to 8 Hz. EDA and heart rate were epoched to the on- and offset of each movie using markers sent by the experimental program. EDA was then further processed to obtain the phasic component, also called skin conductance response (SCR), as this component of EDA is characterized by fast consecutive response-like changes. To do so, we used continuous decomposition analysis as implemented in the Ledalab toolbox for MATLAB (Benedek and Kaernbach, 2010). We use the SCR, as external stimuli mainly affect this component of EDA.

The average SCR and heart rate over each epoch were used as input for the predictive models that use individuals' physiological signals as predictor, as described below. The time-course of the SCR and heart rate were used for computation of inter-subject correlations. In the following parts of the manuscript, when we use the term EDA we refer to the SCR component of the electrodermal signal.

5.2.5.3. Inter-subject correlations

Physiological synchrony was quantified with the use of inter-subject correlations. For each participant's EDA and heart rate during a given movie, inter-subject correlations were computed with the EDA and heart rate of every other participant during that same movie, following our previous procedure (Stuldreher et al., 2023a). When averaging over the inter-subject correlations with all other participants, we obtain a metric we refer to as participant-to-group inter-subject correlation. This metric is used as the predictor variable. From here on, if we refer to inter-subject correlation we refer to this participant-to-group metric.

To test the significance of participant-to-group inter-subject correlation values over chance level, we used the circular shuffle statistic, following (Pérez et al., 2021; Madsen and Parra, 2022). Each participant's physiological signal during each movie was circular shifted by a random amount within the epoch length. The inter-subject correlations and participant-to-group inter-subject correlations were then computed with this circular shuffled data. This procedure was repeated 500 times for each participant and each movie to estimate the chance distribution of inter-subject correlations. Note that a higher number of shuffles (e.g., 10.000 as in Pérez et al., 2021; Madsen and Parra, 2022) would result in a better estimation of the chance distribution. The p-value then is the fraction of circular shuffles with inter-subject correlation values higher than the original unshuffled inter-subject correlations.

As an additional check, we computed participant-to-group inter-subject correlation values using data from non-matching movies to compare the real value to. That is, instead of computing inter-subject correlations of participant x with participant y using data from the same movie m_1, we computed inter-subject correlations between data of participant x from movie m_1 with data of participant y from all movie clips m_2 in M_2, where M_2 are all videos not equal to m_1. For the most stringent test, we then selected the maximum inter-subject correlation value to compare to the real inter-subject correlation values. Comparisons were done using paired sample t-tests.

Note that we do not necessarily expect inter-subject correlations higher than chance level for all times at night. We expect low, or even absent inter-subject correlations if participants do not attend to the movies because of sleepiness.

5.2.5.4. Hierarchical linear analysis

We used an hierarchical linear model (HLM) for the main statistical analysis. An HLM was selected since the data is organized hierarchically and clustered within individuals. Therefore, the assumption of independent observations required for a regular regression is violated.

We performed the hierarchical linear regression in four steps, shown in Table 5-2. In the first step, that serves as a baseline, only the dependent variable is added to the model. In the second step, the level two variable individual is added to the model, to investigate if allowing the means of the dependent variable to vary across individuals leads to improvement of the model. In the third step, the predictor is added with a fixed slope and random intercept. This step allowed us to investigate whether the predictor has a main predictive effect on the dependent variable. In the fourth and final step, the predictor variable is added with a random slope and random intercept, meaning that the model

Table 5-2. Steps of the hierarchical linear model analysis

Step	Formula
1	$V_{dep} \sim 1$
2	$V_{dep} \sim 1 + (1 \mid \text{participant})$
3	$V_{dep} \sim p_1 + (1 \mid \text{participant})$
4	$V_{dep} \sim p_1 + (p_1 \mid \text{participant})$

V_{dep}: dependent variable, either log transformed PVT lapse probability scores or number of correctly answered on questions about the movies, p_1: predictor variable 1, either inter-subject correlations in EDA or heart rate, mean EDA or heart rate

allows the relationship between the predictor and the dependent variable to vary between individuals.

We first built models with number of correct answers on questions about the movie as dependent variable, individual as level two variable and inter-subject correlations in EDA or heart rate as predictor. This was done to investigate whether the previously found relation between narrative retention and inter-subject correlations could be replicated in the current setting.

We then built models with PVT lapse probability as a dependent variable, individual as a level two variable and inter-subject correlations in EDA or heart rate as predictor. This analysis was conducted to answer the main research question, to explore the potential association between inter-subject correlation and performance on a subsequent task requiring a different form of attention.

Then, we used individuals' mean EDA or heart rate as predictor of PVT lapse probability. This was done to compare the predictive value of inter-subject correlations to individuals' physiological activation.

Before performing the HLM analyses, inter-subject correlation scores were centered using grand-mean sampling to avoid collinearity problems. PVT lapse probability was log transformed as untransformed data were positively skewed.

5.3. Results

5.3.1. Significance of inter-subject correlations as measure of attentional engagement

Before exploring potential associations between inter-subject correlations and attention, we establish to what extent the inter-subject correlations are higher than one would expect based on chance using two approaches. First, we investigated whether real inter-subject correlations were higher than inter-subject correlations obtained from circular shuffled data. Figure 5-2 depicts the inter-subject correlations in EDA and heart rate for each individual and each movie clip ordered from low to high. In each panel, each marker refers to a single participant. Filled markers depict inter-subject correlations that are significantly higher than chance level, open markers depict inter-subject correlations that are not significantly higher than chance level. For heart rate, during each movie on average 73.2% (SD: 12.3%) of the participants show significant inter-subject correlations. In EDA, during each movie on average 36.6% (SD: 18.8%) of the participants show significant inter-subject correlations. Supplementary Figure 1 shows the same inter-subject correlations and compares these to the chance level distributions obtained through 500 instances of circular shuffling. We then investigated whether real inter-subject correlations where higher than inter-subject correlations obtained when movies did not match between participants. Figure 5-3 depicts the inter-subject correlations in EDA and heart rate for each individual and each movie and compares these to inter-subject correlations obtained after a permutation in which non-matching movies were used. In each panel each line refers to data of a single participant. Blue lines depict participants where the permuted inter-subject correlations are lower than the original values, orange lines depict participants where the permuted inter-subject correlations are higher than the original values. A set of paired sample t-tests showed that the real inter-subject correlation values were significantly higher than the permuted ones in most cases. The t-test statistics can be found in Supplementary Table 2.

5.3.2. Inter-subject correlations as predictor of number of correctly answered questions about the movie.

Our next step was to check whether we could replicate previous reports of an association between inter-subject correlations and the number of correctly answered questions about the presented movie clips. Figure 5-4 shows how the number of correct answers varies over the course of the night. It suggests a slight drop in the number of correct answers up to 05:00 o'clock, followed by an increase up to 07:00 o'clock. Table 5-3 shows the statistical parameters of the HLM analysis used to identify a potential association between inter-subject correlations and the number of correct answers. For each step, the table displays the minus 2 Log Likelihood statistic (-2LL), degrees of freedom (DF) and the Akaike Information Criterion (AIC). The -2LL is a statistic that can be used to compare whether a model is a significant improvement of another model (Field, 2013). The AIC is a more complex measure that can also be used to compare models. The model that best fits the data will display lowest AIC. The table also shows the Chi-square (χ^2) -2LL change statistic, that describes whether the change in -2LL is significantly different from the previous model. Also displayed

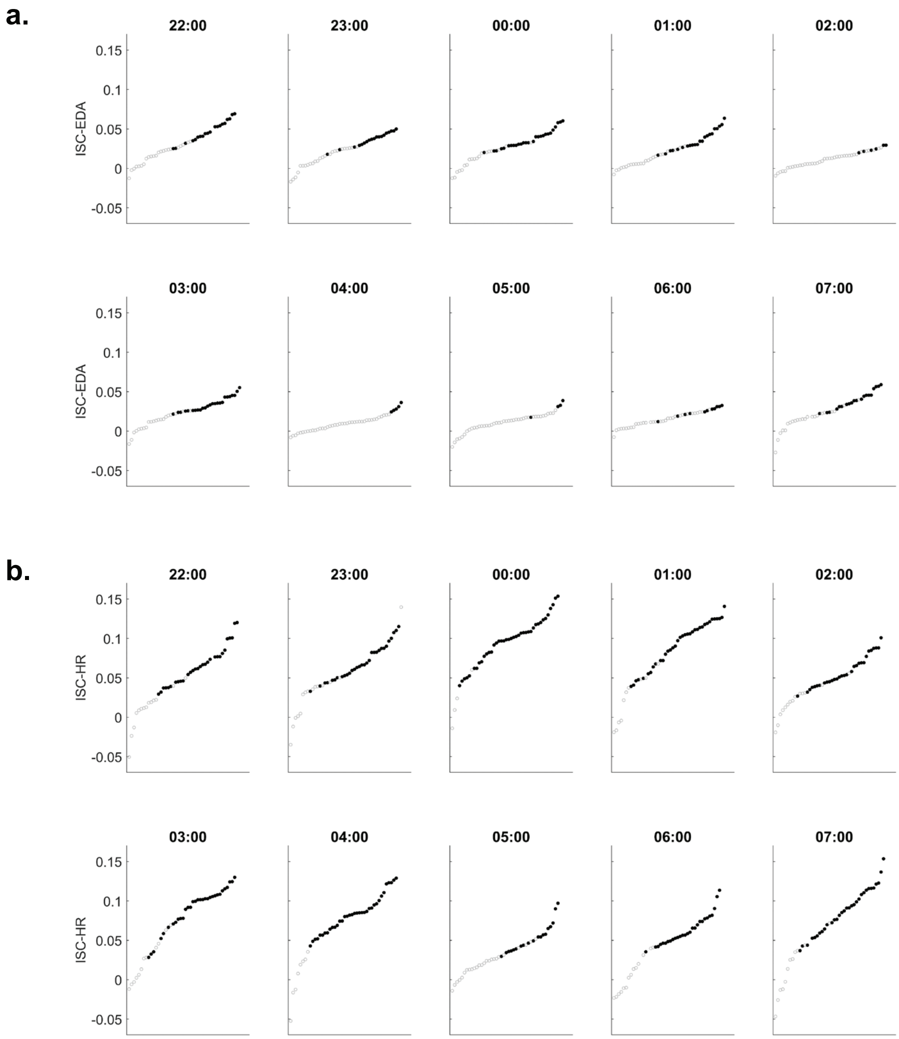

Figure 5-2. Inter-subject correlations in a. EDA (ISC-EDA) and b. heart rate (ISC-HR) for each individual and each movie clip ordered from low to high. Each marker refers to the participant-to-group inter-subject correlations of a participant. Filled markers depict inter-subject correlations significantly higher than chance, open markers depict inter-subject correlations not higher than chance level.

is the percentage of total explained variance and the -statistics of the fixed effects.

In step one, only the dependent variable (i.e., the number of correct answers) was added to the model to serve as a baseline. The AIC is in this model is 1593.5.

In step two, the level two variable participant was added to the model to in-

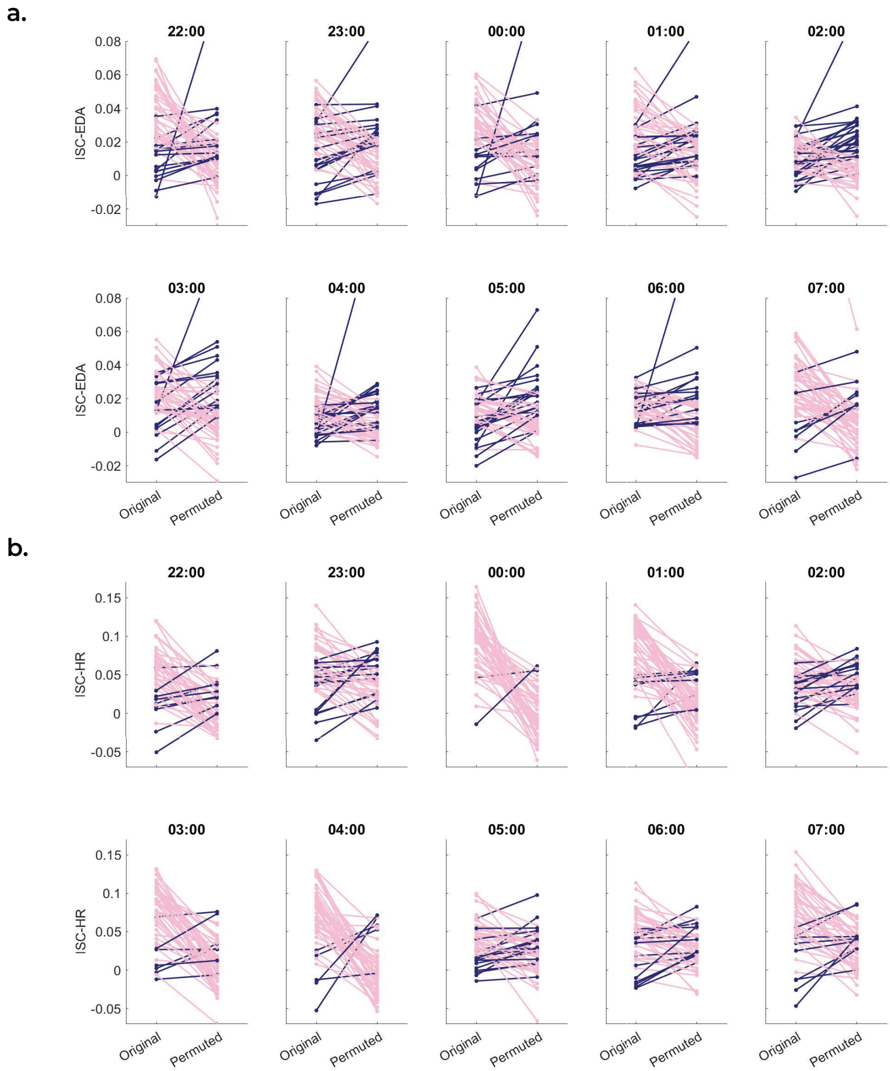

Figure 5-3. Inter-subject correlations in a. EDA (ISC-EDA) and b. heart rate (ISC-HR) for each individual and each movie compared to a permuted inter-subject correlations obtained by not matching the movies between participants. Pink lines depict participants where the permuted inter-subject correlations are lower than the original values, blue lines depict participants where the permuted inter-subject correlations are higher than the original values.

vestigate whether allowing for individual differences improves the model prediction. The AIC is in this model is 1595.5, and the model did not significantly improve compared to step one, χ^2 (1, N = 439) = 0, p = 1.

In step three, the predictor inter-subject correlations in either EDA (step 3a) or heart rate (step 3b) was added to the model with fixed slope and random in-

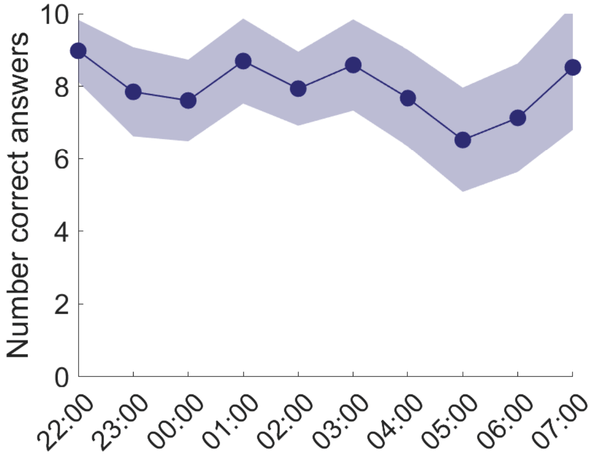

Figure 5-4. Number of correct answers to questions about the content of the movie over the course of the night. Markers depict the mean across participants, shaded area depicts the standard deviation.

Table 5-3. Statistical parameters of the two hierarchical linear models using inter-subject correlations in either EDA (ISC-EDA) or heart rate (ISC-HR) as predictor of the number of correct answers about the content of the movies.

Step	Predictor	-2LL	DF	AIC	χ^2 -2LL change	R^2	t fixed effect
1		-794.76	2	1593.2			
2		-794.76	3	1595.5	(1, N = 439) = 0, p = 1	0	112.48
3a	ISC-EDA	-791.89	4	1591.8	(1, N = 439) = 5.75, p = .016	.013	2.41*
3b	ISC-HR	-783.74	4	1575.5	(1,N = 439) = 22.04, p < .001	.049	4.75***
4a	ISC-EDA	-783.82	6	1579.6	(2, N = 439) = 16.15, p < .001	.084	2.70**
4b	ISC-HR	-779.5	6	1571.1	(2, N = 439) = 8.40, p = .015	.088	4.03***

* p < .05, ** p < .01, *** p < .001

Abbreviations: -2LL: = minus 2 Log Likelihood statistic; DF = degrees of freedom; AIC = Akaike Information Criterion

tercept. Both for EDA, χ^2 (1, N = 439) = 5.75, p = .016, and heart rate, χ^2 (1, N = 439) = 22.04, p < .001, the prediction significantly improved compared to model step 2. Inter-subject correlations thus have a main predictive effect of the number of correctly answered questions about the movie clips, both for EDA (β = .17, p = .017) and heart rate (β = .33, p < .001).

In step four, the predictor was added with random slope and random intercept, meaning that the model allows the relationship between the predictor and dependent variable to vary across participants. The prediction significantly improved compared to step three, both for EDA (step 4a), χ^2 (2, N = 439) = 16.15, p < .001, and heart rate (step 4b), χ^2 (2, N = 439) = 8.40, p = .015. This indicates that the specific association between inter-subject correlations and number of correct answers is individual specific.

We repeated the four steps of the analysis using the permuted inter-subject correlation values obtained after using non-matching movies. If physiological synchrony indeed reflects the level of attention to the same presented stimulus, no association between the permuted inter-subject correlation values and the number of correct answers is expected. Indeed, there was no significant main predictive effect of the permuted inter-subject correlation values of EDA and heart rate on the number of correct answers on the content of the movie. The statistical parameters of the HLMs used for this analysis can be found in Supplementary Table 3.

5.3.3. Inter-subject correlations in EDA or heart rate as predictor of vigilance

Next, we investigated the potential association between inter-subject correlation in EDA and heart rate and performance on a subsequent attentional task, i.e. the PVT. A premise for our analysis is that vigilant performance is affected by sleep deprivation and thus varies over the course of the night. Figure 5-5 shows the PVT lapse probability over the course of the night. The median lapse probability gradually increases from close to zero up to 0.4 at 05:00 o'clock. This is followed by a strong decrease to the initial level at 06:00 o'clock. This overall pattern is consistent with previous findings (Hudson et al., 2020). Also consistent with previous findings are the large individual differences in vigilant performance during the night (Hudson et al., 2020), indicated by the shaded area in the figure.

A second premise for inter-subject correlations to be predictive of vigilant performance throughout the night is that the inter-subject correlations vary throughout the night. Figure 5-6 shows inter-subject correlations in EDA and heart rate over the course of the night. Inter-subject correlations in heart rate do not seem to follow a consistent pattern throughout the night. Inter-subject correlations in EDA seem more consistent with the overall lapse probability. The figure suggests a decreasing trend of inter-subject correlations in EDA up to 05:00 o'clock, followed by an increase from 06:00 o'clock.

Table 5-4 summarizes the results of the four steps of HLM analyses used to in-

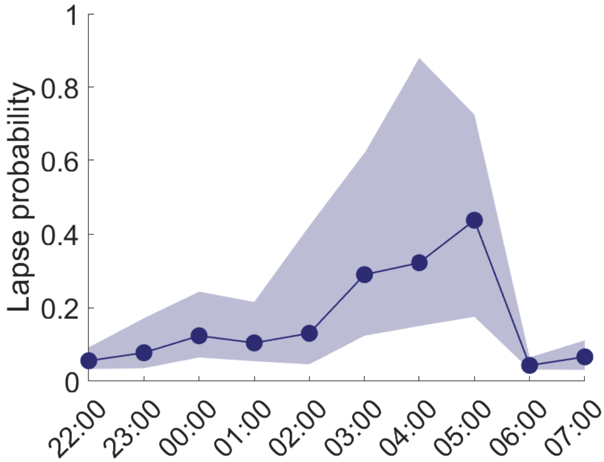

Figure 5-5. PVT lapse probability over the course of the night. The markers depict the median across participants, shaded area depict the 25^{th} to 75^{th} percentile.

vestigate how well inter-subject correlations in either EDA or heart rate predict PVT lapse probability. In step one, only the dependent variable PVT lapse probability was added to the model to serve as a baseline. The AIC is in this model is 701.69.

5

In step two, the level two variable participant was added to the model to inves-

Table 5-4. Statistical parameters of the two hierarchical linear models using inter-subject correlations in either EDA (ISC-EDA) or heart rate (ISC-HR) as predictor of PVT lapse probability.

Step	Predictor	-2LL	DF	AIC	$\chi 2$ -2LL change	R^2	t fixed effect
1		-348.85	2	701.69			
2		-288.12	3	582.24	(1,N = 439) = 121.45, p < .001	.365	-17.18***
3a	ISC-EDA	-275.59	4	559.18	(1, N = 439) = 25.06, p < .001	.414	-5.12***
3b	ISC-HR	-287.63	4	583.27	(1, N = 439) = 0.97, p = .323	.366	-0.99
4a	ISC-EDA	-275.43	6	562.87	(2, N = 439) = 0.31, p = .856	.414	-5.23***
4b	ISC-HR	-287.61	6	587.22	(2, N = 439) = 0.05, p = .977	.366	-1.01

*** p < .001

Abbreviations: -2LL: = minus 2 Log Likelihood statistic; DF = degrees of freedom; AIC = Akaike Information Criterion

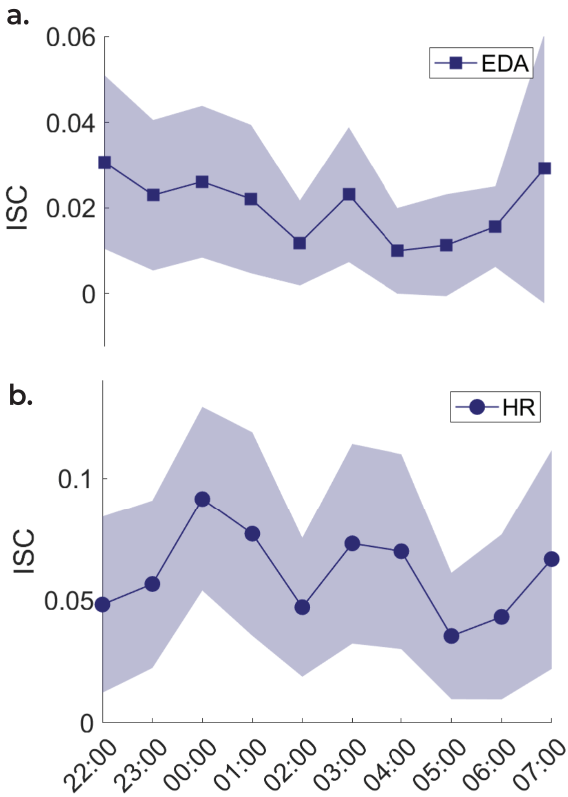

Figure 5-6. Inter-subject correlations in a. EDA and b. heart rate (HR) over the course of the night. Markers depict the mean across participants, shaded area depicts the standard deviation.

tigate whether allowing for individual differences improves the model prediction. The AIC is in this model is 582.24. There is a highly significant change in the -2LL, χ^2 (1, N = 439) = 121.45, p < .001. This step of the model explains 36.5% of the variance in PVT lapse probability. The fixed effect of individual on PVT lapse probability is highly significant (p < .001).

In step three, the predictor, being inter-subject correlations in either EDA (step 3a) or heart rate (step 3b), was added to the model with a fixed slope and random intercept. The results in this step indicate whether the predictor has a main predictive effect on PVT lapse probability. Adding the predictor with a fixed slope means that the model assumes that for all individuals, the relationship between the predictor and PVT lapse probability is identical (Grinsted et al., 2004). In this step, the AIC value is 559.18 for EDA and 583.27 for heart rate. The -2LL value for EDA is -275.59, a significant improvement compared to the previous step, χ^2 (1, N = 439) = 25.06, p < .001. The total variance explained is 41.4% and the fixed effect is highly significant (p < .001). For heart rate, the -2LL value (-287.63) is not significantly lower than the model in step two, χ^2 (1, N = 439) = 0.97, p = .323. The total variance explained is 36.6 percent and fixed effect is not significant.

In step four, the predictor variable was added with a random slope and a random intercept, meaning that the model allows the relationship between inter-subject correlations and PVT lapse probabilities to vary between individuals. Neither for EDA (step 4a) nor for heart rate (step 4b) the model in this step was significantly better than the model in step three (EDA: AIC = 562.87, -2LL = -275.43, -2LL χ^2 (2, N = 439) = 0.31, p = .856; heart rate: AIC = 587.22, -2LL = -287.61, -2LL χ^2 (2, N = 439) = 0.05, p = .977).

5.3.4. Individuals' mean EDA or heart rate as predictor of vigilance

To allow comparison of the predictive value of inter-subject correlations with individual physiological responses, we repeated steps three and four of the above analyses with mean EDA or heart rate as predictor. Figure 5-7 shows the median EDA and heart rate across participants during the movies over the course of the night. Shaded area depicts the 25th to 75th percentile, to reflect the variation across individuals. Both median EDA and heart rate seem to gradually decrease up to about 02:00, and then steadily increase, except for a relatively high median heart rate at 2:00. Indeed, median EDA and median heart rate are significantly and positively correlated (r = .18, p < .001).

Table 5-5 depicts the HLM statistics corresponding to the analyses with individual mean EDA and heart rate values as predictors. When added as predictor with fixed slope, both for EDA (step 3a) and heart rate (step 3b) the model significantly improves over models in which EDA or heart rate were not included as predictor (EDA: χ^2 (1, N = 439) = 5.44, p = .019, p = .016; heart rate: χ^2 (1, N = 439) = 9.45, p = .002. Both EDA and heart rate have a main significant predictive effect on lapse probability (p < .05, p < .01, respectively). Neither for EDA nor for heart rate the model improved when the predictors were added with random slope to allow variation in the relationship between the predictor and PVT lapse probability across participants.

When comparing the AIC scores of the predictions using inter-subject correlations and individuals' physiological activation, for EDA the former appears a better predictor. For heart rate, mean heart rate has more predictive value than inter-subject correlations in heart rate.

5.3.5. Individuals' mean EDA and inter-subject correlations in EDA as predictor of vigilance

As both mean EDA and inter-subject correlations in EDA had a significant main predictive effect on PVT lapse probability, we investigated the potential add-

5

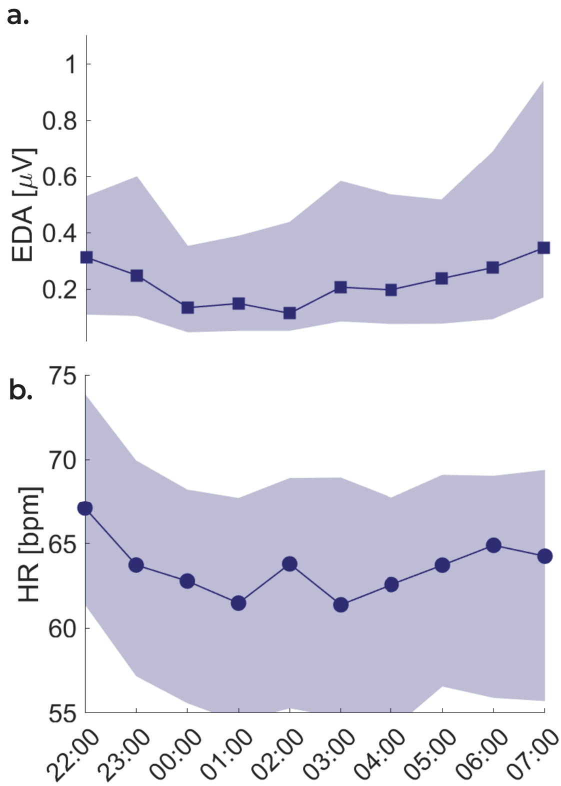

Figure 5-7. a. Phasic EDA and b. heart rate (HR) during the ten movies over the course of the night. The markers depict the median across participants, shaded area depict the 25th to 75th percentile.

ed value of inter-subject correlations in EDA over one's individual mean phasic EDA. Table 5-6 shows the statistical parameters of the hierarchical linear model using both individuals' mean EDA and inter-subject correlations in EDA as predictor with random intercept and fixed slope. It shows that that the prediction of lapse probability is significantly better with mean EDA and inter-subject correlations in EDA compared to only mean EDA ($p < .001$).

5.4. Discussion

5.4.1. Physiological synchrony as predictor of vigilant attention

The main aim of the current work was to investigate whether variations in vigilant attention that arise from sleep deprivation can be captured by inter-subject correlations in EDA and heart rate as measures of physiological synchrony. To do so we investigated whether inter-subject correlations in the phasic component of EDA and heart rate could predict the variations in subsequent vigilant attentional performance throughout a night of sleep deprivation. We found that inter-subject correlations in EDA had a significant main predictive

Table 5-5. Statistical parameters of the hierarchical linear models using individuals' average EDA or heart rate (HR) as predictor of PVT lapse probability.

Step	Predictor	-2LL	DF	AIC	$\chi2$ -2LL change	R^2	t fixed effect
1		-348.85	2	701.69			
2		-288.12	3	582.24	(1, N = 439) = 121.45, p < .001	.365	-17.18***
3a	EDA	-285.4	4	578.81	(1, N = 439) = 5.44, p = .019	.378	-2.35*
3b	HR	-283.4	4	574.79	(1, N = 439) = 9.45, p = .002	.380	-3.09**
4a	EDA	-285.29	6	582.58	(2, N = 439) = 0.23, p = .891	.378	-2.30*
4b	HR	-282.81	6	577.61	(2, N = 439) = 1.18, p = .554	.381	-3.22**

* p < .05, ** p < .01, *** p < .001

Abbreviations: -2LL: = minus 2 Log Likelihood statistic; DF = degrees of freedom; AIC = Akaike Information Criterion

Table 5-6. Statistical parameters of the hierarchical linear models using individuals' average EDA or heart rate (HR) as predictor of PVT lapse probability.

Step	Predictor	-2LL	DF	AIC	$\chi2$ -2LL change	R^2	t fixed effect
3c	EDA ISC-EDA	-273.04	5	556.08	(1, N = 439) = 24.73, p < .001	.425	-2.28* -5.08***

* p < .05, ** p < .01, *** p < .001

Abbreviations: -2LL: = minus 2 Log Likelihood statistic; DF = degrees of freedom; AIC = Akaike Information Criterion

effect on the performance of a vigilant attention task. This was not the case for inter-subject correlations in heart rate.

Using two separate permutation methods, we established that the inter-subject correlations in EDA and heart rate were higher than expected based on chance for most participants and most movies.

We first replicated the relation between inter-subject correlations in signals recorded during the narrative and narrative retention. Inter-subject correlations in heart rate and EDA were both positively correlated with the number of correctly answered questions about the movie clips. In previous work we only found this relation to be statistically significant for inter-subject correlations in heart rate (Stuldreher et al., 2020b, 2023a). Now we find that in EDA such inter-subject correlations are associated with performance as well. Additionally, we extended previous work by showing that inter-subject correlations do not only capture performance differences between individuals, who received different attentional instructions, or with different personal characteristics, but that inter-subject correlations also capture the variations in performance within an individual. The prediction of the number of correct answers using inter-subject correlations improved when the hierarchical linear model allowed the relation to vary between individuals. This indicates that the relation between inter-subject correlation and attentional performance is different for everyone and confirms the benefit of personalized models in neuroergonomics (Dehais et al., 2020).

Importantly, we also find that the inter-subject correlations in EDA during narrative movie clips predict performance on a consecutive vigilant attention task. This suggests that the momentary attentional processing capabilities captured by inter-subject correlations during the presented narrative reflect longer lasting variations in general attentional capability of individuals that also affects other types of tasks. Although significant, the association between inter-subject correlations and vigilant attention, and the predictive value of inter-subject correlations in the model, was modest. This may be explained by a discrepancy in the attentional processes that are captured by inter-subject correlations in response to relatively engaging movies and the subsequent PVT, and the effect sleep deprivation has on each of these processes. Lapses in attention, i.e. short moments of inattentiveness, have been considered the main reason for a decline in cognitive functioning through sleep deprivation (Williams et al., 1959; Kjellberg, 1977), though later it was found that in between lapses cognitive functioning was also impacted through slowing of cognitive processing (Kjellberg, 1977; Dorrian and Dinges, 2005) and fluctuations in alertness (Doran et al., 2001). Still, sleep deprivation especially impacts cognitive functioning in long, simple and monotonous tasks requiring reaction speed or vigilance (Alhola and Polo-Kantola, 2007; Lim and Dinges, 2008; Hudson et al., 2020). The PVT is specifically designed to show a strong effect of sleep deprivation. It is a long monotonous task, that captures the lapses in attention and does not require cognitively complex or emotional processing. Inter-subject correlations, on the other hand, have been reported to capture the cognitive processing of a shared stimulus and to be modulated by the level of “attentional engagement” with the stimulus (Dmochowski et al., 2012; Stuldreher et al., 2020b; Madsen and Parra, 2022). The concept attentional engagement implies a broad type of attention and entails processes like logical reasoning, emotional processing, empathy elicitation and low level visual processing (Dmochowski et al., 2012). The stimuli during which we monitored inter-subject correlations were not simple and monotonous, but relatively engaging movies. Thereby, attending to these movies does not require strong vigilant abilities as the movies’ features attract attention automatically in addition to the internal guidance of attention to the movie. Attentional engagement and the inter-subject correlations that are said to capture it, may thus not be influenced as strongly by sleep deprivation as the psychomotor vigilance task. Sleep deprivation generally does not show such a strong decline in performance for tasks that encourage participants to remain engaged and attentive compared to vigilance tasks (Pilcher et al., 2007).

As we presented all participants with the same order of movies, we cannot separate effects of sleep deprivation on inter-subject correlation with the effect of individual movies on inter-subject correlations in the current experiment. In a previous experiment with six of the ten current movies we did not find the same pattern in the inter-subject correlations in EDA and heart rate as in the current study (Stuldreher et al., 2023a), suggesting it is sleep deprivation that affects the inter-subject correlations over the night and not movie-specific features.

5.4.2. Individuals' physiological activity as predictor of vigilant attention

We compared the predictive value of inter-subject correlations in heart rate and EDA to the predictive value of individual's mean EDA and heart rate. Mean EDA and heart rate both predicted vigilant performance. Since inter-subject correlations in heart rate did not predict performance at all, the mean heart rate thus had a higher predictive value. The mean EDA explained less variance of the vigilant attention than inter-subject correlations in EDA. Also when comparing the AIC, inter-subject correlations in EDA seemed to provide a better model fit than the mean value. To further investigate the potential added value of inter-subject correlations, we added both individuals' mean EDA and inter-subject correlations in EDA to an HLM. We compared performance to that of the model using only individuals' mean EDA. The predictive performance was found to significantly improve when inter-subject correlations were added, indicating that interpersonal analysis of EDA is of added value when interested in monitoring attention.

Our finding that EDA and heart rate were negatively associated with subsequent lapse probability is in line with the established relation between decreased vigilant performance when arousal decreases. Miró et al. (2002) reported a steady decrease in skin conductance throughout the first sleep deprived night, indicating that arousal decreases over the course of a sleep deprived night. Posada-Quintero et al. (2017) found a high negative correlation between the mean values of the skin conductance level and reaction time among sleep deprived individuals, indicating that the arousal decrease indeed results in worse task performance. Van Den Berg and Neely (2006) found similar effects using heart rate, as they reported a strong negative correlation between heart rate and reaction time in sleep deprived individuals.

5.4.3. Differences between results using EDA and heart rate

For EDA, both the model using inter-subject correlations and the model using individuals' mean activity predict vigilant performance. However, for heart rate, only the mean activity could significantly predict vigilant performance. It thus appears that for heart rate, it is specifically the congruency in timing of the response across individuals that cannot capture decreased vigilant performance. We are not sure what underlies this observation. It is not the case that inter-subject correlations in heart rate could not capture attention altogether. For the majority of participants, inter-subject correlations in heart rate were higher than one would expect based on chance, indicating that attending to a shared stimulus causes fluctuations in heart rate to synchronize across participants. Additionally, inter-subject correlations in EDA and heart rate had a significant main predictive effect on the number of corrects answers on questions about the movies. For this latter analysis, the predictive effect of inter-subject correlations was actually larger for heart rate then EDA.

The discrepancy in findings using EDA and heart rate points in the direction that inter-subject correlations in these signals both capture different aspects of attentional processing. Higher inter-subject correlations in heart rate were related to better memory retention during the movies. Also in previous work, we and others found heart rate to be associated with memory retention of the presented stimulus (Stuldreher et al., 2020b; Pérez et al., 2021; Madsen and Parra, 2022). In previous work, inter-subject correlations in heart rate did not reflect the occurrence of emotional sounds attracting attention bottom-up among individuals instructed not to focus on them, but to focus instead on the simultaneously presented audiobook (Stuldreher et al., 2020a). It thus appears that inter-subject correlation in heart rate are more associated with higher order cognitive processes of attentional engagement than with shorter moments of attentiveness driven by sensory processes. Higher inter-subject correlations in EDA were related to better memory retention during the movies and better subsequent vigilant performance . We previously did not find an association between inter-subject correlations in EDA and memory retention for narratives (Stuldreher et al., 2020b). Here we do find this association, but still to a lesser degree than for heart rate. In previous work, inter-subject correlations in EDA occurred after emotional sounds attracting bottom-up attention (Stuldreher et al., 2020a). We speculate that compared to inter-subject correlations in heart rate, inter-subject correlations in EDA are more associated with shorter moments of attentiveness due to discrete, arousing sensory events and less so to longer periods of attentional engagement driven by higher-order cognitive processes. This could also explain why inter-subject correlations in EDA were sensitive to sleep deprivation, but inter-subject correlation in heart rate were not. Short moments of inattentiveness, or lapses in attention, are thought to be the main contributor to the cognitive decline induced by sleep deprivation (Williams et al., 1959; Kjellberg, 1977). Such lapses may be associated in reduced electrodermal responses to otherwise arousing events in the movie in sleep-deprived individuals. As the lapses do not occur at the same points in time across individuals, they would result in reduced inter-subject correlations in EDA.

5.4.4. Conclusions

We replicated findings that physiological synchrony during presented narratives is associated with performance on questions about the narratives. In addition, we found that physiological synchrony was associated with performance on a subsequent vigilant attention task. These findings confirm the association between physiological synchrony and attention. Physiological synchrony captures the attentional processing during the narratives, and proves valuable for capturing more general changes in the attentional state of monitored individuals. The discrepancy in findings using EDA and heart rate suggest that physiological synchrony in these measures captures different aspects of attentional processing. Physiological synchrony in EDA may especially reflect short moments of attentiveness caused by arousing sensory events. Synchrony in heart rate may rather reflect longer intervals of attention driven by higher-level cognitive control. Individuals' mean EDA and heart rate were also associated with performance on the subsequent vigilant attention task. For EDA, inter-subject correlations explained more variance if vigilant performance than individual's mean activity. Using inter-subject correlations in EDA in addition to mean EDA to predict vigilant performance yielded better performance than prediction based on mean EDA alone. Our results are an important step towards the use of physiological synchrony as implicit measure of shared attention, as we show that variations in the attentional abilities within individuals can be captured. We see physiological synchrony as potential tool to monitor the changes in attentional engagement within and between individuals in a group. It may for instance be used to assist teachers in evaluating the attentional engagement of students in an (online) classroom or to track the shared attention among cooperating teammates.

Acknowledgments

We would like to thank all the colleagues that contributed to the data acquisition.

Part II:

Physiological synchrony from lab to life

6.

A comparison between laboratory and wearable sensors in the context of physiological synchrony

Van Beers, J.J. , Stuldreher, I.V., Thammasan, N., Brouwer, A.M. (2020). A comparison between laboratory and wearable sensors in the context of physiological synchrony. In *Proceedings of the 2020th International Conference on Multimodal Interaction*, Utrecht, The Netherlands, 604-606. doi: 10.1145/3382507.3418837

Abstract

Measuring concurrent changes in autonomic physiological responses aggregated across individuals (Physiological Synchrony; PS) can provide insight into group level cognitive or emotional processes. Utilizing cheap and easy to use wearable sensors to measure physiology rather than their high-end laboratory counterparts is desirable. Since it is currently ambiguous how different signal properties (arising from different types of measuring equipment) influence the detection of PS associated with mental processes, it is unclear whether, or to what extent, PS based on data from wearables compares to that from their laboratory equivalents. Existing literature has investigated PS using both types of equipment, but none compared them directly. In this study, we measure PS in electrodermal activity (EDA) and inter beat interval (IBI, inverse of heart rate) of participants who listened to the same audio stream but were either instructed to attend to the presented narrative (N = 13) or to the interspersed auditory events (N = 13). Both laboratory and wearable sensors were used (ActiveTwo electrocardiogram (ECG) and EDA; Wahoo Tickr and EdaMove 4). A participant's attentional condition was classified based on which attentional group they shared greater synchrony with. For both types of sensors, we found classification accuracies of 73% or higher in both EDA and IBI. We found no significant difference in classification accuracies between the laboratory and wearable sensors. These findings encourage the use of wearables for PS based research and for in the field measurements.

6.1. Introduction

Autonomic physiological responses can provide informative insights into an individual's cognitive and emotional state. When aggregated across multiple individuals, group level dynamics may be investigated through similarities in their physiological activity. This concept is known as physiological synchrony (PS).

One prevalent domain in which autonomic PS has been extensively employed is that of interpersonal interactions (Palumbo et al., 2017). Use of PS can also be envisioned in various human computer interactions, such as in competitive multiplayer video games. Interest in autonomic PS is gaining traction in part due to the rapid development and growing maturity of wearable sensor technology (Mukhopadhyay, 2015; Seshadri et al., 2019; van Lier et al., 2020) compounded with the possibility to combine a myriad of different wearable devices. Moreover, synchrony within different physiological modalities could be reflective of different processes, such as stress or empathy (Palumbo et al., 2017), and may provide insight into the mechanisms driving PS. It is suspected that PS can be used to measure shared attention (Brouwer et al., 2019), where shared attention may be an underlying explaining factor of other findings such as those by (Gashi et al., 2019) on audience engagement. Attention itself plays an important role in learning capabilities (Jiang and Chun, 2001; Jamet et al., 2008), task performance (Pashler et al., 2001), and social interactions (Andrade et al., 2009). Jamet et al. demonstrated the benefits of using attention guiding techniques to facilitate learning, resulting in improved performance for retention (e.g. memory) based tasks (Jamet et al., 2008). As such, PS may be of interest as a tool to monitor attention continuously and unobtrusively in the classroom to assist students with learning disabilities, or to improve upon existing teaching methods (Dikker et al., 2017).

Wearable sensors are typically unobtrusive, affordable, and mobile, enabling 'in the field' research which may provide for more realistic insights into natural human behavior. However, these benefits usually come at the cost of a diminished signal. For instance, the Wahoo Tickr, a wearable used to measure heart rate (HR), high frequency HR information is lost due to the low sampling rate and on board processing (Borovac et al., 2020). It is still unclear what physiological signal aspects are relevant for measuring PS and how the limitations imposed by wearables influence their ability to measure meaningful affective/cognitive PS. Therefore, we chose to directly compare PS obtained through both laboratory and wearable sensors. Indeed, there are a few studies which compare wearable and laboratory grade equipment, such as that conducted by Ragot et al. on recognizing emotion (Ragot et al., 2018). However, these lie outside the domain of PS and thus may rely on different signal aspects.

To directly compare laboratory and wearable equipment, data obtained from an experiment described in (Brouwer et al., 2019; Stuldreher et al., 2020b) were used. In this experiment, participants were instructed to listen to the same audio track, but to attend to different stimulus aspects with the aim of determining the selective attention in groups using PS. In (Brouwer et al., 2019; Stuldreher et al., 2020b), EDA and HR data obtained from laboratory equipment (ActiveTwo) were analyzed. During the experiment, EDA and HR were concurrently measured with wearable sensors (Wahoo Tickr, HR; EdaMove4, EDA). The current study elaborates on this experiment and directly compares autonomic PS results between wearable data and their laboratory counterparts, when subject to the same conditions and analysis methods. Therein, we aim to evaluate the feasibility of the use of wearables, spanning two physiological modalities, in the domain of PS. To the best of our knowledge, this is the first study that compares PS from wearable data with PS from high end laboratory equipment.

6.2. Methods

6.2.1. Participants

Participants (N = 27, aged between 18 and 48), with no self-reported problems in hearing or attention, were recruited from the research institute's (TNO) participant pool. All participants signed an informed consent form prior to the experiment and were given a small monetary reward after the experiment. Data of one participant was removed due to failed recordings. The study was approved by the TNO Institutional Review Board (TCPE) and the TU Delft Human Research Ethics Committee.

6.2.2. Materials

For the laboratory equipment, both EDA and electrocardiogram (ECG) were measured via an ActiveTwo system (BioSemi, Amsterdam, Netherlands) at 1024 Hz. For EDA, two passive gelled Nihon Kohden electrodes were placed on the ventral side of the distal phalanges of the middle and index finger on participants' left hand. For ECG, two active gelled Ag AgCl electrodes were placed at the right clavicle and lowest floating left rib. Regarding the wearable equipment, EDA was recorded through an EdaMove4 (movisens GmbH, Karlsruhe, Germany) at 32Hz while HR was measured with a Wahoo Tickr (Wahoo Fitness, Atlanta, Georgia, USA) at 1Hz. The EdaMove4 was attached by two self-adhesive electrodes placed on the palm on participants' left hand. The Wahoo Ticker was fitted around the chest of participants after applying gel on its sensors. The Wahoo Tickr outputs a filtered HR signal with a minimum increment of 1 bpm derived from a measured electrical signal and thus does not provide raw inter beat intervals (IBI). The signal resolution (1 bpm) and sampling rate (1 Hz) are independent of each other, but the amount of information contained in the signal is dependent on both. Thus, the Wahoo Tickr lacks high frequency HR information (e.g. respiratory sinus arrhythmia RSA), that are present in the ActiveTwo ECG.

6.2.3. Stimuli and design

Each participant completed the experiment individually and listened to the same audio file. This file was composed of a 66 min audiobook (a Dutch thriller written by Corine Hartman: 'Zure koekjes') with interspersed auditory stimuli (beeps and affective sounds). These short stimuli were randomly ordered with intervals between stimuli ranging from 35 to 55 seconds. Half of the participants were assigned to attend to the narrative (NA) of the audiobook and to ignore all other stimuli. The other half of the participants were asked to focus on the short stimuli (SSA) and ignore the narrative.

The affective sounds, of 6 second durations, are taken from the International Affective Digitized Sounds (IADS) (Bradley and Lang, 2007a): a collection of acoustic stimuli normatively rated for emotion, valence and dominance. Examples include sounds of a crying baby or the cheers of a sports crowd. 12 neutral sounds, 12 pleasant sounds and 12 unpleasant sounds were elected. Beeps were presented in blocks of 30 seconds, with every two seconds a 100ms high (1kHz) or low (250Hz) pitched beep. SSA participants were tasked with counting the number of high and the number of low tones (De Dieuleveult et al., 2018). 27 blocks of sounds were presented.

6.2.4. Analysis

Data processing was done using MATLAB R2018b (Mathworks, Natick, MA, USA). ActiveTwo EDA measurements were downsampled to 32 Hz. For both ActiveTwo and EdaMove4, the phasic component of the EDA response was extracted for further analysis using the Ledalab toolbox for MATLAB (Benedek and Kaernbach, 2010). Studies on EDA typically show a certain number of 'non responders' (Braithwaite et al., 2013), or weak responders participants with a low EDA magnitude and near zero phasic response. Weak responders in our study were identified through visual inspection by the individual authors of the manuscript. Data of these participants were not discarded since the weak responses of these participants appeared to contain information pertaining to the shape of the response, which may be useful for synchrony. However, the phasic responses of the EdaMove4 weak responders were contaminated with peaks arising from noise and jitter due to on board processing, distorting the signal shape. Therefore, the full EDA traces of the EdaMove4 weak responders were filtered using a Savitzky Golay filter with a three second window and the phasic components were recomputed. Since the experiment event markers are expressed in the ActiveTwo timeline, we accounted for delays between EdaMove4 and ActiveTwo, among others arising from on board processing on EdaMove4. The phasic response obtained via EdaMove4 was time corrected through a normalized cross correlation with the phasic response from ActiveTwo. Here, the lag maximizing the correlation of the two signals is assumed to represent the accumulated delay (inconsistent across participants).

ECG measurements acquired from ActiveTwo were first downsampled to 256Hz, then high pass filtered at 0.5Hz. R peaks of the ECG signal were detected following (Pan and Tompkins, 1985). The resulting semi timeseries of consecutive IBIs were subsequently interpolated and resampled at 32Hz to transform them into a timeseries. The Wahoo Tickr HR signal was first upsampled to 32 Hz, then time corrected through a normalized cross correlation with the ActiveTwo derived HR (i.e. inverse of IBI). The pre-processed Wahoo Tickr HR was then converted to IBI.

Regardless of sensor type and physiological signal, inter-subject correlations (ISC) were determined using a moving window, as introduced by (Marci et al., 2007). A window of size 15 seconds traverses the signal at 1 second increments with Pearson correlations calculated over successive windows. The overall correlation between two responses is given by the natural logarithm of the sum of all positive correlations divided by the absolute value of the sum of all negative correlations. Classifications were based on the average ISC of a participant with all members from the NA group and all members from SSA group, excluding the participant in question. Participants were classified with the attentional group that they were more correlated with (i.e. shared the highest ISC). Paired sample t-tests were conducted to determine whether the NA ISC and SSA ISC were significantly different within each attentional group for EDA and IBI. Chance level classifications were determined through surrogate data with 100 instances of randomly shuffled attentional group labels. To evaluate if the classification accuracies between the laboratory and wearable sensors are statistically different, an exact McNemar's test was used. This test is suitable to compare paired nominal data with small sample sizes, such as ours (Foody, 2009; Adedokun and Burgess, 2011).

6.3. Results

Figure 6-1 illustrates that the patterns in ISC are similar between the laboratory and wearable sensors, and that overall, withingroup ISC is higher than between group ISC for both types of equipment. Figure 6-1 also presents results for both NA and SSA participants. For EDA, the within group ISC is significantly higher

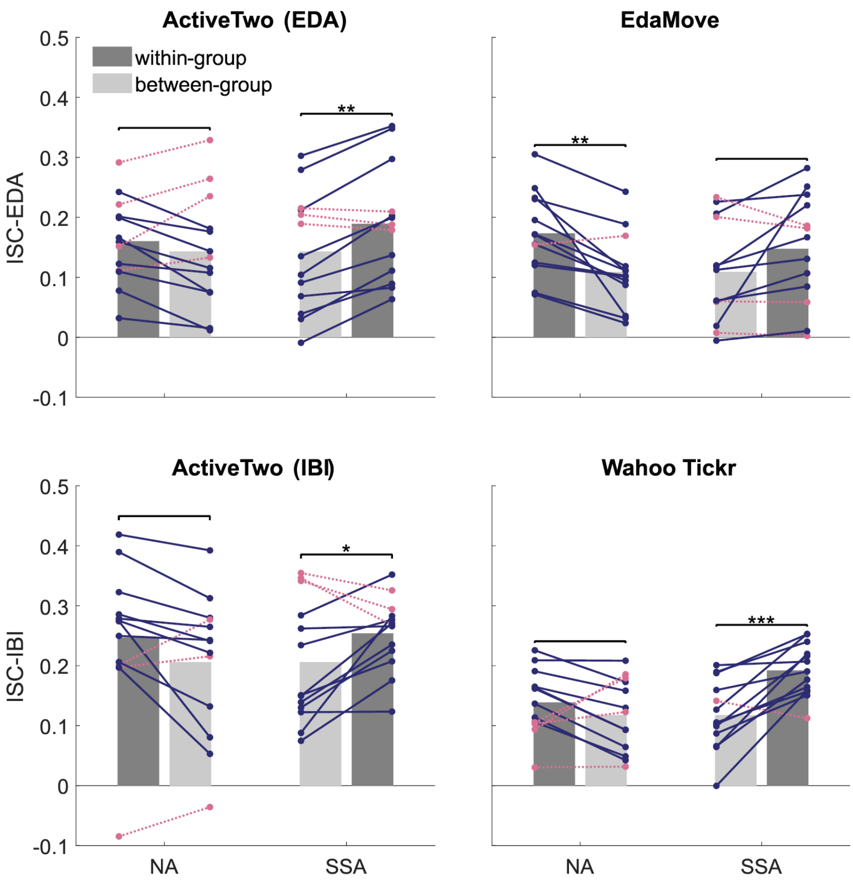

Figure 6-1. Within-group and between-group inter-subject correlations (ISC) of electrodermal activity (EDA, top) and inter-beat interval (IBI, bottom) for both attentional groups (NA, left bars; SSA, right bars) derived from laboratory (ActiveTwo, left column) and wearable (EdaMove & Wahoo Tickr, right column) sensors. Also illustrated are connected dots which represent the individual participants. Full blue lines indicate higher within-group ISC. Dotted pink lines denote higher between-group ISC. Paired sample t-tests were used to determine if within-group ISC are significantly higher than between-group ISC (* $p < .05$,** $p < .01$,*** $p < .001$).

than the between group ISC for NA participants with the EdaMove4 ($t(12) = 4.16$, $p = .001$) but not for the ActiveTwo ($t(12) = 0.74$, $p = .476$). For SSA participants, within group ISC is significantly higher than between group ISC for the ActiveTwo ($t(12) = 4.07$, $p = .002$) but not for the EdaMove4 ($t(12) = 1.98$, $p = .072$). Regarding IBI, the significance of the patterns in group level ISC are consistent between the ActiveTwo and Wahoo Tickr. For SSA participants, the within-group ISC is significantly higher than the between group ISC (ActiveTwo: $t(12) = 2.27$, $p = .043$; Wahoo Tickr: $t(12) = 4.75$, $p < .001$). Within group ISC is not significantly

Table 6-1. Overall classification accuracy (in percentage) of correctly identified participant attentional conditions based on their inter-subject correlations (ISC). The chance level values, and associated standard deviations, are given in brackets. The corresponding p-values are also presented.

	ActiveTwo	Wearables (EdaMove4, Wahoo Tickr)
EDA	73 (50 ± 10) p = .009	81 (52 ± 9) p = .005
IBI	77 (50 ± 11) p = .009	81 (50 ± 9) p = .001

higher than between group ISC for NA participants (ActiveTwo: $t(12) = 2.02$, $p = .066$; Wahoo Tickr: $t(12) = 1.38$, $p = .192$).

Table 6-1 presents the percentage of participants whose attentional condition (i.e. NA or SSA) was correctly identified. The corresponding chance level classification accuracies are also shown in accompanying brackets. For EDA, the classification accuracy is significantly above chance for both ActiveTwo ($t(99) = 2.39$, $p = .009$) and EdaMove4 ($t(99) = 2.91$, $p = .005$). Likewise, the IBI classification accuracy is significantly above chance for both ActiveTwo ($t(99) = 2.43$, $p = .009$) and Wahoo Tickr ($t(99) = 3.38$, $p = .001$). Rather than performing worse, wearables tend to outperform their laboratory counterparts. EdaMove4 classifies 81% of the participants correctly as opposed to 73% with ActiveTwo. Similarly, Wahoo Tickr classifies 81% of the participants correctly, in comparison to 77% with ActiveTwo. However, an exact McNemar's test showed no statistical difference between the classification accuracy of ActiveTwo and EdaMove4, $p = .727$, or between ActiveTwo and Wahoo Tickr, $p = 1.000$.

6.4. Discussion

Through this study, we have shown that PS in selective attention can be derived from wearable sensors, EdaMove4 and Wahoo Tickr, equally well as their laboratory based counterparts.

Our results are especially notable for the Wahoo Tickr, given its poor resolution (1 bpm) and sampling rate (1Hz). This suggests that wearables with lower bitrates may also be appropriate for PS based research, broadening the potential applications of autonomic PS. The relatively good performance of the Wahoo Tickr could suggest that the very low to low frequency HR (i.e. 0.003 to 0.15Hz) (Quer et al., 2016) is an influential feature for determining synchrony in selective attention. The lower frequency components of the ActiveTwo and Wahoo Tickr HR traces are mostly coincident, hence, the presence of high frequency HR (e.g. due to breathing) could act as 'noise' and may explain some of the differences in classification performance between these sensors. Consequently, future work should investigate methods to remove breathing from the ActiveTwo data to compare results more directly with the Wahoo Tickr. Under conditions of movement, synchrony in HR due to shared attention may be strengthened, if the movements are associated with shared attention such as in (Quer et al., 2016), or overshadowed when unrelated as seen in (Verdiere et al., 2020). However, any synchrony in HR induced by quick breathing pattens, as with (Quer et al., 2016), will not be captured by the wearable sensor used here.

Differences between EdaMove4 and ActiveTwo can, in part, be explained through the ‘weak responders’. In total, there were three weak responders for ActiveTwo, two of which were in the SSA group, and seven weak responders for EdaMove4, five of which were in the SSA group. The large concentration of weak responders among the EdaMove4 SSA participants may explain the difference in significance of the group level ISC between ActiveTwo and EdaMove4 seen at the top of Figure 6-1. Phasic responses of the EdaMove4 weak responders may have lacked some synchrony relevant features (e.g. peaks) which were either filtered out or were not present in the initial signal, resulting in poorer classification performance. For instance, two participants who were weak responders for EdaMove4 but not for ActiveTwo were misclassified with EdaMove4 data and correctly classified with ActiveTwo data. This misclassification of weak responders is not unique to EdaMove4 since ActiveTwo also misclassifies some weak responders. In general, weak responders are difficult to classify due to a lack of informative features. This lack of features may also artificially suppress the magnitude of the group level ISC, leading to unreliable classifications which extend beyond these weak responders. To mitigate this, a different physiological modality (such as IBI) may be used to compliment the classification result. In the current study, all but one of the EdaMove4 weak responders were correctly identified by the Wahoo Tickr, motivating the use of various modalities to augment classification accuracies.

For participants who were not weak responders, any discrepancies in performance between ActiveTwo and EdaMove4 may be explained by the long recovery time of the EdaMove4 (Borovac et al., 2020). We suspect that the long recovery time is due to the large adhesive pads of EdaMove4 which impair the evaporation of sweat. In this region, the magnitude of the phasic response is locally reduced while the noise level remains constant. This culminates in a lower signal to noise ratio within the affected region and mirrors challenges observed with weak responders. Moreover, this locally reduced response may artificially inflate synchrony since this feature has a large temporal footprint and is present across all participants (high chance of overlap between participants). Due to this, using EdaMove4 may be limited to experiments which aim to measure synchrony across temporally sparse events, or that combine various physiological modalities.

The exact McNemar's tests show that there is no statistical difference in classification accuracy between wearables and laboratory equipment when measuring PS in shared attention, and the trend is even such that wearables perform better rather than worse. Clearly, our findings encourage the use of wearables for PS based experiments and for in the field research. Limitations of this study are that the experiment was conducted in laboratory conditions with minimal movement and that only two types of wearables were used for comparison. Therefore, it is yet unclear as to how appropriate other wearables are for computing PS and how suitable wearables in general for more active applications.

6.5. Conclusion

The current study indicates that measuring PS in shared attention with laboratory and wearable sensors can result in similar performance between the two. PS derived from the wearable sensors used in this study distinguished between the two attentional conditions (NA and SSA) equally well as PS obtained from laboratory equipment, in both physiological modalities (EDA and IBI). Since wearables are less obtrusive and are inherently mobile, these results motivate the use of wearable sensors for both in the lab and in the field measurements, such as for measuring PS in an audience during artistic performances or students in a classroom.

7.

Robustness of physiological synchrony in wearable electrodermal activity and heart rate as a measure of attentional engagement to movie clips

Stuldreher, I.V., van Erp, J.B.F., Brouwer, A.M. (2023). Robustness of physiological synchrony in wearable electrodermal activity and heart rate as a measure of attentional engagement to movie clips. *Sensors*, ***23***, 3006. doi: 10.3390/s23063006

Abstract

Individuals that pay attention to narrative stimuli show synchronized heart rate (HR) and electrodermal activity (EDA) responses. The degree to which this physiological synchrony occurs is related to attentional engagement. Factors that can influence attention, such as instructions, salience of the narrative stimulus and characteristics of the individual, affect physiological synchrony. The demonstrability of synchrony depends on the amount of data used in the analysis. We investigated how demonstrability of physiological synchrony varies with varying group size and stimulus duration. Thirty participants watched six 10 min movie clips while their HR and EDA were monitored using wearable sensors (Movisens EdaMove 4 and Wahoo Tickr, respectively). We calculated inter-subject correlations as a measure of synchrony. Group size and stimulus duration were varied by using data from subsets of the participants and movie clips in the analysis. We found that for HR, higher synchrony correlated significantly with the number of answers correct for questions about the movie, confirming that physiological synchrony is associated with attention. For both HR and EDA, with increasing amounts of data used, the percentage of participants with significant synchrony increased. Importantly, we found that it did not matter how the amount of data was increased. Increasing the group size or increasing the stimulus duration led to the same results. Initial comparisons with results from other studies suggest that our results do not only apply to our specific set of stimuli and participants. All in all, the current work can act as a guideline for future research, indicating the amount of data minimally needed for robust analysis of synchrony based on inter-subject correlations.

7.1. Introduction

Individuals that attend to narrative stimuli (e.g. movie clips or auditory narratives) show synchronized neurophysiological responses such as brain potentials and heart rate (HR) (Ki et al., 2016; Stuldreher et al., 2020b; Pérez et al., 2021). This is referred to as physiological synchrony. Stronger physiological synchrony among participants is generally related to higher degrees of shared attentional engagement (Dmochowski et al., 2012, 2014; Cohen and Parra, 2016; Cohen et al., 2018; Pérez et al., 2021). Physiological synchrony is for instance higher among individuals who actively attend to a presented narrative compared to individuals who focus attention inward on a distracting task (Ki et al., 2016; Pérez et al., 2021). Additionally, when engagement decreases due to being presented with a narrative for the second time, physiological synchrony also decreases (Dmochowski et al., 2012). Physiological synchrony as marker of attentional engagement is even informative on the level of an individual within a group: the more the physiological responses of an individual synchronize with those of others attending to the narrative, the better this individual can recall the narrative (Cohen and Parra, 2016; Stuldreher et al., 2020b; Pérez et al., 2021).

Narrative-driven synchrony is most studied and most pronounced for neuroimaging modalities, such as the electroencephalogram (EEG), magnetoencephalogram (MEG) or functional magnetic resonance imaging (fMRI) (Hasson et al., 2004; Dmochowski et al., 2012; Liu et al., 2017). Brain-to-brain synchrony is often quantified through inter-subject correlations (Parra et al., 2019). Such inter-subject correlations amongst attending individuals are significantly higher than chance upon presentation of narrative stimuli (Poulsen et al., 2017), are reduced when individuals are distracted from the narrative(Ki et al., 2016), distinguish between individuals with different selective attentional focus to part of the presented stimulus (Stuldreher et al., 2020b) and are predictive of the occurrence of attentionally relevant stimuli in time (Stuldreher et al., 2020a). In addition, brain-to-brain synchrony has been related with numerous behavioral metrics, such as stimulus retention, efficacy of advertising and efficacy of communication (Dmochowski et al., 2014; Schmälzle et al., 2014; Cohen et al., 2018; Stuldreher et al., 2020b). This relation between brain-to-brain synchronization, attention and behavioral outcomes has been established for both auditory and audiovisual narratives (Cohen and Parra, 2016).

Narrative-driven synchrony has also been established in measures that reflect autonomic nervous system activity, such as HR and electrodermal activity (EDA) (Stuldreher et al., 2020b, Pérez et el., 2021). We refer to this as body-to-body synchrony. Although the brain is most closely involved in attentional processing, body measures can also index attention through reflecting arousal: the physio-

7

logical state of activation of the body. Arousal and attention are closely related and share a common neural substrate (Critchley, 2002). HR and EDA are sensitive to changes in arousal and thereby also associated with changes in attention. Indeed, HR and EDA respond to arousing, emotionally relevant events that are attentionally prioritized (Lang et al., 1998; Bradley and Lang, 2000) and contribute to the monitoring of driver attention (Awais et al., 2017; Najafi et al., 2023).

Similar to brain-to-brain synchrony, it has indeed been found that body-to-body synchrony reflects attentional engagement. Inter-subject correlations in such measures are significantly higher than chance upon presentation of narrative stimuli (Pérez et al., 2021), are reduced when individuals are distracted from the narrative (Pérez et al., 2021), distinguish between individuals with different selective attentional focus to part of the presented stimuli (Stuldreher et al., 2020b) and are predictive of the occurrence of attentionally relevant stimuli in time (Poulsen et al., 2017). Although to a lesser degree than brain-to-brain synchrony, body-to-body synchrony has been related to some behavioral metrics, such as stimulus retention (Stuldreher et al., 2020b; Pérez et al., 2021).

While body-to-body synchrony is associated with attentional engagement, it appears to be less robust and to a lesser degree related to attention than brain-to-brain synchrony. It has been argued that this may be so because body-to-body synchrony rather reflects arousal or emotional engagement, whereas brain-to-brain synchrony more directly reflects attention (Golland et al., 2014; Steiger et al., 2019; Stuldreher et al., 2020b). In addition, the potential information density of most brain measures strongly exceeds that of body measures in terms of modulation frequency and number of sensors. In general the number of participants with significant inter-subject correlations (i.e. correlations exceeding chance level expectations) was lower for HR than for EEG (Madsen and Parra, 2022). Inter-subject correlations in EDA and HR also distinguished less reliably between two selective attentional conditions than inter-subject correlations in EEG, though inter-subject correlations in HR were correlated with performance metrics in a similar way as inter-subject correlations in EEG (Stuldreher et al., 2020b).

In comparison with methods based on machine learning models, physiological synchrony is a potentially valuable method to implicitly monitor attention in real life situations. Analyzing physiological synchrony allows the use of ecological stimuli (movies and narratives). Additionally, there is no need to train a model beforehand on previously collected (personal) data. While as reviewed above, brain-to-brain synchrony seems more sensitive than body-to-body synchrony, the latter has advantages from the perspective of user comfort and costs. The use of wearable EDA and HR sensors allows for the monitoring of physiological synchrony in real-world settings, such as classrooms or conferences (Gashi et al., 2019; Liu et al., 2021). The need to move to real-life settings has been stressed many times (Brouwer et al., 2015b). While physiological synchrony as determined using wearables seems very suitable for real-life research and applications, questions regarding the analysis of physiological synchrony remain. The computational approach used to quantify physiological synchrony may affect potential outcomes. Algumaei for instance reported that physiological synchrony in the electrocardiogram (ECG) had predictive value of team performance using a multidimensional recurrence quantification analysis, but not using dyadic linear cross correlations (Algumaei et al., 2023). Linear cross correlations, on the other hand, are successfully employed to assess the attentional engagement towards narrative stimuli (Stuldreher et al., 2020b; Pérez et al., 2021). In settings where individuals are presented with the same narrative stimulus, nonlinear analyses may not be required as there are no asymmetric relationships to be captured. More complex nonlinear analysis may also require more data to be collected to set the additional modeling parameters, such that inter-subject correlations are the more suitable approach.

However, also for such inter-subject correlations, the requirements in terms of amount of data needed to obtain robust inter-subject correlations remains unclear. There are still few studies employing inter-subject correlations in HR and EDA and reporting requirements in terms of data used. Pérez et al. (Pérez et al., 2021) report that for single 60-second movie clips only a few participants show significant inter-subject correlations, but when aggregating over all 16 movie clips that they used in their study, the majority of participants show significant inter-subject correlations. Stuldreher et al. (2020b) found that when selecting parts of an entire 66 minute audio stimulus, generally less participants could be classified in the correct attentional condition. Though these results indicate that significance of inter-subject correlations depends on the amount of data used, it remains unclear what the specific relationship is between the amount of data and inter-subject correlations. There are no criterions for minimum group size or minimum stimulus length for sensible physiological synchrony results. In addition, it is unclear whether it is wise to increase the amount of data through recording from more people, or longer stimuli if the option is there.

We here explore the amount of data that is required to obtain robust results of body-to-body synchrony (HR and EDA) recorded with wearable equipment during watching movies. The amount of data is varied by the number of participants and the duration of (audiovisual) stimuli. In addition, we varied attentional instruction such that participants were either asked to attend to the movie, or to respond to a certain visual cue. As described later, this manipulation did not, or hardly, affect any outcome measure and is not the focus of the paper, but will be discussed at the end.

7.2. Materials and methods

7.2.1. Participants

Thirty participants (14 female), between 19 and 64 years old, with an average age of 39.4 years and a standard deviation of 15.8 years, were recruited through the institutes participant pool. Before performing the study, approval was obtained from the TNO Institutional Review Board (IRB). The approval is registered under reference 2020-117. All participants signed informed consent before participating in the experiment, in accordance with the Declaration of Helsinki. After successful participation, participants received a small monetary compensation for their time and traveling costs.

7.2.2. Materials

EDA and HR were recorded using wearable systems. EDA was recorded using EdaMove 4 (Movisens GmbH, Karlsruhe, Germany) worn at the wrist of the non-dominant hand. The EdaMove 4 uses two solid gelled Ag/AgCl electrodes (MTG IMIELLA electrode, MTG Medizintechnik, Lugau, Germany, W55 SG, textured fleece electrodes, 55 mm diameter) recording signals from the palmar surface of the hand. A constant direct current (DC) voltage of 0.5 V was applied to the skin. Measurements were conducted at a sampling rate of 32 Hz with an input range of 2 – 100 µS and with a resolution of 14 bits. HR was recorded using the commercially available and sport-oriented Tickr chest-strap (Wahoo Fitness, Atlanta, GA, USA). The device reports HR in beats per minute (bpm), at a rate of one value per second and a resolution of one bpm. Data were streamed over Bluetooth to a smartphone for local saving on the device through the Wahoo Fitness application (version 1.36.0.291). Both EdaMove 4 and Tickr have been demonstrated before to provide signal quality close to high-end lab equipment (Borovac et al., 2020), and to be suitable to measure attention-modulated physiological synchrony (Van Beers et al., 2020).

7.2.3. Design

Participants performed the experiment one by one. All participants were presented with six movie clips of approximately 10 minute duration (09:48 ± 00:41 minutes). The details and URLs can be found in Table 7-1. The movie clips were selected from the Dutch YouTube channels NPO3 and KORT! and featured 10-minute stories. The presentation order was randomized across participants. We chose these clips as we are not aware of any affective movie databases containing six approximately 10-minute videos of neutral valence and moderate arousal, preferably in Dutch. Although the clips are not standardized in terms of

Table 7-1. Description of shown movie clips, categorized by name, duration and URL.

Name	Duration	URL
Chauffeur	09:45	https://www.youtube.com/watch?v=jaFmvyH7dW8
El Mourabbi	09:04	https://www.youtube.com/watch?v=X9bJou2gKxo
De Chinese Muur	09:50	https://www.youtube.com/watch?v=yjGFuhPy3Qo
One of the boys	10:58	https://www.youtube.com/watch?v=PsGAuhgQ97k
Samual	09:45	https://www.youtube.com/watch?v=VUseoqCVnj4
Turn it around	09:26	https://www.youtube.com/watch?v=beC7IpQpTz4

valence and arousal, the movies are comparable in terms of length and type. All contain an emotional narrative that develops throughout the 10-minute video, as judged at face value by two of the authors. We avoided films with strongly arousing content, such as the death of characters, physical violence and verbal violence. The films are of the arthouse genre.

Besides the robustness of inter-subject correlations as a function of measurement duration and group size, we also aimed to study the effect of a double task on inter-subject correlations. For this, every 2-10 seconds a millisecond counter (red font) was displayed in the center of the screen on top of the movie. This counter disappeared after 5 seconds or upon a button press of the participant. Alternating between movie clips, participants were instructed to attend to the movie clip and to ignore the counter or to respond as quickly as possible to the appearance of the counter by pressing the spacebar (i.e., performing the psychomotor vigilance task (PVT); (Thomann et al., 2014)). These conditions are referred to as movie attending (MA) and task attending (TA), respectively. In the TA condition, PVT performance was determined by the mean reaction time after appearance of the counter.

To gain insight in the effect of the attentional instructions on the selective attentional performance of participants and to relate attentional performance to physiological synchrony, participants in both conditions were asked to answer 10 questions about the narrative of the movie clips immediately after each movie clip. These questions and the correct answers (in Dutch) can be found in the supplementary Table S2., which can be found online at: https://www.mdpi.com/article/10.3390/s23063006/s1.

7.2.4. Analysis

7.2.4.1. Pre-processing

All data and scripts are available on https://github.com/ivostuldreher/robustness-of-physiological-synchrony. Data processing was performed using MATLAB 2021a software (Mathworks, Natick, MA, USA). Both EDA and HR were time-locked to the onset of the movie clips using locally saved timestamps. Before starting the experiment, we synchronized device clocks by using the same online clock on all devices. EDA was imported using proprietary scripts using the unisens4matlab toolbox. We followed the pre-processing procedure of recent work using the same sensor (Thammasan et al., 2020). Periods of signal loss due to loose electrodes were identified and marked for removal later on in the process. The signal was filtered using a 3-second Savitzky-Golay filter to overcome the quantization noise in the signal. The fast changing phasic component and slowly changing tonic component of the EDA were then separated using Continuous Decomposition Analysis as implemented in the Ledalab toolbox for MATLAB (Benedek and Kaernbach, 2010). In further analysis, we use the phasic component of EDA as this component is mainly related to responses to external stimuli. With EDA we will from now on refer to the phasic component of EDA. After decomposition of the signal, we removed those parts of the signal that were marked for removal earlier on. If more than 30% of data recorded from an individual were marked for removal, the entire recording was removed for further analysis. Based on this criterion data of three out of the thirty participants were removed from further analysis. From three additional participants, part of the EDA data were removed as they contained periods marked as artifactual. From these participants 0.8%, 6.0% and 10.8% of data were removed. In total, we thus used datasets of 27 individuals in further analysis, of which for three participants 0.8%, 6.0% and 10.8% of the EDA data were removed.

The HR data were exported from the Wahoo Fitness application in a .fit format and imported in MATLAB using proprietary scripts based on FIT SDK 21.38.00 (FIT SDK RELEASE NOTES, 2020). Data of one participant was lost due to a failed recording. Suspicious samples in the data were removed based on two criteria: 1) samples higher than 200 bpm or lower than 30 bpm because we consider such values unrealistic in the current setting and 2) samples more than 25% different from the value 1 second before such changes in HR are considered unrealistic (Cheung, 1981). This was only true for one participant, for whom 0.3% of data were removed. Data of participants where samples equal to the previous sample are 50 times more prevalent than samples different from the previous sample were also removed from further analyses, as they indicate a malfunctioning HR sensor. This affected six participants. HR data of three participants were completely removed from further analyses, for the three other participants data recorded during two of the six movie clips were removed. In sum, the HR data of 26 participants were used in further analysis, of which for three participants data of two out of six movie clips was discarded.

EDA and HR data were then epoched and time-locked to the start of each movie clip as further analysis are conducted on the physiological data recorded during the movie presentation. Data were time-locked to the movie based on the computer time at the start of each movie presentation, that were saved to a .csv file and cut to the duration of each movie.

7.2.4.2. Inter-subject correlations

We computed physiological synchrony between participants during each movie clip using inter-subject correlations, following methods used in earlier work (Stuldreher et al., 2020b, 2020a; Pérez et al., 2021). Dyadic inter-subject correlations were computed for all unique dyads. Inter-subject correlations were computed in 15 second windows sliding at 1 second increments over the entire epoch of interest (here the entire movie clip), that were subsequently averaged across the entire epoch. For each physiological measure and each epoch, we thus obtain an N×N matrix of inter-subject correlations, where N represents the number of participants. For each participant, participant-to-group inter-subject correlations were computed by averaging over all values in a row, excluding the diagonal.

To investigate whether the movies and attentional instruction affected inter-subject correlation, a two-way ANOVA with independent variables movie and attentional condition was performed.

7.2.4.3. Circular-shuffle based significance test

To test the significance of participant-to-group inter-subject correlation values, we used the circular shuffle statistic, following (Pérez et al., 2021). Each participant's physiological signal was circular shifted by a random amount within the epoch length. The inter-subject correlations and participant-to-group inter-subject correlations were then computed with this circular shuffled data. This procedure was repeated 50 times for each participant. Statistical significance of the inter-subject correlations was assessed using a one-sided independent sample t-test comparing the actual correlations to those resulting from the shuffled data (significance threshold $p < .05$).

7.2.4.4. Effect of stimulus duration and group size on inter-subject correlation significance

Our aim is to investigate how stimulus duration and group size affect the inter-subject correlation significance.

To investigate the effect of stimulus duration, we artificially created shorter and longer movie clips. The first step was to combine the six 10-minute movie clips after each other into 60-minute movie clips. 720 different orders of combining the 10-minute movie clips are theoretically possible, namely six factorial. Due to computational constraints we chose to combine the movie clips in six different

7

orders based on the following Latin square, where each row corresponds to one of the six orders used:

1	2	3	4	5	6
2	5	4	6	1	3
5	1	6	3	2	4
3	4	5	1	6	2
6	3	2	5	4	1
4	6	1	2	3	5

The second step was to only select the first x minutes of data, where we varied x: 30 seconds, 1 minute, then in 2 minute increments up to 19 minutes and then from 24 minutes in 4 minute increments to 60 minutes. For each subsegment of data, we performed the abovementioned circular shuffle analysis to investigate the statistical significance of the inter-subject correlations.

To investigate the effect of group size on inter-subject correlations, we varied the group size from two to 27 participants, at one participant increments. As there are many subsets of participants possible, for each group size we selected 50 random subsets of participants using the 'randsample' function as implemented in Matlab. For each subset of participants we investigated for all of the above subsegments of data which fraction of participants showed significant synchrony.

7.3. Results

7.3.1. Effect of movie and condition on inter-subject correlation

To investigate whether the movies and attentional instruction affected inter-subject correlation, a two-way ANOVA with independent variables movie and attentional condition was performed. There was no effect of attentional instruction on inter-subject correlations in HR ($F(1,5) < 10^{-4}$, $p = .979$) and EDA ($F(1,5) = 0.15$, $p = .698$). While there was a main effect of movie on inter-subject correlations, both for HR ($F(1,5) = 7.57$, $p < .001$) and EDA ($F(1,5) = 3.79$, $p < .003$), there was no interaction with attentional instruction (HR: $F(1,5) = 0.88$, $p = .498$, EDA: $F(1,5) = 1.25$, $p = .286$). In the remaining analyses we therefore collapse over the two conditions.

7.3.2. Significance of inter-subject correlations

We first investigated the significance of inter-subject correlations in response to each of the movie clips separately. Figure 7-1 depicts the HR and EDA participant-to-group inter-subject correlations for each participant and each movie clip compared to a circular-shuffle-based chance level distribution. The figure illustrates the effect of movie clip on inter-subject correlation as described in the

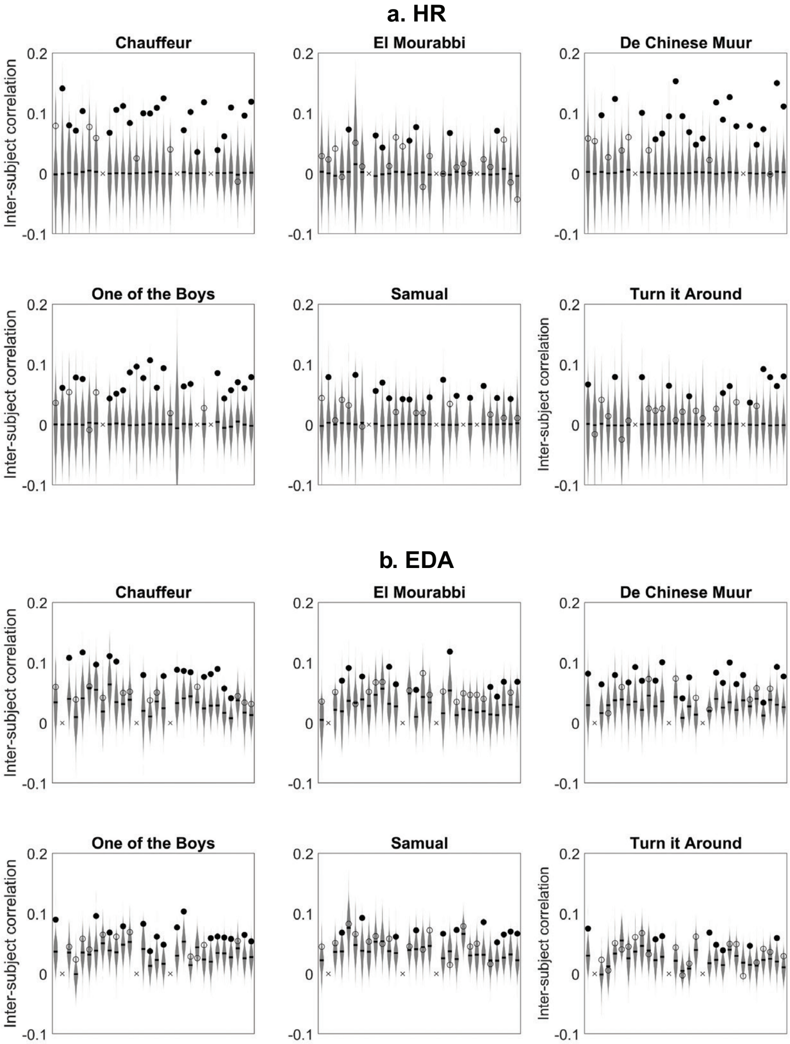

Figure 7-1. Participant-to-group physiological synchrony for each participant's a. HR and b. EDA. In each window, each marker refers to a participant. Filled markers depict inter-subject correlations exceeding chance level correlations based on 500 trials of circular shuffle (depicted by the grey distributions), open markers depict inter-subject correlations not exceeding chance level. Crosses depict missing data.

previous section. The majority of participants show significant HR inter-subject correlations in response to the movies "Chauffer", "De Chinese Muur" and "One of the Boys" while this is not the case for "El Mourrabi", "Samuel" and "Turn it Around". Inter-subject correlations in EDA show a similar pattern.

7.3.3. Dependency of significant inter-subject correlation on stimulus duration and group size

We then investigated how the percentage of participants with significant inter-subject correlations varied with stimulus duration and group size. Figure 7-2 shows how the percentage of participants with significant inter-subject correlations varies with stimulus duration and group size, for HR (top plots) and EDA (bottom plots) when averaged over the six movie orders and the fifty random subsets of participants. The left plots depict how the percentage varies with stimulus duration, for different group sizes. The right plots depict how the percentage varies with group size, for different stimulus durations. Additionally, Figure 7-3 shows the standard deviation of this percentage of participants with significant inter-subject correlations across the different movie orders as a function of stimulus duration in the left plots and across the fifty random subsets of participants as a function of group size in the right plots.

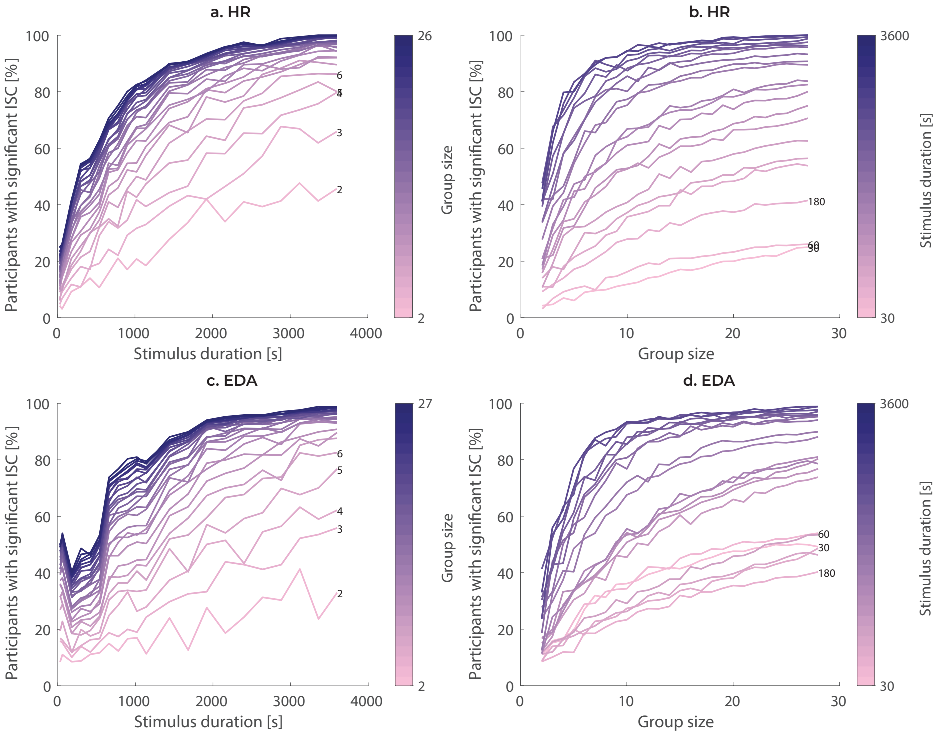

Figure 7-2. Percentage of participants with significant inter-subject correlations for HR (top; a, b) and EDA (bottom; c, d) as a function of stimulus duration (left; a, c) and participant group size (right; b, d), averaged over the six movie orders and 50 subsets of participant combinations. The color of the lines refers to the group size in the left plots and to the stimulus duration in the right plots, as do the numbers on the right side of some of the lines.

In Figure 7-2, the left plots show a generally increasing percentage of participants with significant inter-subject correlations with increasing stimulus duration, for all group sizes and for both EDA and HR. A small exception occurs in EDA when including only the first 30 seconds or one minute of data in the analysis; the percentage of participants with significant inter-subject correlations is actually higher for these stimulus durations than when including two to four minutes of data. The left plots in Figure 7-3 show that the percentage of participants with significant inter-subject correlations is strongly affected by the specific movie clip, depicted by the relatively large standard deviations for short stimulus durations (including only a single movie clip).

The right plots in Figure 7-2 show a generally increasing percentage of participants with significant inter-subject correlations with increasing group size, for all stimulus durations and for both EDA and HR. We observe that a larger group size leads to less dependence on the specific sample of participants, indicated by the decreasing standard deviation for increasing group size as depicted in

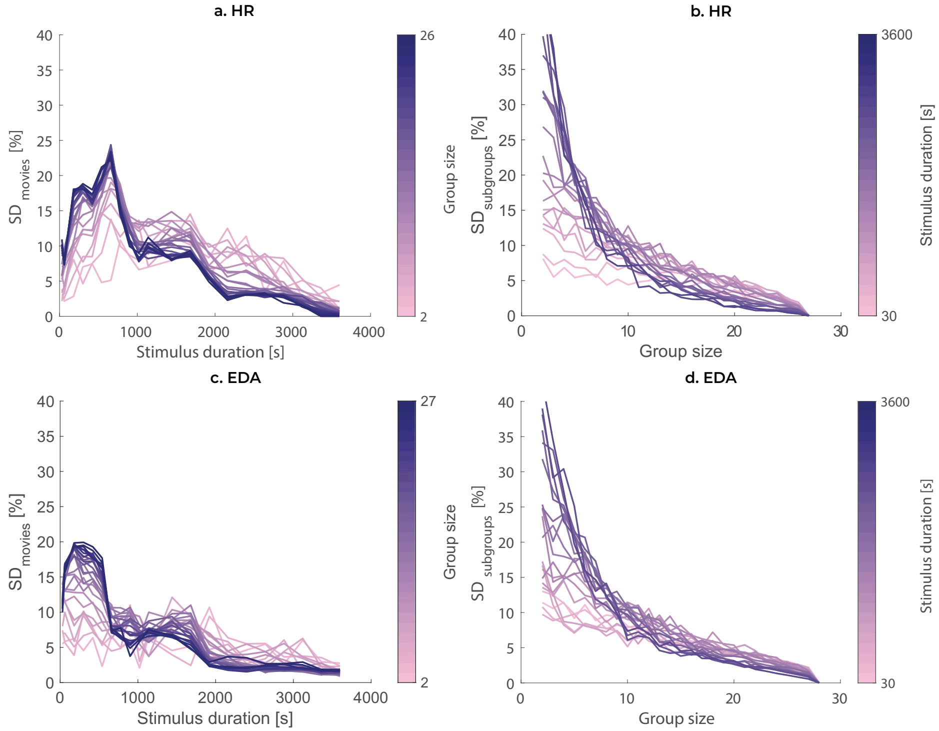

Figure 7-3. Standard deviation across movies (left a, c) and subgroups (right; b, d) of the percentage of participants with significant inter-subject correlations for HR (top; a, b) and EDA (bottom; c, d) as a function of stimulus duration (left; a, c) and participant group size (right; b, d). The color of the lines refers to the group size in the left plots and to the stimulus duration in the right plots.

7

the right plots of Figure 7-3.

7.3.4. Comparing effects of stimulus duration and group size on significant inter-subject correlation

The previous section describes two methods of increasing the total amount of data used (more movies – stimulus duration, and more participants – group size) and show the expected pattern that more data leads to a higher percentage of participants with significant inter-subject correlations. This raises the question if it is wise to increase the number of participants or the length of the stimulus if the option is there. Therefore, we expressed both methods of varying the amount of data in terms of total duration of included data. For instance, a group size of 10 participants and stimulus duration of 40 minutes leads to a total duration of included data of 400 minutes and a group size of 20 participants and stimulus duration of 20 minutes also leads to a total duration of included data of 400 minutes. Figure 7-4 shows the effects of both stimulus duration and group size in terms of total amount of data included as expressed in minutes for HR (top) and EDA (bottom). Each dot in the plot refers to the average percentage of participants with significant inter-subject correlations for a specific stimulus duration and group size. The color of the dots varies with group size, the size of the dots varies with stimulus duration. The shaded areas depict the standard deviation across the fifty random subsets of participants and six movie orders. The dots close in total amount of data included are also close in the percentage of participants with significant inter-subject correlations; dots are all on the same curve. The graphs thus show that while increasing the amount of data increases the percentage of participants showing significant inter-subject correlations, it does not seem to matter whether the amount of data is increased by increasing the stimulus duration or the number of participants.

7.3.5. Correlation between performance measures and inter-subject correlation

Last, we investigated whether the inter-subject correlation correlate with performance on questions about the content of the movie clips. Correlations between inter-subject correlations and performance on questions about the content of the movie clips are non-significant, as indicated in Table 7-2. However, when collapsing over all movies, there is a significant correlation between HR participant-to-group synchrony and number of correct movie questions (last row in Table 7-2).

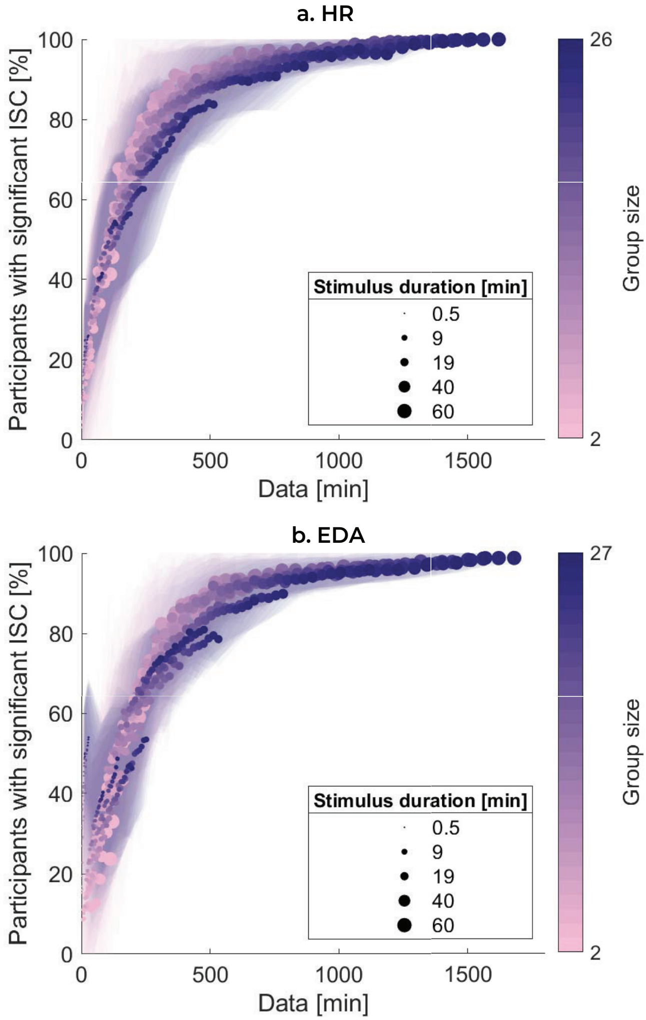

Figure 7-4. Fraction of participants with significant inter-subject correlations for a. HR and b. EDA as a function of the total minutes of data included in analysis, varied through varying stimulus duration and group size. For each datapoint, the stimulus duration is reflected by the marker size and the group size is reflected by the marker color.

7.4. Discussion

7.4.1. Significance of inter-subject correlations for different stimuli

In the current work we investigated the robustness of physiological synchrony in EDA and HR when presented with narrative stimuli. This is of interest because physiological synchrony can reflect shared attentional engagement. A first premise for a robust relation between physiological synchrony (assessed through inter-subject correlations) and attention is significant inter-subject

Table 7-2. Spearman correlations between number of correct answers on questions about the content of the movie and participant-to-group physiological synchrony

Movie	HR	EDA
Chauffeur	r = .13, p = .522	r = .02, p = .905
El Mourabbi	r = .14, p = .502	r = .14, p = .490
De Chinese Muur	r = .27, p = .171	r = .12, p = .541
One of the boys	r = .12, p = .550	r = .18, p = .341
Samual	r = .06, p = .751	r = .12, p = .571
Turn it around	r = .20, p = .320	r = -.08, p = .659
Overall	**r = .20, p = .010**	**r = .09, p = .234**

correlations. That is, inter-subject correlations that are higher than one would expect based on chance. We found that inter-subject correlations depend on the specific movie stimulus. For some movies 78% of participants show significant participant-to-group inter-subject correlations, while for others only 30% of participants show significant inter-subject correlations. We do not know what movie characteristics cause this difference. It may be so that some movies were more engaging than others. It may also be that low-level characteristics drawing attention in a bottom-up fashion were more prevalent in some movies than in others. Inter-subject correlations are known to be affected by low-level characteristics of the movie that draw attention in a bottom-up fashion (Golland et al., 2014; Steiger et al., 2019; Stuldreher et al., 2020a). Previously, we found that moments of high synchrony when listening to an audiobook did not correspond with scenes identified as overall 'attentionally engaging' by an independent group of listeners. Instead, we suggested that moments of high synchrony corresponded with relatively low-level engaging moments, such as swear words or salient intonation (Stuldreher et al., 2020a). With the use of EEG, it has been reported that moments of high inter-subject correlations corresponded to short suspenseful moments in the movie (Dmochowski et al., 2012; Poulsen et al., 2017). Though the present study did not show an effect of attentional task (either performing the PVT or not), several studies, including (Stuldreher et al., 2020b), demonstrate the effect of top-down guided attention on physiology. Such findings refute the idea that physiological synchrony is completely caused by involuntary, bottom-up drawn attention. However, the exact relation between inter-subject correlations in HR and EDA and experienced attentional engagement is not clear yet.

7.4.2. Significance of inter-subject correlations with varying amount of data included

Our main question was how the prevalence of significant inter-subject correlations depends on the duration of the stimulus and on the size of the participant group. As expected, with increasing stimulus duration and increasing group size, the percentage of participants that show significant inter-subject correlations also increases. When using data of all participants, but stimulus durations of up to ten minutes, the percentage of significant inter-subject correlations is movie dependent and low on average (< 50%). Aggregating over two or three ten-minute movies already results in fairly robust significance levels of above 80%. Similarly, when using all 60 minutes of stimuli, for small group sizes up to 10 participants the percentage of significant inter-subject correlations is dependent on the specific sample of participants and low on average. With larger group sizes, the fraction of significant inter-subject correlations is high on average (> 85%) and less dependent on the specific sample.

It appears that it does not matter in which way the total amount of data is reached. Robust inter-subject correlations (80% of participants with significant inter-subject correlations) is reached with a total amount of data of roughly 10 hours for both HR and EDA. It does not matter whether this total amount of data is reached by increasing the number of participants or by increasing the stimulus duration. Note that for small group sizes the percentage of participants with significant synchrony strongly depends on the specific sample of participants (see Figure 7-3). We therefore suggest researchers conducting similar studies to include at least 10 participants.

A small exception occurs in EDA when including only the first 30 seconds or first minute of each movie stimulus in the analysis. The significance of inter-subject correlations is actually higher with only the first minute of data than when including the first two up to the first four minutes of data (Figure 7-2, bottom left panel). When examining how EDA inter-subject correlations change over the course of the movies (see supplementary Figure S1), we observe high correlation values in the first 60 seconds of the each of the six movies. The figures in (Golland et al., 2014) show a similar pattern, with relatively high inter-subject correlations in the first 60 seconds of the stimulus. We think this is related to the very high levels of EDA at the beginning of each stimulus movie, that are necessarily followed by a sharp drop (see bottom panel in supplemenary Figure S2). Such EDA patterns are commonly found in studies displaying stimuli to observers, and presumably related to arousal associated with the presentation of ‘something new’. Also note that HR shows a similar, be it more modest pattern of high values at the start of each movie, followed by a drop (top panel in supplementary Figure S2). We think arousal and attentional engagement are closely linked. For the special case of the start of a stimulus, high levels of EDA, and high inter-subject correlations, may reflect general stimulus-driven arousal and engagement, that is somewhat apart from attentional engagement with the stimulus’ content.

We cannot be sure whether our results are generalizable to similar studies or a finding specific for our sample of participants and videos. There is a very limited number of studies assessing inter-subject correlations in EDA or HR of observers, let alone report their significance. However, two recent studies assessing significance of inter-subject correlations in HR present results similar to ours. Pérez et al. (Pérez et al., 2021) report significance of the inter-subject correlations in HR of 27 participants presented with 16 one-minute audiobook fragments. When aggregating over all audiobook fragments, thus including 27 × 16 = 432 minutes of data in total, 63% of participants show significant inter-subject correlations. This point coincides with the curve in Figure 7-4, indicating the results presented here may be comparable to results found using other stimuli. Madsen and Parra (Madsen and Parra, 2022) obtain significant inter-subject correlations in HR for 66% participants, while using 920 minutes of data recorded during viewing of instructional videos in total. This reported prevalence of significance is on the lower end compared to our average results, but is still higher than the minimum percentage of participants with significant inter-subject correlations that we found for the same amount of data. These results suggest that the reported relation between amount of included data and inter-subject correlation significance applies to data beyond our specific set.

A premise for inter-subject correlations when presented with narrative stimuli is that individuals process narrative stimuli comparably (Madsen and Parra, 2022). While our sample of participants varied quite widely in age and contained about equal numbers of males and females, samples of participants that contain even more differing individuals may start to violate the premise of comparable processing. For instance, individuals with autism spectrum disorders (ASD), depression or first-episode psychosis, are known to show more varying neural patterns and thus reduced neural inter-subject correlations during naturalistic stimulus presentations (Hasson et al., 2009; Salmi et al., 2013; Guo et al., 2015; Mäntylä et al., 2018). When measuring inter-subject correlations across a group of individuals among which there are individuals with non-typical neural patterns, we expect that adding more data by including longer stimuli will result in a lower fraction of participants with significant inter-subject correlations than when adding more data by including more individuals with typical neural patterns. Conversely, when measuring inter-subject correlations across a group of ‘typical’ individuals, adding more data by using a longer stimulus is expected to lead to a higher fraction of participants that show significant synchrony than when adding ‘non-typical’ participants.

7.4.3. Inter-subject correlations as measure of attentional engagement

For validating inter-subject correlation as a measure of attentional engagement, and with an eye on potential application of this marker, a demonstrated relation between inter-subject correlation and behavioral, attentional outcomes is important. For individual movie clips, we did not find significant correlations between participant-to-group inter-subject correlations and correctly answered movie questions in this study. When aggregating over all six movie clips, the correlation was significant for HR ($r = .20$, $p = .010$), though not for EDA ($r = .09$, $p = .234$). These results are in line with higher proportions of participants showing significant inter-subject correlation for HR than for EDA. It indicates that also for a robust relation between inter-subject correlations and attention, sufficient amounts of data are needed.

The finding presented here is consistent with our previous work, in which inter-subject correlations in HR were predictive of selective attentional performance, but inter-subject correlations in EDA were not (Stuldreher et al., 2020b). The correlations between inter-subject correlations and number of correctly answered questions presented here are low ($r = .20$ when aggregated across movies). Inter-subject correlations thus only explain a small part of the variance in the answers to questions about the content of the video. Note that the strength of the relation between attentional engagement (as estimated through inter-subject correlation) and performance strongly depends on the sensitivity of the behavioral measures. E.g., when questions are used that are too easy or too hard, an association will not be found. Furthermore, whereas answers to questions on the movie contain information on content-level attention, inter-subject correlations most likely also capture more low-level attentional characteristics, driven among others by stimulus saliency (Stuldreher et al., 2020a) and emotional engagement (Golland et al., 2014; Steiger et al., 2019).

7.4.4. Is higher inter-subject correlations always better?

In the current work we tested how participant group size and stimulus duration influence inter-subject correlations and their significance, with the goal of maximizing the fraction of participants with significant inter-subject correlations. This does not mean that higher inter-subject correlation is always better from the point of view of its potential value as a measure of attentional engagement. In fact, in the theoretical case that physiological signals of all participants are perfectly synchronized, the inter-subject correlations would lose their value since there is no physiological variability to relate to behavioral variability (Hedge et al., 2018). However, if the inter-subject correlations do not exceed chance level because of a lack of data, any differences in attentional engagement due to an intervention or personal differences will also not be visible. An important question is where the sweet spot lies where a stimulus evokes just enough synchrony to obtain significant inter-subject correlations, but not enough to saturate the individual signals of interest (Finn et al., 2020). This sweet-spot may also be context dependent. For instance, when evaluating how effective a certain stimulus is in attracting attention compared to a second stimulus, the inter-individual differences are of less interest than when investigating which group of participants shows most engagement with a certain stimulus.

7.4.5. Conclusion

This study explored how the amount of data included in analyses influences the fraction of participants that show significant inter-subject correlations and showed that the source of the data (i.e. more participants or longer stimuli) is irrelevant. Future research should investigate the generalizability of this finding over different sets of stimuli and participants. We hope our results can help guide future researchers when setting up studies on narrative driven inter-subject correlations. Through our work we hope to contribute to the shift in research from controlled laboratory settings and high-end equipment, to real-life settings and wearable sensors.

Acknowledgments

We would like to thank Mannes Poel for his valuable input on the discussion. This work was supported by TNO's Early Research Program Body-Brain Interaction.

8.

Unsupervised clustering of individuals sharing selective attentional focus using physiological synchrony

Stuldreher, I.V., Merasli, A., Thammasan, N., van Erp, J.B.F., Brouwer, A.M. (2022). Unsupervised clustering of indiviuals sharing selective attentional focus using physiological synchrony. *Frontiers in Neuroergonomics*, **2**, 750248. doi: 10.3389/fnrgo.2021.750248

Abstract

Research on brain signals as indicators of a certain attentional state is moving from laboratory environments to everyday settings. Uncovering the attentional focus of individuals in such settings is challenging because there is usually limited information about real-world events, as well as a lack of data from the real-world context at hand that is correctly labeled with respect to individuals' attentional state. In most approaches, such data is needed to train attention monitoring models. We here investigate whether unsupervised clustering can be combined with physiological synchrony in the electroencephalogram (EEG), electrodermal activity (EDA), and heart rate to automatically identify groups of individuals sharing attentional focus without using knowledge of the sensory stimuli or attentional focus of any of the individuals. We used data from an experiment in which 26 participants listened to an audiobook interspersed with emotional sounds and beeps. Thirteen participants were instructed to focus on the narrative of the audiobook and 13 participants were instructed to focus on the interspersed emotional sounds and beeps. We used a broad range of commonly applied dimensionality reduction ordination techniques - further referred to as mappings - in combination with unsupervised clustering algorithms to identify the two groups of individuals sharing attentional focus based on physiological synchrony. Analyses were performed using the three modalities EEG, EDA, and heart rate separately, and using all possible combinations of these modalities. The best unimodal results were obtained when applying clustering algorithms on physiological synchrony data in EEG, yielding a maximum clustering accuracy of 85%. Even though the use of EDA or heart rate by itself did not lead to accuracies significantly higher than chance level, combining EEG with these measures in a multimodal approach generally resulted in higher classification accuracies than when using only EEG. Additionally, classification results of multimodal data were found to be more consistent across algorithms than unimodal data, making algorithm choice less important. Our finding that unsupervised classification into attentional groups is possible is important to support studies on attentional engagement in everyday settings.

8.1. Introduction

Research on brain signals as indicators of mental state, such as attention, is moving from laboratory environments to everyday settings. This comes with several challenges. Firstly, contextual information about the environment and the people acting in it is limited. It is, for instance, usually unknown what events occur in the environment that are of potential interest to individuals. This complicates the process of uncovering attentional state by referring to known events through traditional analysis of event-related brain potentials. Secondly, everyday settings make it difficult to acquire suitable data to train algorithms that uncover mental state. Machine learning techniques have increased our ability to uncover complex mental states even with limited contextual information, but user-specific data from a similar context is required to train well-performing machine learning models. In a supervised machine learning approach, a model is trained with data recorded when information was available about events, and about the mental state of the individuals, to enable discrimination between the mental states of interest for unseen data collected when contextual information is limited. Such paradigms have been widely applied, for instance to recognize the emotional response to videos (Soleymani et al., 2012, 2016), to distinguish between different mental workload conditions (Hogervorst et al., 2014) or to estimate the attentional state of individuals (Abiri et al., 2019; Vortmann et al., 2019). The requirement of context-specific training for discrimination between mental states is the major drawback of supervised machine learning (Arico et al., 2018). Especially in everyday settings, the ground truth mental state information needed in the training phase is often not available (Brouwer et al., 2015b).

We here focus on further exploring an alternative approach that requires little information about the individuals' environment and does not require training. This approach is based on the interdependence of physiological signals in groups of individuals and may be used to probe attentional engagement. A number of everyday settings exist in which groups of individuals share their attention to some degree. An example is a group of students listening to the instruction of a teacher in a classroom. The degree to which physiological signals in such groups of individuals uniformly change is often referred to as physiological synchrony (Palumbo et al., 2017). It has been related to the attentional engagement of individuals in a group, for example when presented with the same narrative stimulus, such as a movie or audio clip (Hasson et al., 2010; Dmochowski et al., 2012). Ki et al. (2016) found that when a participant was actively attending to a movie, his or her electroencephalogram (EEG) was more synchronized with the EEG of others attending to the same movie than when the participant's attention was focused inwardly on a mental arithmetic task. Perez et al. (2021) found similar results when using heart rate instead of EEG. Stuldreher et al. (2020b) found that physiological synchrony in EEG, heart rate and electrodermal activity (EDA) could not only distinguish between different attentional conditions within an individual, but could also distinguish between participants who had received different selective attentional instructions toward the exact same external stimulus. That is, a majority of participants showed more physiological synchrony with others who received the same attentional instructions than with others who received opposite attentional instructions.

Previous work indicates that both similarities in emotional and cognitive processing may underlie physiological synchrony across individuals. Poulsen et al. (2017) found that moments in time with high physiological synchrony often coincided with emotionally arousing scenes of presented movie clips, suggesting that emotional engagement underlies high physiological synchrony. Stuldreher et al. (2020a) showed that not only presentation of emotionally arousing sounds led to high physiological synchrony, but also the presentation of to-be-counted beeps, suggesting shared cognitive processing can also underlie high physiological synchrony. Dmochowski et al. (2014) showed that physiological synchrony over time was predictive of the number of tweets and viewership during a popular television series, where emotional and/or cognitive engagement may have resulted in being compelled to view the stimulus. The contribution of shared emotional or cognitive processing of specific stimuli to the overall interpersonal physiological synchrony seems to depend on the specific physiological measure. Stuldreher et al. (2020a) found that moments of high physiological synchrony in EEG corresponded with the occurrence of cognitive processing, but not with emotionally arousing events. Moments of high physiological synchrony in heart rate, on the other hand, corresponded well with emotionally arousing events, but not with cognitive processing. Nonetheless, physiological synchrony in all of the above measures was shown to distinguish between groups with different selective attentional focus toward the same narrative stimulus (Stuldreher et al., 2020b).

Physiological synchrony thus enables monitoring the degree of attentional engagement without training of a model, and without detailed information about the environment. However, researchers up to now have only identified the specific attentional focus of an individual by putting physiological synchrony in context of other individuals of whom the attentional focus is known, such as inwardly vs. outwardly focused attention (Cohen and Parra, 2016; Ki et al., 2016; Pérez et al., 2021) or one of two specific types of selective attentional instructions (Stuldreher et al., 2020b). In everyday settings, such knowledge is not always available. For example, it is not known a priori who out of a group of students are attending to key elements of the lecture, and who are attending to what is happening in the classroom around them, which would be required to classify an unknown individual into one or the other attentional group following the earlier used methods. For such cases, we require unsupervised identification of groups of individuals sharing attentional focus.

Unsupervised learning techniques may be used to find clusters of individuals sharing attentional state. Unlike supervised learning, unsupervised learning techniques are not based on a model that is trained on a labeled dataset. Instead, these techniques form clusters of samples that are proximate in a high-dimensional space (Grira et al., 2004). Numerous algorithms are available, from well-known algorithms such as traditional k-means (Lloyd, 1982), and its more modern iterations (Yu et al., 2018; Sinaga and Yang, 2020), to spectral clustering (Von Luxburg, 2007) or hierarchical clustering (Ward, 1963). Complementary to data clustering are ordination techniques, that pre-order objects in such a way so that similar objects are close to each other and dissimilar objects are far away from each other. Often used are the algorithms that are part of the family of multidimensional scaling (Borg et al., 2018).

Unsupervised learning techniques have been explored before in research using physiological measures to assess mental state. For instance, Schultze-Kraft et al. (2016) successfully employed unsupervised learning techniques to classify either low or high operator workload in a laboratory setting based on EEG signals. Existing work focuses on within-subject classification of mental state (Carreiras et al., 2016; Schultze-Kraft et al., 2016; Maaoui and Pruski, 2018). To the best of our knowledge, unsupervised clustering of individuals sharing their attentional focus has not been demonstrated before.

The goals of the current work are therefore two-fold. First, we establish the feasibility of unsupervised clustering of individuals based on physiological synchrony, to automatically identify groups of individuals sharing attentional focus without pre-knowledge of attentional focus of any of the individuals. Clustering performance is evaluated by using ground truth information on attentional state. Second, we investigate how performance depends on the type of physiological measure used. While distinguishing between different attentional conditions using synchrony in EEG, EDA, and heart rate has been explored before (Stuldreher et al., 2020b), we do not know how such results transfer to an unsupervised approach. Additionally, we test performance when multiple physiological measures are combined. We investigate all of this with the use of a broad range of classic and more modern unsupervised learning techniques. A secondary goal therefore is to compare clustering performance across algorithms.

In this study, we use the (Stuldreher et al., 2020b) publicly available dataset (https://osf.io/8kh36/) in which ground-truth information about the attentional state of individuals is available. Though such information is generally not available in everyday—and if it is, one would use supervised learning techniques due to their higher performance compared to unsupervised learning (Schultze-Kraft et al., 2016; Arico et al., 2018) – we here need the ground truth information to reflect on the performance of this novel approach. We also investigate the use of the silhouette coefficient as a potential way to evaluate unsupervised clustering performance in scenarios where no ground-truth information is available (Rousseeuw, 1987).

In sum, we investigate whether attentional focus can be determined using unsupervised clustering, and if so, whether clustering performance depends on the type of physiological modality (EEG, EDA, and heart rate).

We hypothesize that:

1) Attentional focus can be determined using unsupervised clustering techniques.

2) Classification accuracies are higher when using EEG rather than EDA or heart rate.

3) Combining modalities into a multimodal approach leads to higher classification accuracies than unimodal approaches because a multimodal approach includes information of more mental processes in the classification decision.

4) The silhouette coefficient is correlated with clustering accuracy.

When testing these hypotheses, we use multiple clustering algorithms. An additional exploratory research question is how performance depends on clustering algorithm.

8.2. Methods

8.2.1. Participants

Twenty-seven participants (17 female), aged between 18 and 48 years (M = 31.6, SD = 9.8 years), took part in the experiment. They were recruited through the participant pool of the research institute where the study took place. None of the participants reported problems with hearing. Prior to the experiment all participants signed an informed consent, in accordance with the Declaration of Helsinki. All participants received a small monetary compensation for their participation in the experiment and for traveling costs. Data from 26 participants were further processed due to a recording failure in one case. The experiment was approved by the TNO Institutional Review Board. The approval is registered under reference 2018-70.

8.2.2. Materials

Electroencephalogram, EDA, and electrocardiogram (ECG) were recorded at 1,024 Hz using an ActiveTwo Mk II system (BioSemi, Amsterdam, Netherlands). Electroencephalogram was recorded with 32 active Ag/AgCl electrodes, placed on the scalp according to the 10–20 system, together with a common mode sense active electrode and driven right leg passive electrode for referencing. The electrode impedance was maintained below 20 kOhm. For EDA, two passive gelled Nihon Kohden electrodes were placed on the ventral side of the distal phalanges of the middle and index finger. For ECG, two active gelled Ag/-AgCl electrodes were placed at the right clavicle and lowest floating left rib. Electrodermal activity and heart rate were also recorded using wearable systems (Movisens EdaMove 4 and Wahoo Tickr, respectively). These data are not discussed here, but are publicly available on https://osf.io/8kh36/ and compared to the data recorded using the ActiveTwo in van Beers et al. (2020).

8.2.3. Stimuli and design

Participants performed the experiment one by one. Each participant listened to the same audio file, composed of a 66 min audiobook (a Dutch thriller "Zure koekjes," written by Corine Hartman) interspersed with other auditory stimuli. The 13 participants in the audiobook attending (AA) group were asked to focus on the narrative of the audiobook and ignore all other stimuli or instructions. The 13 participants in the stimulus attending (SA) group were asked to focus on the other stimuli, perform accompanying tasks, and ignore the audiobook. The order of interspersed stimuli was randomly determined, but was identical for each participant. Intervals between the end of one stimulus and the onset of the next one varied between 35 and 55 s (M = 45, SD = 6.1 s). The short auditory stimuli were affective sounds, blocks of beeps, and the instruction to sing a song. For the exact types and order of interspersed stimuli we refer the reader to Stuldreher et al. (2020b).

After the experiment, all participants were asked to answer two questionnaires. In the first questionnaire, participants used a slider on a horizontal visual analog scale running from "not at all" to "extremely" to rate their mental effort, distraction and emotion during the short emotional sounds. The second questionnaire was on the content of the stimuli: participants were asked to report as many emotional sounds as they could remember, they were asked to estimate the average number of beeps in a block, and they were asked questions about the content of the narrative. For more details we refer the reader to Stuldreher et al. (2020b).

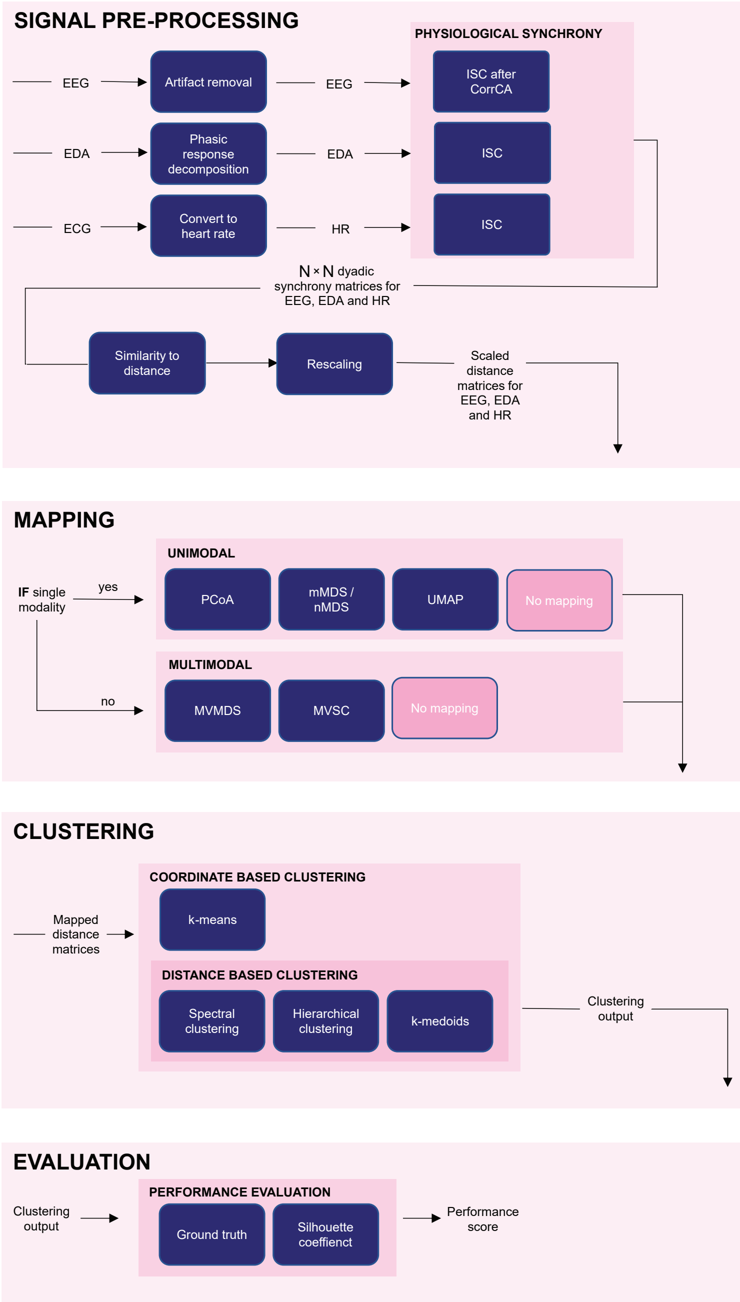

Figure 8-1. Overview of the processing pipeline, divided in signal pre-processing, mapping, clustering, and evaluation. EEG, electroencephalogram; EDA, electrodermal activity; ECG, electrocardiogram; HR, heart rate; PS, physiological synchrony; PCoA, principle coordinate analysis; mMDS, metric multidimensional scaling; nMDS, non-metric multidimensional scaling; UMAP, uniform manifold approximation and projection; MVMDS, multiview multidimensional scaling; MVSC, multiview spectral clustering.

8.2.4. Analysis

An outline of the complete analysis is depicted in Figure 8-1. In the sections below, each part of the analysis is elaborated upon separately.

8.2.4.1. Signal pre-processing

Data pre-processing was done using MATLAB 2019a software (Mathworks, Natick, MA, USA). Electroencephalogram was pre-processed using EEGLAB v14.1.2 for MATLAB (Delorme and Makeig, 2004). To remove potentials not reflecting sources of neural activity, like ocular or muscle-related artifacts, logistic infomax independent component analysis (Bell and Sejnowski, 1995) was performed. Electroencephalogram was first downsampled to 256 Hz and high-pass filtered with the passband edge at 1 Hz using the standard finite-impulse-response filter implemented in EEGLAB function pop_eegfiltnew. This relatively high cut-off frequency has shown to work better for independent component analysis compared to lower cut-off frequencies (Winkler et al., 2015). Data were then notch filtered at 50 Hz, again using the standard finite-impulse-response filter implemented in EEGLAB function pop_eegfiltnew. Channels were re-referenced to the average channel value. Independent component analysis was performed and the Multiple Artifact Rejection Algorithm (Winkler et al., 2011) was used to classify artifactual independent components. Components that were marked as artifactual were removed from the data. Then, samples whose squared amplitude magnitude exceeded the mean-squared amplitude of that channel by more than four standard deviations were marked as missing data ("NaN") in an iterative way with four repetitions to remove outliers. By doing so, 0.8% of data were marked as missing.

Electrodermal activity was downsampled to 32 Hz. The fast changing phasic and slowly varying tonic components of the signal were extracted using Continuous Decomposition Analysis as implemented in the Ledalab toolbox for MATLAB (Benedek and Kaernbach, 2010). In further analyses we use the phasic component, as this component of the EDA signal is mainly related to responses to external stimuli (Boucsein, 2012).

Electrocardiogram measurements were processed to acquire the inter-beat interval (inversely proportional to heart rate). After downsampling to 256 Hz, ECG was high-pass filtered at 0.5 Hz. R-peaks of the ECG signal were detected following Pan and Tompkins (1985), resulting in a semi-time series of consecutive inter-beat intervals. This inter-beat interval semi-time series was transformed into a time series by interpolating consecutive intervals at 32 Hz.

8.2.4.2. Physiological synchrony

We computed inter-subject correlations in the time-domain as a measure of physiological synchrony. Rather than treating EEG signals separately, the inter-subject correlations were evaluated in the correlated components of the EEG (Dmochowski et al., 2012, 2014). The goal of the correlated component analysis is to find underlying neural sources that are maximally correlated between participants, using linear combinations of electrodes. The technique is similar to the more familiar principal component analysis, differing in that projections capture maximal correlation between sets of data instead of maximal variance within a set of data. After computing the correlated components based on data from all 26 participants, EEG data of each participant were projected on the component vectors. Inter-subject correlations between a participant with all other participants were then computed as the sum of correlations in the first three component projections, as correlations in higher order projections are often close to chance level (Ki et al., 2016). The result is a N × N matrix inter-subject correlations of all possible pairs of participants. The correlation values were normalized by dividing all correlation values by the diagonal value - in this case three, as we computed physiological synchrony as the sum of correlations in the first three correlated components.

For EDA and heart rate, we also computed inter-subject correlations in the time-domain as a measure of physiological synchrony. Pearson correlations were calculated over successive, running 15 s windows at 1 s increments. The overall correlation between two participants was computed as the natural logarithm of the sum of all positive correlations divided by the sum of the absolute values of all negative correlations. Again the correlation matrices were normalized. Originally, the diagonal here contained infinite values (as there are no negative correlations, the denominator in the ratio is zero). We therefore chose to replace these cells with finite values in such a way that the ratio between the diagonal value and the mean of the matrix was the same for the matrices of EDA and heart rate as for EEG. Then again, all correlations were divided by the diagonal value.

Clustering algorithms usually require distance matrices. Thus, correlation matrices were transformed into distance matrices before applying clustering algorithms. Several transformations exist (Groenen and van de Velden, 2005). We followed the suggestion of Gower and Legendre (1986) and computed the values in the distance matrix as the square root of one minus the values in the correlation matrix.

As the off-diagonal correlation values were close to zero, and thus the off-diagonal distance values close to one, we applied a linear transformation of each off-diagonal coefficient like in interval multidimensional scaling (Borg and Groenen, 2005) to evenly distribute the values between zero and one.

8.2.4.3. Mapping

Various ordination methods, or "mappings," have been proposed to create distance matrices. Mapping is complementary to data clustering in such way that objects are ordered so that similar objects are close to each other and dissimilar objects are far away from each other. Mappings toward a different (lower-dimensional) space can be of value for visualization of clusters, and they can improve clustering performance (Kent et al., 1979). The dimension of the mapped space can be chosen arbitrarily. We chose the output mapping to be in two-dimensional space, which is most common in literature and easy to interpret. We applied different, commonly known mappings, of which an overview can be found in Table 8-1.

8.2.4.4. Clustering

After mapping, or skipping the mapping, we applied a range of classical clustering algorithms (Table 8-2). Not all combinations of mapping and clustering yielded valid results. Some methods, for example, are not deterministic, but provide different outcome maps for different initializations. We therefore used multiple random initializations and averaged over the clustering results for each initialization. However, this approach did not converge when using k-means on the raw distance matrices.

8.2.4.5. Evaluation: clustering quality assessment

To assess the clustering quality we compared found clusters to attentional condition labels (AA or short SA), so that the clustering performance can easily be assessed. To investigate whether clustering performance is above chance level, we conducted a permutation analysis with shuffled group-labels, so that we can compare the clustering accuracy to accuracies obtained for 100 trials with randomized group-labels. We determined the significance level using a one-tailed non-parametric Mann Whitney U-Test. Chance level distributions were determined for all algorithm combinations. The threshold for significantly higher clustering accuracies compared to chance were found to be either 65% (17 out of 26 participants correctly clustered) or 70% (18 out of 26 participants correctly clustered). We selected the strictest significance level (i.e., 70%) to compare all classification results to.

In real-world conditions, ground-truth information on the attentional state is often not a-priori available, which makes it hard to tell how well-unsupervised clustering of attentional states works in a particular condition. Therefore, we explored an alternative measure of evaluating clustering performance, known as the silhouette coefficient (Rousseeuw, 1987). This index measures the compactness and separation of clusters and may be informative as a confidence metric of the clustering outcome. A confident clustering outcome would be associated

Table 8-1. Overview of the used ordination techniques.

Method	Description	Reference
Principle Coordinate Analysis (PCoA)	Also referred to as classical multidimensional scaling, PCoA intends to preserve the distances in the distance matrix in the output mapping. To do so, for each participant the objective is to find coordinates in a lower dimensional space that minimize the strain with the original values.	(Groenen and Borg, 2014)
Metric Multidimensional Scaling (mMDS)	mMDS is a superset of the PCoA that generalizes the optimization procedure, where instead of strain often the stress is minimized. The minimization problem is solved iteratively as there exists no exact solution.	(Borg and Groenen, 2005)
Non-metric Multidimensional Scaling (nMDS)	Unlike PCoA, nMDS distorts the distances in the ordination solution. However, it preserves the rank of dissimilarities by minimizing the non-metric stress in an iterative approach.	(Kruskal, 1964)
Uniform Manifold Approximation and Projection (UMAP)	UMAP is a non-linear manifold learning technique originally developed as dimensionality reduction. It emphasizes local distances over global distances. As the UMAP algorithm is also able to deal with ground-truth information known about some of the data points, it can either be used as all other methods or with self-supervised learning (SSL). With SSL, at the algorithm initialization, no labels are known. When the first mapping and clustering are done, a known label is assigned to the participant which is the closest to one of the cluster center. This procedure is then repeated, each time adding the participant closest to one of the cluster centers that has not been labeled yet.	(McInnes et al., 2018)
Multiview Multidimensional Scaling (MVMDS)	Multi view dimensionality reduction solutions have emerged to solve problems where various samples of the same observation are collected, as is the case here with synchrony in EEG, EDA and heart rate. MVMDS is a multimodal extension of PCoA – with only one matrix as input, results are identical – that intends to find the common eigenvectors across the different distance matrices	(Trendafilov, 2010)
Multiview Spectral Clustering (mvSC).	Like MVMDS, this technique is a multimodal extension of the spectral clustering ordination. It computes the common eigenvectors of the Laplacian of the dissimilarity matrices.	(Kanaan-Izquierdo et al., 2018)

Table 8-2. Overview of used clustering algorithms.

Method	Description	Reference
k-means	Probably the best known clustering algorithm, k-means aims to partition all observations into clusters (here k = 2), in which each observation belongs to the cluster with the nearest mean. Solutions are found iteratively.	(Lloyd, 1982)
k-medoids	This adaptation of k-means is based on the same principle, but rather than minimizing the distance between data points and the cluster center, that is not necessarily one of the input data points, k-medoids chooses data points as centers and minimizes the distance between data points and this medoid.	(Bauckhage, 2015)
Spectral clustering	This technique makes use of the spectrum – or eigenvalues – of the similarity matrix to perform dimensionality reduction before clustering using traditional algorithms like k-means. Therefore, this algorithm somewhat combines a mapping and clustering algorithm in one.	(Von Luxburg, 2007)
Hierarchical clustering	As the name suggests, this algorithm builds a hierarchy of clusters in a bottom-up fashion. Initially, each data point thus belongs to its own cluster. The clusters are progressively merged according to similarity criteria called linkage. We here use Ward linkage that finds new clusters by minimizing the sum of squared differences within the merged clusters.	(Ward, 1963)

with tight and well-separated clusters, depicted by a silhouette coefficient near one, whereas an unconfident clustering outcome would be associated with broadly spread and overlapping clusters, depicted by a silhouette coefficient near zero. The silhouette coefficient cannot be determined for all combinations of mapping and clustering methods, for example when using random initializations before mapping over which has to be averaged, as can be de the case with nMDS, mMDS, or UMAP.

8.3. Results

8.3.1. Clustering performance using physiological synchrony in either EEG, EDA, or heart rate

A complete overview of clustering performance for all used combinations of mapping algorithms and clustering algorithms based on physiological synchrony in either EEG, EDA, or heart rate is presented in Supplementary Table A1. It shows the clustering accuracy, the misclassified participant IDs and silhouette coefficient, wherever available. Figure 8-2 visualizes the clustering accuracies across the eight mapping methods and the three clustering methods that could all be combined with each other, as well as results when using no mapping, for which we could determine results for two out of the three clustering methods. Classification accuracies above the black line are significantly higher than chance level. That is, classification accuracies of 70% or higher are significantly higher than the chance level distribution at $p < .05$. The best performance is obtained using physiological synchrony in EEG [Mdn = 73%, Inter Quartile Range (IQR) = 12% across algorithms], with a maximum clustering accuracy of 85% when using spectral clustering on the raw distance matrix or after applying PCoA ordination. For EDA, a median performance of 58% (IQR = 8%) was obtained; best EDA performance was reached using k-means with nMDS mapping (65%). For heart rate, median performance was 62% (IQR = 4%); best performance was reached using hierarchical clustering with nMDS or mMDS mapping (73%).

We determined the silhouette coefficient as a potential alternative measure of clustering performance. The results in Supplementary Table A1 do not suggest that a high silhouette coefficient corresponds with high clustering accuracy as evaluated using knowledge of the attentional instruction groups. This impression is confirmed by a lack of correlation between clustering accuracy and silhouette coefficient ($r = 0.06$, $p = .543$). We can note, however, that the silhouette coefficient is generally higher after mapping (Mdn = 0.33, IQR = 0.04) than without mapping (Mdn = 0.13, IQR = 0.13).

8.3.2. Clustering performance combining physiological synchrony in EEG, EDA, and heart rate

Supplementary Table A2 presents the clustering results when combining physiological synchrony in multiple modalities for all possible mapping-clustering combinations. It shows clustering accuracies, misclassified participant IDs and silhouette coefficient when combining EEG and EDA, EEG and heart rate, EDA and heart rate, and all three modalities. Figure 8-3 presents an overview of the accuracies for each mapping-clustering combination. Again, the dashed black line at 70% depicts significance level compared to chance. The best clustering performance of 92% is reached for EEG combined with heart rate when k-means and MVMDS or MVMDS-with-rescaling are used; as well as for the combination of EEG, heart rate, and EDA when MVMDS with rescaling is used with spectral or hierarchical clustering.

Table 8-3 shows statistical comparisons of classification accuracy between single modality (EEG, EDA, or heart rate) to all other multimodal combinations. Adding modalities increases performance, except when EDA is complemented with heart rate, or heart rate with EDA. Combinations of EDA and heart rate results in median clustering accuracy of 58% (IQR = 8%).

While adding other modalities to EEG results in higher clustering performance, perhaps more important is that clustering performance seems more robust

Table 8-3. Test statistics of comparison between classification results using different combinations of physiological measures.

EEG vs. EEG – EDA	EEG vs. EEG – HR	EEG vs. EDA – HR	EEG vs EEG – EDA - HR
t(37) = -2.77, p = .009	t(37) = -5.77, p < .001	t(37) = 4.40, p = <.001	t(37) = -4.78, p = <.001
EDA vs. EEG – EDA	**EDA vs. EEG – HR**	**EDA vs. EDA – HR**	**EDA vs EEG – EDA - HR**
t(38) = -10.92, p = <.001	t(38) = -19.08, p < .001	t(38) = -0.22, p = .825	t(38) = -14.91, p < .001
HR vs. EEG – EDA	**HR vs. EEG – HR**	**HR vs. EDA – HR**	**HR vs EEG – EDA - HR**
t(38) = -7.61, p < .001	t(38) = -12.88, p < .001	t(38) = 1.66, p = .104	t(38) = -10.61, p < .001

across algorithms. When combining EEG with EDA (Mdn = 81%) IQR is 4%; when combining EEG with heart rate (Mdn = 85%), IQR is 3% whereas IQR is 12% when using EEG only. When combining all three metrics, performance is as consistent as when combing EEG with heart rate only (Mdn = 85%, IQR = 3%).

8.3.3. Comparing clustering performance with other measures reflective of attentional engagement

Even though we specified the attentional instructions in the current study, we should note that we cannot be sure the attentional focus of participants is always as specified in the instructions. An incorrect classification may therefore not necessarily mean that the algorithms provided the wrong output, it may also be the case that the incorrectly classified participants did not follow their attentional instructions. To explore this possibility, we examined whether participants that were incorrectly classified by the majority of the methods for EEG performed worse on performance measures reflective of their attentional focus (number of correctly answered questions about the content of the narrative, number of correctly described emotional sounds, estimated number of averagely presented beeps), than participants that were correctly classified by the majority of the methods for EEG.

Seven participants were misclassified for more than 50% of the methods and designated as “often misclassified” (ID’s: 2, 3, 8, 10, 16, 18, 25). Table 8-4 provides the performance characteristics of often misclassified and often correctly classified participants and test statistics comparing the two. In the SA group, participants that were often misclassified described significantly less emotional sounds correctly than participants that were often correctly classified, which indeed suggests that misclassified SA instructed participants did not attend to the emotional sounds very well. For the other two performance measures no significant differences were found.

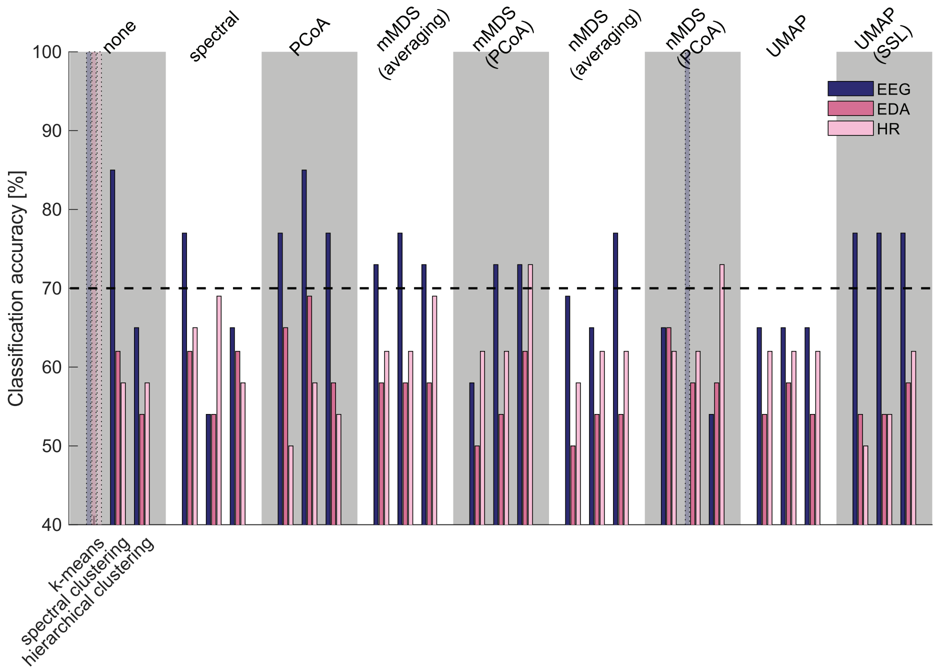

Figure 8-2. Clustering accuracies utilizing physiological synchrony in EEG, EDA, heart rate for different combinations of mappings (top-axis) and clustering methods (bottom-axis). Transparent top-to-bottom bars represent missing data. The dashed black line depicts significance level compared to chance level classification accuracies.

8.4. Discussion

We here showed that by applying unsupervised learning techniques to physiological synchrony, groups of participants sharing selective attentional focus can be identified from a set of participants with one of two different selective attentional instructions. This confirms hypothesis 1. Obtained results were found to depend on the physiological modality on which clustering was based.

8.4.1. Clustering performance using physiological synchrony in either EEG, EDA, or heart rate

We hypothesized that in line with previous research on physiological synchrony, from the three physiological measures EEG would perform best (hypothesis 2). Indeed, with the use of EEG, best performance was obtained. The maximum classification accuracy was 85% which is well above the threshold of 70% above which classification is significantly higher than chance level. However, performance varied strongly across clustering algorithms, with accuracies as low as 54% for some of the algorithms used.

Applying the clustering algorithms to physiological synchrony in EDA or heart

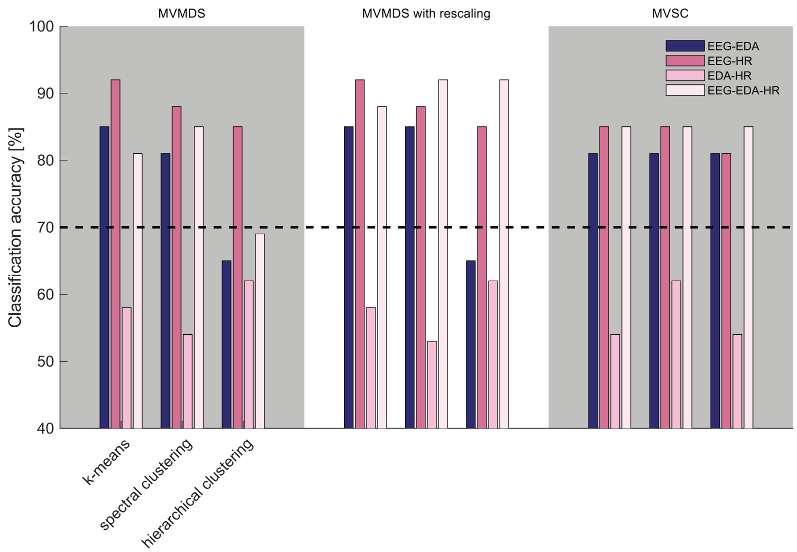

Figure 8-3. Clustering accuracies utilizing physiological synchrony in EEG and EDA, EEG and HR, EDA and heart rate, and EDA, EDA, and heart rate for different combinations of mappings (top-axis) and clustering algorithms (bottom-axis). The dashed black line depicts significance level compared to chance level classification accuracies. (bottom-axis). The dashed black line depicts significance level compared to chance level classification accuracies.

rate resulted in lower classification accuracies than in EEG, and generally led to performance near theoretical chance level. This is in line with other work, where synchronous changes in peripheral modalities have been shown to reflect attentional engagement with narrative stimuli less robustly than EEG (Ki et al., 2016; Stuldreher et al., 2020b; Pérez et al., 2021; Madsen and Parra, 2022).

8.4.2. Effect of multimodal combination of physiological measures on clustering performance

We hypothesized that combining modalities in a multimodal approach would enhance clustering performance compared to a unimodal approach, because different modalities capture different underlying mental processes (hypothesis 3). We partly accept this hypothesis. Indeed, when combining heart rate and EEG, EDA and EEG, or heart rate, EDA, and EEG, the clustering accuracy for combined modalities is higher than when using either of the modalities alone (Table 8-3). When combining heart rate and EEG, or heart rate, EDA, and EEG, the best obtained clustering accuracy across algorithms was also higher than when using either of the measures alone. When combining EDA with heart rate, classification accuracies were not higher compared to EDA or heart rate alone and thus still did not exceed chance level. Importantly, we found that when combining multiple physiological measures, results were not only gen-

Table 8-4. Test statistics comparing performance on questions reflective of attentional focus of the often incorrectly classified participants and the often correctly classified participants.

	Often correct participants AA	Often incorrect participants AA	Often correct participants SA	Often incorrect participants SA
Number of correctly answered narrative questions	Mdn = 5, IQR = 3	Mdn = 6.5, IQR = 2.5	x	x
	W = -57, p = .445			
Number of reproduced affective sounds	x	x	**Mdn = 7, IQR = 6**	**Mdn = 4, IQR = 2.3**
			W = 83.5, p = .028	
Difference between average number of estimated beeps and true number of beeps	x	x	Mdn = 2.5, IQR = 13	Mdn = 1, IQR = 9
			W = -74.5, p = .475	

erally higher but also more consistent across the range of mapping and clustering approaches. This was even the case when combination of modalities did not increase maximum classification accuracies, as for the combination of EEG and EDA. Thus, a multimodal approach resulted in classification performance that is less dependent on the specific algorithm choice. This observation advocates a multimodal approach in everyday settings where for unimodal data, the patterns of variation in algorithm performance may be different than the ones found here.

8.4.3. Factors underlying performance differences between modalities

We found that identifying two attentional groups in our study works best when physiological synchrony in EEG is used rather than EDA and heart rate. As mentioned in the introduction, we previously found that inter-subject correlations in EEG were especially sensitive to well-timed events inducing top-down modulation of attention, more so than to emotional sounds attracting attention bottom-up (Stuldreher et al., 2020a). This and related work showed major pre-frontal and parietal components contributing to inter-subject correlations in EEG when attending to narrative stimuli (Dmochowski et al., 2012; Cohen and Parra, 2016; Ki et al., 2016). Exactly these cortical areas are of major importance in top-down conscious attention processing (Vuilleumier and Driver, 2007). Inter-subject correlations in EDA and heart rate were modulated more by emotional sounds attracting attention bottom-up than by events that caused top-down modulation of attention (Stuldreher et al., 2020a). Other work also suggests that autonomic synchrony during presentation of narrative stimuli is mostly linked with emotional processing of these stimuli (Golland et al., 2014; Steiger et al., 2019). Electrodermal activity and heart rate are largely innervated by midbrain structures, such as the hypothalamus, amygdala and insula (Thayer et al., 2009; Boucsein, 2012) that are hard to capture using EEG. Such midbrain structures have been related strongly to bottom-up emotional modulation of attention (Behrmann et al., 2004; Vuilleumier and Driver, 2007). The fact that in our study, the difference between attentional groups was induced by instructions that affected attention in a cognitive, top-down manner, may have led to the finding that inter-subject correlations in EEG can here better distinguish between the groups with different selective top-down attentional conditions than inter-subject correlations in EDA and heart rate. Future work should investigate whether inter-subject correlations in EDA and heart rate are more suitable than EEG to distinguish between groups with different attentional conditions driven by emotional.

While physiological synchrony in EEG was found to be most informative of attentional group, adding other modalities generally led to higher and more robust performance. We see two potential explanations for the more robust clustering performance when combining modalities. It may be so that combining multiple modalities compensates for potential noisy observations in any of the modalities. Recent work of Madsen and Parra (2021) showed that physiological synchrony in EEG and heart rate in response to instructional videos are co-modulated. Thus, one noisy measurement may be compensated for by another measurement. Alternatively, more robust performance when combining modalities is expected when the physiological measures reflect different aspects of attentional engagement, so that by combining modalities in a multimodal fashion, one captures more aspects of the shared attentional engagement.

8.4.4. Effect of mapping and clustering approach on clustering performance

In our study, best classification results were obtained when using ordination techniques, here referred to as mapping methods, before applying clustering algorithms compared to directly using clustering algorithms on the distance matrices. This observation is supported by the silhouette coefficient, a measure of compactness and separability of the clusters, indicating less separable clusters when directly applying a clustering algorithm on the distance matrices than when using mapping methods. The low separability of clusters in the raw distance matrices may also explain why clustering results obtained with methods like k-medoids, that has random initializations, are different for different runs with different initializations and often did not converge. The only algorithm that provides good results when directly applied on the distance matrices is spectral clustering. This supports the notion that mapping before clustering is important, as the spectral clustering algorithm itself already computes a map before applying a clustering algorithm.

We cannot pinpoint the best mapping method for a general case. Here, clustering performance generally was best using PCoA or the multimodal equivalent MVMDS before applying clustering algorithms. Future work would have to show whether these findings are generalizable across use cases and physiological synchrony computation choices.

When using only a single modality, performance is strongly depends on the mapping method, and to a lesser degree on the clustering algorithm. With the exception of spectral mapping and no mapping conditions, clustering accuracy difference between the best and worst performing clustering algorithm with the same mapping is only around 15%. We cannot pinpoint a single clustering algorithm that performs best for each mapping.

8.4.5. Evaluation of clustering performance

In the current study we could employ the known attentional instructions to evaluate clustering performance. However, as noted before, we cannot be sure that the attentional focus always corresponded to the instructions. An incorrect classification may therefore not necessarily mean that the algorithms provided the wrong output, it may also be the case that the incorrectly classified participants did not follow their attentional instructions. We examined whether participants that were incorrectly classified by the majority of the methods for EEG scored worse on performance measures reflective of their instructed attentional focus (number of correctly answered questions about the content of the narrative for NA participants, number of correctly described emotional sounds, and estimated number of averagely presented beeps for SA participants). Indeed, we found that often incorrectly classified SA participants performed worse on the retention of the emotional sounds than the other participants, though there was no difference for the other two measures.

In real-world applications where unsupervised methods as proposed in the current work may be most applicable, the ground truth attentional condition is often not available. We therefore investigated whether the silhouette coefficient, a measure of the separability of the found clusters, may be used as a reliable metric of clustering performance. Unfortunately, a higher classification accuracy did not correspond to a higher silhouette coefficient and vice-versa. We thus reject hypothesis 4.

Since we have no reliable metric to evaluate clustering reliability when ground-truth labels are not known in real-world use cases, we suggest the use of a multimodal approach when applying unsupervised clustering algorithms on physiological synchrony data. Our results show that a multimodal approach is less prone to incorrect results that can occur for specific algorithm choices.

8.4.6. Future work

In this study, we sought to identify two clusters as participants were instructed to either attend to the audiobook or to the interspersed stimuli. However, it may be the case that some participants did not attend well to any of the presented information, a proposition that is substantiated by the observation that some participants showed low synchrony with both the audiobook and stimulus-attending groups and answered questions on the content of both presented streams of information well below the average (Stuldreher et al., 2020b). Others may have attended to both the audiobook and the interspersed stimuli. These two types of participants not necessarily fall into one of the two attentional groups considered in the current work, and may have negatively impacted the clustering performance, especially in algorithms such as k-means, where classification is strongly influenced by extreme values (Gupta et al., 2017). For more realistic clusters and better applicability in real-world environments, future work should evaluate clustering performance in relation to varying the numbers of clusters. Possible ways to approach this would be to take into account non-attending participants beforehand, by using outlier detection (He et al., 2003), to pre-specify three or four clusters in input of the clustering algorithm (e.g., to take into account non-attending participants and all-attending participants beforehand), or to use algorithms like mean shift (Comaniciu and Meer, 2002) or DBSCAN (Schubert et al., 2017) that automatically determine how many clusters appear in the data.

Another issue not addressed in current work is that many algorithms - such as k-means - tend to provide equal-sized clusters. This effect was not damaging in our study because we expected that the true clusters were about equal size, but in cases where this is not the case, results might be influenced by this algorithm bias. Future work should investigate how unequally sized clusters influence results and should explore algorithms that are less prone to such bias.

Future work should also explore other metrics for the assessment of clustering quality. In the current work the silhouette coefficient did not correspond well to ground truth performance. Potential metrics are distance to the cluster centroid or focus on clusters borders.

Finally, from a mathematical point of view, using other ways of computing the synchrony between the physiological signals could help improving clustering performance. In the current work, simple Pearson correlations were used to compute synchrony between two time-series, but computation of meaningful physiological synchrony, and therewith clustering performance, may be enhanced using other methods such as Dynamic time warping (Berndt and Clifford, 1994). Computing the correlation between two high dimensional signals can lead to the curse of dimensionality, a phenomenon that occurs in clustering with high-dimensional data, where data are more uniformly spread in high dimensions compared to lower dimensions when using a classical distance measure such as Euclidean distance (Bellman, 1967). Dynamic time warping was constructed with the aim of avoiding the curse of dimensionality, which could potentially lead to better clustering results.

8.5. Conclusion

We here combined physiological synchrony and unsupervised learning techniques with the aim to identify groups of individuals sharing the same selective attentional focus. Clustering performance well above chance level was reached when using EEG, but above chance level accuracies were not reached when using EDA or heart rate alone. Obtained results differed depending on the used mapping and clustering algorithm, but applying mapping before clustering generally led to better results. Combining information from multiple modalities resulted in a higher classification performance in cases where EEG was combined with heart rate and/or EDA, and resulted in more robust performance across different types of mapping and clustering algorithms, making clustering results less dependent on the specific algorithm choice. These results may enable researchers to study attentional engagement in everyday settings. We suggest researchers to use a multimodal approach due to its robustness to specific algorithm choice, enabling more consistent and generally better clustering results.

9.

Physiological synchrony in heart rate and electrodermal activity in the classroom

Stuldreher, I.V., Thammasan, N., Schreuders, E., Tjew-A-Sin, M., de Geus, E.J.C., van Erp, J.B.F., Brouwer, A.M. (2021). Physiological synchrony in the classroom. In *Proceedings of the 2021 Neuroergonomics Conference*, München. doi: 10.1145/3382507.3418837

9.1. Background

Physiological synchrony (PS) refers to the degree to which physiological measures of multiple individuals uniformly change (Palumbo et al., 2017). In a controlled laboratory study, we found that physiological synchrony reflects selective attention: The electroencephalogram (EEG), heart rate and electrodermal activity (EDA) of participants showed similar changes when they paid attention to similar aspects of an auditory stimulus (Stuldreher et al., 2020b). PS may also be informative of attentional engagement in educational settings. Dikker et al. (2017) found that students' EEG was more synchronized with each other when they were more, rather than less engaged during a semi-regular biology class. However, lessons were adapted specifically to the study and EEG sensors are still considered to be rather obtrusive. In the current work, we use data of two earlier conducted studies to assess PS in EDA and heart rate among students during regular classes. As a first step in probing whether PS in peripheral measures might be used for monitoring attention in secondary school education, we compared PS between students in the same versus different classrooms, and aimed to distinguish students attending the same class from students attending different classes.

9.2. Methods

Study 1 (Thammasan et al., 2020) was conducted at two schools in the Netherlands. The study was approved by the institutional ethics research board of Tilburg University. Data of 86 adolescents (14.9 ± 0.5 years) coming from 17 different classes were collected during a regular school day. EDA (palm-based) and heart rate were monitored using the Movisens EdaMove 4 (Movisens GmbH, Karlsruhe, Germany) and Wahoo Tickr (Wahoo Fitness, Atlanta, GA, USA), respectively.

Study 2 was conducted at one secondary school in the Netherlands. The study was approved by the Scientific and Ethical Review Board of the Faculty of Behavior & Movement Sciences, VU University Amsterdam. Data of 29 adolescents (13.8 ± 0.4 years) coming from 21 different classes were collected for 24 hours, including a regular school day. EDA (palm-based) and electrocardiogram (ECG) were recorded using the VU University Ambulatory Monitoring System (VU-AMS). ECG peaks were detected following Pan and Tompkins (1985) and transformed into heart rate time-series.

Data from both studies were epoched to the on- and offset of lessons. We further refer to each epoch as a 'student', where one such epoch represents one unique combination of one lesson (or classroom) and one student. Statistics of the number of students in both studies are in Table 9-1. For each student, we

Table 9-1. Number of students (included data sets) for the two studies. Note that due to (partially) failed recordings, the number of included data sets differ between heart rate and EDA.

	Study 1		**Study 2**	
	Heart rate	**EDA**	**Heart rate**	**EDA**
Total number of students	75	59	11	13
Same lesson (mean SD per lesson)	4.71 ± 1.94	1.78 ± 1.65	1.02 ± 0.55	1.23 ± 0.44
Different lesson (mean SD per lesson)	333.28 ± 1.94	99.09 ± 74.58	115.17 ± 35.35	135.77 ± 0.44

computed PS using inter-subject correlations with all other students from that study, following Stuldreher et al. (2020b). For each student, PS toward students attending the same lesson was computed by averaging over synchrony scores with all other students in the same classroom. PS toward students not attending the same lesson was computed by averaging over synchrony scores with all other students not in the same classroom. Figure 9-1AB depicts this processing pipeline. Using paired sample t-tests, we tested for each study and each physiological measure whether physiological synchrony was higher for students when paired with others attending the same classroom than when paired with others not attending the same classroom. In addition, we examined for each student whether classification into 'same' or 'different' classroom based on PS worked out correctly.

9.3. Results

Figure 9-1CD summarizes the results.

For study 1, PS for students was significantly higher when paired with students attending the same class than when paired with students attending other classes, both for heart rate ($t(74) = 5.36$, $p < .001$) and EDA ($t(58) = 1.89$, $p = .032$). Classifying an individual to the group to which they showed a higher PS resulted in classification accuracies of 77.3% (heart rate) and 61.0% (EDA).

For study 2, PS for students was not significantly higher when paired with students attending the same class than when paired with students attending other classes, both for HR ($t(10) = 0.48$, $p = .322$) and EDA ($t(12) = 0.53$, $p = .303$). Classification accuracies were 36.4% (heart rate) and 38.5% (EDA).

The null-finding in study 2 may be explained by the lower amount of data compared to study 1 (both in terms of number of shared lessons and non-shared lessons; see Table 9-1). Indeed, when re-examining data from study 1 using the same characteristics as for study 2, by randomly sampling 1000 times from the

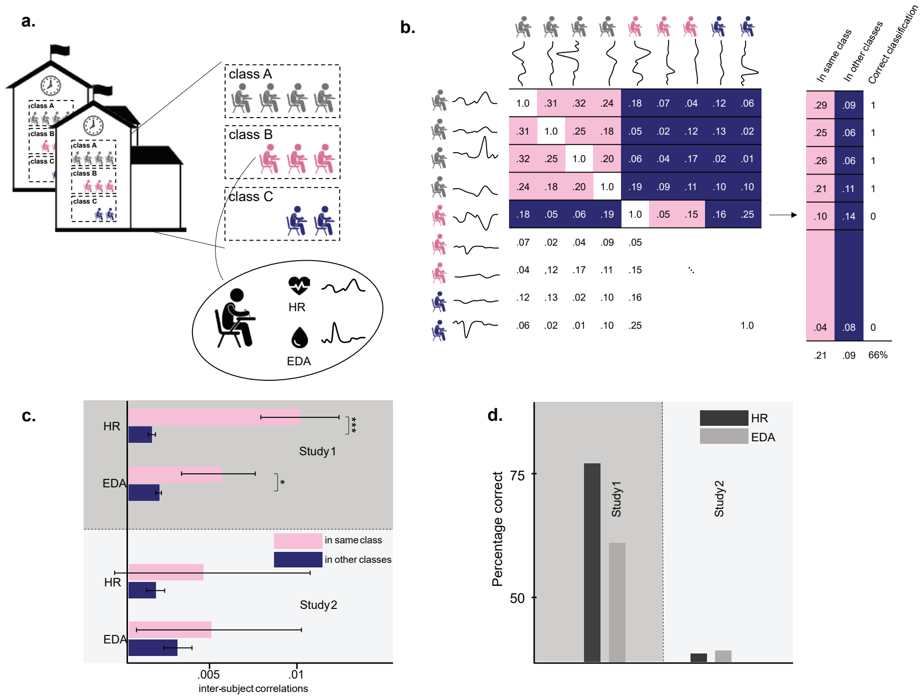

Figure 9-1. Overview of the processing pipeline and results. a. depicts that data were collected in two different studies in multiple groups of students per school. In each classroom, a varying number of volunteering students was equipped with heart rate (HR) and EDA sensors. b. shows how physiological synchrony (PS) was determined. For both HR and EDA, PS was computed using inter-subject correlations for all possible pairs of students from that study, resulting in a N x N matrix, were N refers to the number of collected datasets (= number of actual students X number of lessons) in the study. For each student, an average synchrony toward students attending the same lesson was computed by row-wise averaging over synchrony values with all other students in the same classroom (pink cells) and an average synchrony toward students not attending the same lesson (blue cells). c. shows the inter-subject correlations to students attending the same lesson and attending other lessons for HR and EDA for studies 1 and 2. Error bars depict standard error of the mean. d. shows the percentage of correctly identified groups, when classifying the student to the group (same class/other class) that he or she showed highest synchrony with.

complete dataset, PS is no longer significantly larger for students attending the same class compared to students attending different classes for both heart rate ($t(10) = 1.64$, $p = .066$) and EDA ($t(7) = 0.91$, $p = .196$). Classification accuracies in this case are 54.7 ± 12.8% for heart rate and 60.7% ± 17.0% for EDA.

9.4. Discussion

We here showed that PS in heart rate and EDA can be picked up using wearables in the classroom – students following the same lesson in the same classroom show stronger PS for both measures compared to students in other classrooms. This reached significance in data from one of the two studies that we analyzed, and the same trend is visible in the other study. We suspect that the main cause of the null result in one of the studies is the low number of participants, particularly those in the same classroom (on average only 1.02 classmate rather than 4.72 for heart rate and 1.23 rather than 1.78 for EDA) - see spread in 'same class' PS values for study 2 in Figure 9-1C. Heart rate seems to be a more robust measure than EDA. This was not specifically expected on the basis of our earlier work, using the same wearables as in Study 1 in a controlled, laboratory setting (Van Beers et al., 2020) where both measures performed about equally well, or EDA somewhat better. EDA might be relatively sensitive to noise in real life environments. Note that this study does not yet look into PS yet as a measure of attention and PS may have (partly) been caused by similar patterns in students' physical activity. The next step in examining and validating PS in heart rate and EDA as potential markers of selective attention in the real classroom environment, would be to register an independent, alternative measure of selective attention.

10.

General discussion

In this thesis we aimed to uncover whether different types of attention modulation are captured by physiological synchrony and to what extent physiological synchrony may be used as a tool to monitor attentional engagement in real-life settings. In part I, we studied how different causes of varying attention affect physiological synchrony in brains and bodies. In part II we tried to bridge the gap from lab to life, by determining minimally necessary conditions that are important for successful monitoring of inter-subject correlations in real-life conditions.

10.1. Part I: attentional modulations and physiological synchrony in brains and bodies

In part I, we addressed the research question: "How do different manipulations of attention affect physiological synchrony in brains and bodies?"

10.1.1. Bottom-up and top-down attentional processing reflected by synchronous brains and bodies

In Chapters 2 and 3 we studied the respective influences of sensory bottom-up and higher-level top-down attentional processes on inter-subject correlations as a measure of physiological synchrony in EEG, EDA and heart rate. Variations in attention originating from bottom-up and top-down attentional processes were both found to affect inter-subject correlations, though their respective influence varied between EEG, EDA and heart rate.

Previous findings of synchrony in brains and bodies across individuals presented with the same narrative stimulus could largely be explained by bottom-up attentional mechanisms drawing attention. In this thesis we explicitly confirmed that moments in a stimulus that attract attention through bottom-up mechanisms related to the emotional relevance result in higher levels of physiological synchrony.

However, the current thesis also demonstrated repeatedly that top-down attentional processes modulate the occurrence of such physiological synchrony. For instance, a) individuals with the same selective attentional focus have higher inter-subject correlations in EEG, EDA and heart rate than individuals with different selective attentional focus; b) depending on the instructed focus of attention certain auditory events result in increased inter-subject correlations or not; and c) inter-subject correlations in EEG, EDA and heart rate predict performance on post-stimulus tests.

Through our results we thus provide explicit evidence that physiological synchrony upon the presentation of narrative stimuli is not only a result of the bottom-up features of the stimulus guiding individuals' attention. Instead, we confirm the already in previous literature presented hypothesis that top down processing guiding attentional engagement also modulates the occurrence of physiological synchrony. Our findings are in line with recent findings by others. Rosenkranz, Holtze and colleagues showed that the inter-subject correlations in EEG of individuals were higher with other individuals focusing on the same than with individuals focusing on a different speaker in a cocktail party paradigm experiment (Rosenkranz et al., 2021; Holtze et al., 2022). Pérez, Madsen and colleagues reported that inter-subject correlations in EEG, EDA, heart rate, eye movements and pupil size were higher when individuals actively attended a stimulus than when focusing attention inward on a mental arithmetic task (Pérez et al., 2021; Madsen and Parra, 2022, 2023). The same authors also reported similar relations between inter-subject correlations and post-stimulus test performance for EEG, heart rate, eye movement and pupil size.

In our and the abovementioned studies, it was shown that shared attentional processing is sufficient for the occurrence of physiological synchrony. Madsen and Parra (2022) come to a similar conclusion in their study, namely that the cognitive processing of a shared stimulus is sufficient for the occurrence of inter-subject correlations. We do want to highlight the subtle difference in these conclusions. Although cognitive processing of a shared stimulus causes physiological synchrony to occur, attention is essential for the cognitive processing of a shared stimulus. Without attention directed to a stimulus, be it consciously through top-down mechanisms or involuntarily through sensory mechanisms, there will not be cognitive processing of the stimulus. Also important to note here is that through using the terminology attention instead of cognitive processing we also emphasize that it is not only voluntary cognitive processing that drives physiological synchrony, but also attention that is attracted by emotional or arousing events. In our view, a better description for the prerequisites of inter-subject correlations therefore is that the occurrence of inter-subject correlations is driven by attentional processing of a shared stimulus.

10.1.2. Discrepancies between synchronous brains and bodies in their ability to reflect attention

Notable in our findings is that not only brain-to-brain synchrony reflects attentional engagement, but also body-to-body synchrony reflects attentional engagement. In our studies inter-subject correlations in EDA and heart rate were higher than expected based on chance during the presentation of a shared stimulus, inter-subject correlations in EDA and heart rate distinguished between individuals with different selective attentional focus, inter-subject correlations in EDA and heart rate predicted the occurrence of attentionally relevant events and inter-subject correlations in EDA and heart rate predicted post-stimulus test performance. Since our findings, others have presented similar results. Pérez et al. (2021) presented that upon the presentation of a shared stimulus, inter-subject correlations in heart rate were higher than one would expect based on chance level, were higher when actively attending the presented narrative than when focusing attention inward on a mental arithmetic task and predicted better recall of the narrative. Madsen and Parra found similar results for heart rate, EDA, pupil size and gaze direction (Madsen and Parra, 2022, 2023).

The observation that also body-to-body synchrony reflects attentional engagement may be surprising to some as the underlying control is different from the signals captured by the EEG and quite diverse among physiological measures. Heart rate is controlled by midbrain structures modulated by input from the amygdala, cingulate and insula (Thayer et al., 2009) and pathways of EDA originate in the posterior hypothalamus (Collins, 1999). However, these deeper brain structures are part of larger networks that are also connected to cortical areas, of which activity is captured by EEG. Heart rate, for instance, is connected through inhibitory pathways from the pre-frontal cortex to the amygdala (Thayer and Lane, 2009; Thayer et al., 2009) and EDA is connected to higher subcortical and cortical brain areas implicated in attention, such as the ventromedial prefrontal cortices and the hippocampus (Critchley, 2002). Madsen and Parra (2022) therefore suggest that a robust brain-body connection is a prerequisite for inter-subject correlations in a measure to reflect shared attentional engagement. The authors tested this hypothesis and confirmed that measures that were coupled to the EEG – heart rate, gaze position, pupil size, saccade rate – showed significant inter-subject correlations, but measures that were not coupled to the EEG – respiration rate, head velocity – did not show these inter-subject correlations. In their hypothesis the authors thus consider EEG as a gold standard when measuring attention, as only metrics related to the EEG may be informative of the cognitive processing of a shared stimulus. However, in our work we show that synchronous autonomic measures can also contain complementary information to synchronous brains. Inter-subject correlations in EDA and heart rate were more sensitive to the occurrence of emotionally salient events than inter-subject correlations in EEG. The specific sensitivity of the autonomic nervous system activity to emotion is often noted (Smith et al., 2004) and can also be understood from an anatomical perspective. The hypothesis that physiological measures need to be closely coupled to the EEG to reflect shared attentional engagement implies that information on the attentional processing is transmitted over pathways from the cortex to such physiological measures. However, deeper brain structures such as the amygdala and insula that are closely involved in control of autonomic nervous system activity, play an important role in emotion experience (Lindquist et al., 2012). These deeper brain structures are part of larger networks with robust connections to the autonomic nervous system, for instance through the hypothalamus as the area for central control of the autonomic nervous system (Marcus et al., 2015). These networks are also connected to cortical areas, of which activity is captured by EEG, see for instance (Zotev et al., 2016), but these pathways are more complex than the connections with the autonomic nervous system.

Our finding that synchrony in brain and body reflect complementary information are at odds with the hypothesis of Madsen and Parra. For physiological signals to be informative of attentional processing, they do not necessarily need to be closely related to EEG. Clearly, cortical areas, such as prefrontal and parietal regions, are essential in establishing attentional engagement. EEG, mainly capturing cortical activity, can thus certainly inform us on the attentional engagement of monitored individuals. However, attention is not only established by cortical areas, but rather is the result of complex networks in the entire nervous system, among which cortical areas and deeper brain areas are involved. Physiological measures that are innervated through different pathways connecting to deeper brain areas inform us on other aspects of attentional processing that are not captured by the EEG. It is the discrepancy between EDA or heart rate and EEG that contains additional information on attentional processing, not the similarity between these measures. Our alternative hypothesis is that inter-subject correlations in brain and body measures can reflect shared attentional engagement, where measures that are innervated through different pathways reflect different aspects of attentional engagement.

Besides the clear differences in respective contribution of attentional processes captured in either brains or bodies we also found discrepancies in results using EDA or heart rate. Compared to inter-subject correlations in heart rate, inter-subject correlations in EDA were more associated with shorter moments of attentiveness due to discrete, arousing sensory events and less so to longer periods of attentional engagement driven by higher-order cognitive processes. Inter-subject correlations in EDA were for instance more responsive to the occurrence of emotional sounds than inter-subject correlations in heart rate, but were not as strongly related to stimulus retention as inter-subject correlations in heart rate. We do not know what underlies this observation and we are not aware of other researchers noting this difference. It may be so that the different autonomic innervations underly these differences. Whereas the heart is innervated by both the sympathetic and parasympathetic branches of the autonomic nervous system, the sweat gland activity captured by EDA is purely innervated by the sympathetic branch (Cacioppo et al., 2000; Boucsein, 2012). Parasympathetic activation has been associated with performance on several cognitive tasks involving attention and working memory (Hansen et al., 2003; Thayer et al., 2009; Smith et al., 2017) and decreased parasympathetic activity has been related to focused attention (Cacioppo et al., 1978; Suess et al., 1994). The higher-order cognitive functioning captured by inter-subject correlations in heart rate may thus be associated with parasympathetic activation that is not reflected by EDA.

10.1.3. Multimodal brain-body measures of physiological synchrony

If our hypothesis that was introduced in the previous section is true, the attentional processes that are captured by inter-subject correlations differ between brain and body and within different bodily measures. The inter-subject correlations in these modalities should then contain complementary information, such that a multimodal metric combining information from brain and body should better reflect attentional engagement. Indeed, when combining brain-to-brain and body-to-body inter-subject correlations in a multimodal approach we could better distinguish between individuals with different selective attentional focus and better detect the occurrence of attentionally relevant events than when using only one of the two. We therefore argue for a multimodal approach of physiological synchrony, in which signals measured from brains and bodies are combined into a single metric of physiological synchrony. Such multimodal metrics seem particularly valuable when contextual information, for instance about the type of stimulus that is presented, is not available.

10.1.4. Capturing momentary attentional state or trait

In most of our experiments, attentional instructions were varied across individuals to generate differences in attentional focus. In real-life settings, attention also varies across and within individuals in less obvious ways. We investigated how physiological synchrony reflects interpersonal and intrapersonal variations in attentional processing and how estimated levels of attention correspond to a momentary attentional state or to a more long-term pattern of attentional processing corresponding to personal trait. Through these studies we did not aim to show that physiological synchrony reflects either momentary attentional state or relates to personal trait, but that the relation between physiological synchrony and attentional engagement is a widely supported phenomenon.

Regarding momentary attentional state, in Chapter 5 we found that inter-subject correlations could predict changes in general attention in time. Inter-subject correlations in EDA obtained during the viewing of short movie clips over the course of a night of total sleep deprivation predicted performance on consecutive vigilant attention tasks. This suggests that the momentary attentional processing capabilities captured by inter-subject correlations in EDA during the presented movie clips are associated with longer lasting variations in the general attentional capability of individuals that also affects other types of tasks.

In Chapter 4 we found that interpersonal variations in attention as captured by physiological synchrony can also originate from the personal trait of monitored individuals. We found positive significant correlations between the level of food neophobia – the hesitance to try and buy new foods – and the inter-subject correlations in EEG during the presentation of a movie about a foreign food.

Together, these results show that the relation between physiological synchrony and attentional engagement is a widely supported phenomenon. It is not only attentional processes driven by the presentation of a shared stimulus, or top-down attentional processes directly related to processing the presented stimulus that affect physiological synchrony. Instead, one's priors and momentary state can modulate attentional processing and therewith they also modulate the level of physiological synchrony. Thereby, physiological synchrony may also be a marker of such priors or momentary attentional state.

Our results add on prior research as, to the best of our knowledge, for the first time we show that physiological synchrony is predictive of one's momentary attentional state. Furthermore, unlike previous research in populations with autism, depression or psychosis (Hasson et al., 2009; Salmi et al., 2013; Guo et al., 2015; Mäntylä et al., 2018), where inter-subject correlations are diminished across the measured populations, in our results inter-subject correlations vary with the intensity of a certain trait, in our case food neophobia. This result is arguably more subtle and better reflects how traits may affect physiological synchrony in the general population. The results also suggest that one's traits do not necessarily affect the occurrence of physiological synchrony in general. Instead, traits can alter the relative importance of presented information, such that populations with specific traits show altered inter-subject correlations for specific, but not all stimuli. Together, results are demonstrative for the complex factors affecting attentional processing that all modulate the level of physiological synchrony.

10.1.5. Factors underlying the occurrence of physiological synchrony

Taken together, in part I of this thesis we showed that the relationship between physiological synchrony and attentional engagement is a broadly carried phenomenon. Attentional engagement is a composite of sensory processes drawing attention in a bottom-up manner and top-down processes modulating attention, which can vary momentarily in time, but are also affected by priors. We showed that physiological synchrony can capture these influences on attentional processes.

Our results, that have been confirmed by others, nuance the large number of studies (see e.g., Palumbo et al., 2017) stating that physiological synchrony is a marker or even driver of social interaction. We show that social interaction is not a prerequisite for the occurrence of physiological synchrony, but that shared attentional processing of the same information is sufficient physiological synchrony to occur. It may simply be shared attentional processing that, at least partly, underlies the positive findings of physiological synchrony in contexts of social interaction.

Our findings do not rule out numerous additional factors that contribute to the occurrence of physiological synchrony in general. We discuss the two most important factors. First, synchronized movement can cause increased levels of physiological synchrony by matched metabolic demands. Synchronous responses in brains or bodies across individuals have occurred in various synchronous and coordinated movements (Funane et al., 2011; Holper et al., 2012; Yun et al., 2012; Müller et al., 2013; Ikeda et al., 2017; Gordon et al., 2020). Second, the mere co-presence of other people sharing a common experience was found to produce higher inter-subject correlations in autonomic measures (Golland et al., 2015; Ardizzi et al., 2020). Golland and colleagues for instance found higher inter-subject correlations between pairs that watched a movie together than between pairs that watched the same movie, but separately (Golland et al., 2015). The authors suggest that this is a result of recursive interpersonal influences during which individual differences of the emotional experience propagated across the co-present viewers, leading to a shared emotional experience. Perhaps additional synchronization among co-present individuals involves the mirror-neuron system. In humans, the perception of a specific emotion also triggers the neural systems responsible for generation of that emotion in the observer. This has for instance been found for fear (De Gelder et al., 2004), disgust (Wicker et al., 2003), anxiety (Prehn-Kristensen et al., 2009) and reward (Mobbs et al., 2009). It is unclear which sensory channels are responsible for the transmission of this emotional information. It may be that emotional cues such as posture or facial expressions are detected in peripheral vision and thereby influence one's own emotional state, as demonstrated in (Bayle et al., 2011; Calvo et al., 2014a, 2014b). Another explanation is that emotional states are mediated by chemosignals, therewith circumventing the conscious allocation of attention (de Groot et al., 2012, 2015).

10.2. Part II: physiological synchrony from lab to life

In part II we addressed the research question: "To what extent may physiological synchrony be used as a tool to monitor attention in real-life settings?" We determined minimally necessary conditions for successful monitoring of inter-subject correlations and explored whether and to what extent information is lost when relying on wearable sensors rather than high-end equipment and recording in real life conditions.

10.2.1. Monitoring physiological synchrony with wearables

In Chapter 6 we compared the inter-subject correlation in EDA and heart rate measured using high-end lab equipment with those monitored using relatively cheap and accessible wearable devices. Although raw signal quality was worse for wearables than for high-end lab equipment, inter-subject correlations monitored using wearables could distinguish between individuals with different selective attentional focus with similar, if not higher, accuracies than when monitored using high-end lab-equipment. Positive results were confirmed in following studies, presented in Chapters 5, 7 and 9. The fact that inter-subject correlations monitored using wearable equipment are as informative of attentional engagement as when monitored using laboratory-grade equipment is not only good news for real-life applications, it also informs us on the signal features that are important for monitoring attentional engagement. In our studies, algorithms integrated in wearable equipment filter high frequency information in hear rate. It thus appears that especially the lower frequencies of the heart rate signals carry information on the attentional engagement of monitored individuals. This was recently confirmed by Madsen and Parra, who found that the heart rate of attentive and distracted individuals mainly varied around 0.1 Hz (Madsen and Parra, 2023).

Important for future studies is to consider the minimum amount of data required for successful monitoring of inter-subject correlations using wearables. In Chapter 7 we explored the minimum amount of data necessary to successfully monitor inter-subject correlations in EDA and heart rate using wearable equipment. To the best of our knowledge, this was the first time that the effect of the amount of data on the resulting inter-subject correlations was systematically studied. As a rule of thumb we found that roughly 10 hours of data recording are needed for robust inter-subject correlations in EDA and heart rate. Here robust means that approximately 80% of individuals show inter-subject correlations that are higher than expected based on chance. We found that it did not matter whether this total amount of data was reached by adding more individuals or by increasing stimulus duration. Although we cannot say for sure that this rule of thumb applies more general than for our specific sample of stimuli and participants, an initial comparison with other recent studies (Pérez et al., 2021; Madsen and Parra, 2022) suggest that these results may apply beyond our set of participants and stimuli. If this general rule of thumb indeed applies, there are implications for use of physiological synchrony in real-life settings. In settings where only a limited number of individuals are monitored for a short amount of time, the signal-to-noise ratio may not be large enough to provide reliable information on the attentional processing of monitored individuals. Physiological synchrony as a tool to monitor attentional engagement thus is most suited in applications where larger groups of individuals are together for elongated periods of time. Think for instance of students in a classroom that are equipped with wearables to monitor attentional engagement during a lesson or a large test-panel testing how engaging a newly-developed advertisement is.

10.2.2. Physiological synchrony and unsupervised clustering

In Chapter 8 we explored the combination of inter-subject correlations with novel unsupervised clustering techniques. We investigated whether individuals with two different selective attentional conditions towards the same stimulus could be clustered together while being blind to the attentional condition for all participants. We showed that this was indeed possible. This is an important step for the use of physiological synchrony as attention monitoring tool in real-life settings, as in such settings there is often no contextual information or information about the attentional state of any of the monitored individuals available a-priori. By combining inter-subject correlations with unsupervised clustering algorithms, clusters of shared attending individuals can for instance be detected. For instance, clusters of individuals that have trouble attending to a specific lesson may then be identified and helped. In the current study we did specify beforehand that there were two clusters of individuals sharing attentional focus, whereas in real-life the number of attentional groups may be unclear. We would therefore suggest to explore the use of algorithms such as mean shift or DBSCAN that automatically determine the appropriate number of clusters (Comaniciu and Meer, 2002; Schubert et al., 2017).

10.2.3. Physiological synchrony in the classroom

As a proof of concept of the use of physiological synchrony in real-life settings we explored whether inter-subject correlations in EDA and heart rate could be monitored in a real-life classroom setting using wearable devices. We found that inter-subject correlations in EDA and heart rate could distinguish between individuals in the same and different classrooms. This does not necessarily mean that this is based on differences in top-down attention, as effects of synchronous body movements or attention drawn by salient sensory events could have played an important role. Still, it is a first indication that inter-subject correlations can be captured without controlling for the information that was presented. There have been some other studies exploring the use of physiological synchrony in classroom or lecture settings (Gashi et al., 2018, 2019; Liu et al., 2021;

Zhang et al., 2021). Liu and colleagues for instance reported that inter-subject correlations in EDA could predict collaboration quality among students in the classroom (Liu et al., 2021). Zhang and colleagues reported that inter-subject correlations in EDA among students in the classroom predicted self-reported attention, self-reported mastery of knowledge, but also the test exam score (Zhang et al., 2021). Our work adds to these other studies in that we show that inter-subject correlations can be captured without controlling for the information that was presented, over the course of multiple lessons during an entire school day.

The development of more accurate wearable EEG headsets and artifact-removal algorithms also allow for monitoring of brain waves in real-life settings. Already in 2017, Dikker and colleagues reported a study in which brain-to-brain synchrony between 12 classroom students was predictive of classroom engagement in a classroom-like setting (Dikker et al., 2017). In the years following, work was extended by showing that brain-to-brain synchrony can also predict students' performance in immediate and delayed retention tests (Davidesco et al., 2019), student-teacher brain-to-brain synchrony predicted social closeness between student and teacher (Bevilacqua et al., 2019) and brain-to-brain synchrony among students was higher in specific lecture segments associated with questions that were answered correctly (Davidesco et al., 2023).

10.2.4. Recent developments and limitations

Neurophysiological measurement equipment is in continuous development. Equipment is shrinking in size while signal quality increases. Especially of value for our aim are the developments of in-ear and around-the-ear EEG sensor systems (Looney et al., 2012; Debener et al., 2015; Goverdovsky et al., 2016), of which the cEEGrid is one example. Using this latter around-the-ear sensor system, Holtze and colleagues recently found that they could identify to which of two audiostreams individuals were listening based on the inter-subject correlations among listeners (Holtze et al., 2022). Around-the-ear EEG sensors may thus be a suitable wearable sensor for monitoring attentional engagement based on physiological synchrony.

Such developments are needed to overcome the current limitations for the use of physiological synchrony as a tool to monitor attentional engagement in real-life settings. We previously mentioned that relatively large amounts of data are still required to obtain robust inter-subject correlations, at least in EDA and heart rate. In addition, though we found clear and consistent relations between physiological synchrony and attentional engagement, the variance in attention explained by physiological synchrony is limited. In our studies the variance explained ranged anywhere from five to forty percent. In our studies the strongest associations were found between physiological synchrony and performance on a post-stimulus test, in which participants answers questions about the content of the presented stimulus. For use in real-life settings this may imply that physiological synchrony can best be used as tool to assess the overall attentional engagement towards presented narrative information, such as a lecture.

10.3. Towards physiological synchrony as a tool to monitor attentional engagement in real-life settings

In the current section we would like to highlight three future directions to move towards physiological synchrony as a tool to monitor attentional engagement in real-life setting, corresponding to applied-scientific, methodological and ethical questions that are still open.

The first research direction is tackling the applied-scientific questions that remain open. In the current thesis physiological synchrony was shown to reflect attentional engagement and as a proof of concept physiological synchrony was monitored in a real life setting, namely among students in the classroom. However, we did not couple these two. Future work could investigate to what extent physiological synchrony in bodily measures such as EDA and heart rate predicts the level of attentional engagement among students in the classroom, parallel to work on brain-to-brain synchrony in the classroom (Dikker et al., 2017; Davidesco et al., 2023). In the current thesis we further showed that attentional processing of a shared stimulus is sufficient for the occurrence of physiological synchrony. We therefore hypothesized that it is also shared attentional processing that, at least partly, underlies the positive findings of physiological synchrony in contexts of social interaction. Future studies should test this hypothesis in settings of social interaction. Such studies could for instance focus on collaborative teamwork settings, in which the attentional focus varies dynamically across teammates and varies in time. Future studies could for instance investigate whether moments of shared attention are predicted by moments of high physiological synchrony or whether physiological synchrony may predict when an individual is in shared attentional engagement with the rest of team. If physiological synchrony indeed reflects shared attention also in settings of social interaction, it would allow the development of a synchrony-based tool for assisting members in a team to collaborate more effectively when needed.

Another question involves real-time monitoring and use of physiological synchrony. Though physiological synchrony can be determined in near real-time, even by use of existing open-source algorithms (e.g., Ayrolles et al., 2021), we are not aware of studies that aim to boost performance by adapting a system in real-time using information about attention of monitored individuals (Brouwer et al., 2023). We believe such real-time systems are feasible and valuable, for instance in classroom settings to inform a student, adapt a virtual classroom setting or by providing information to a human teacher. Future work should explore the effectiveness of such systems.

A second research direction deals with methodological approaches for the establishment of physiological synchrony. In the current thesis we used simple inter-subject correlations to quantify physiological synchrony, as we had no specific hypothesis regarding nonlinear relationships and expected that attending to the same input would result in similar neurophysiological activation at that point in time. In social settings where the neurophysiological response of one individual is expected to precede that of others correlation analyses may be performed at different time lags. By doing so Davidesco and colleagues for instance found that brain-to-brain synchrony between teacher and student was highest when the teacher's brain signals preceded those of the students by 300 ms (Davidesco et al., 2023). More general, there are numerous different analytical approaches that have been explored for the establishment of physiological synchrony, summarized in a number of review studies. Czeszumski and colleagues provide a comprehensive overview of methodological approaches for the quantification of brain-to-brain synchrony (Czeszumski et al., 2020). Palumbo and colleagues provide an overview of methodologic approaches for the quantification of interpersonal autonomic synchrony (Palumbo et al., 2017). Helm and colleagues outline several approaches for measuring physiological synchrony in dyads and relate these to different types of synchrony, being trend, concurrent or lagged synchrony (Helm et al., 2018). Different types of analysis may be most suited for different types of settings. With the move towards more dynamic settings than explored in this thesis, analyses that can deal with these dynamics may be more appropriate than the currently used correlation analysis. Here one can for instance think of nonlinear analysis techniques such as cross-recurrence analysis, as for instance used in (Konvalinka et al., 2011). Alternatively, analysis of synchrony may be geared at longer-lasting brain states than the consecutive event-related potentials that are captured by inter-subject correlations. For instance through assessment of coherence, as previously demonstrated by (Reinero et al., 2021).

We would like to emphasize the importance of developing novel multimodal metrics of physiological synchrony. There is a strong difference in response latency and frequency power spectrum between EEG and autonomic measures complicating the combination into a single multimodal metric. In a literature study conducted in 2019 we did not find any studies combining information of brains and bodies into a single multimodal metric of physiological synchrony, perhaps due to these challenges (Stuldreher et al., 2019). Although averaging over the z-scored inter-subject correlation values from EEG, EDA and heart rate at each point in time allowed for better detection of attentionally engaging events than by using one of these measures (Chapter 3), this simple approach certainly does not capture the full potential of a multimodal brain-body metric of synchrony. Future researchers should investigate applications where inter-subject correlations in brains and bodies can be combined with decision algorithms, to allow multimodal combination of inter-subject correlations from brain and body at the decision level. Alternatively, the use of multivariate regression analysis, which can be used to study the association between multiple independent variables and the dependent variable of interest (Alexopoulos, 2010), may be explored.

A third important consideration of future work is the ethics of the use of physiological synchrony as tool to monitor attentional engagement. Alongside the research conducted in this thesis we explored the ethics of using wearable technology to monitor attention in the classroom, among others by speaking with high-school students on this (Snoek et al., 2022). Ethics are mostly perceived as exploring risks of a technology. However, in a broader sense, it is an exploration of the implications of the technology, both negative and positive. The use of wearables in the classroom may enhance students' engagement (Borthwick et al., 2015; Bower and Sturman, 2015; Sandall, 2016; Engen et al., 2017; Attallah and Ilagure, 2018), it may provide educators with valuable information on what teaching style is experienced as most engaging (Bower and Sturman, 2015; Demir and Demir, 2016; Sandall, 2016; Dikker et al., 2017; Attallah and Ilagure, 2018; Janssen et al., 2021) and it may help to differentiate between educational needs of students (Borthwick et al., 2015; Dikker et al., 2017; Babiker et al., 2019). On the other hand, the use of wearables in the classroom may violate well-known ethical principles, such as autonomy, privacy and consent (Ienca et al., 2018; Mecacci and Haselager, 2019). In addition, the use of wearables in the classroom has the potential to change our relationship with the world, known as Foucauldian ethics. For instance, large-scale use of wearables may change what we value in teachers: technical knowledge over pedagogical qualities. Studies exploring the use of physiological synchrony as tool to monitor attentional engaging should consider the ethics corresponding to their specific application. By involving developers of technology, ethicists and end-users one may improve the development and design of technologies such that the risks and concerns that are identified can be taken away to the satisfaction of involved stakeholders.

10.4. Conclusion

In this thesis we aimed to uncover whether different types of attention modulation are captured by physiological synchrony and to what extent physiological synchrony may be used as tool to monitor attention in real-life settings.

Part I was aimed at answering the research question: “How do different manipulations of attention affect physiological synchrony in brains and bodies?” We found that physiological synchrony reflects multiple types of attention modulations, both sensory and top-down mechanisms of attention. We add to previous work by showing that not only synchronous brain, but also synchronous body metrics reflect this attentional processing and even showed that these reflect complementary aspects of attention. Physiological synchrony was shown to reflect the momentary attentional state of monitored individuals, as well as interpersonal variation in attentional processing originating from interpersonal variations in personal traits. In our work we bridged the scattered literature, by showing that physiological synchrony in brain and body can both reflect shared attention. This may explain numerous findings in a unified coherent way. Specifically, findings of physiological synchrony in social interaction may also be explained by shared attention instead of any esoteric phenomenon.

Part II was aimed at answering the research question: “To what extent may physiological synchrony be used as tool to monitor attention in real-life settings?” Our results were promising for physiological synchrony to be used in real-life settings. Information on attention could be accessed by monitoring brains, but also by monitoring body measures that are more easily accessible in real-life settings. Such body synchrony can also be reliably monitored using wearable equipment, even in real-life settings such as the classroom, as long as sufficient amounts of data are used. Physiological synchrony can be combined with novel algorithmic developments, as demonstrated by the successful combination with unsupervised learning techniques. That said, limitations of physiological synchrony as tool to monitor attentional engagement should be considered. Sufficient data is required for robust monitoring and even physiological synchrony can only explain part of the variance in attention.

To advance physiological synchrony as tool to monitor attention in real-life settings, future work should focus on the applied, methodological and ethical questions that remain unanswered. Future work on scientific questions should investigate how physiological synchrony may reflect attention in more dynamic real-life settings and how real-time monitoring of physiological synchrony may enhance attentional guidance in groups of individuals. Future work on methodological questions should focus on novel metrics to quantify physiological synchrony that may be better suited for dynamic real-life settings. Future work on ethical questions should focus on the ethics of the use of wearables and attention monitoring in real-life settings, such as the classroom.

References

Abiri, R., Borhani, S., Jiang, Y., & Zhao, X. (2019). Decoding Attentional State to Faces and Scenes Using EEG Brainwaves. *Complexity*, *2019*. https://doi.org/10.1155/2019/6862031

Adedokun, O. A., & Burgess, W. D. (2011). Analysis of Paired Dichotomous Data: A Gentle Introduction to the McNemar Test in SPSS. *Journal of MultiDisciplinary Evaluation*, *8*(17). https://doi.org/10.56645/jmde.v8i17.336

Alexopoulos, E. C. (2010). Introduction to multivariate regression analysis. *Hippokratia*, *14*.

Algumaei, M., Hettiarachchi, I., Veerabhadrappa, R., & Bhatti, A. (2023). Physiological Synchrony Predict Task Performance and Negative Emotional State during a Three-Member. *Sensors*, *23*(4). https://doi.org/10.3390/s23042268

Alhola, P., & Polo-Kantola, P. (2007). Sleep deprivation: Impact on cognitive performance. In *Neuropsychiatric Disease and Treatment* (Vol. 3, Issue 5).

Aliakbaryhosseinabadi, S., Kamavuako, E. N., Jiang, N., Farina, D., & Mrachacz-Kersting, N. (2017). Classification of EEG signals to identify variations in attention during motor task execution. *Journal of Neuroscience Methods*, *284*. https://doi.org/10.1016/j.jneumeth.2017.04.008

Anderson, J. R. (1985). A series of books in psychology. Cognitive psychology and its implications. *Decision Support Systems*, *38*(4).

Andrade, B. F., Brodeur, D. A., Waschbusch, D. A., Stewart, S. H., & McGee, R. (2009). Selective and sustained attention as predictors of social problems in children with typical and disordered attention abilities. *Journal of Attention Disorders*, *12*(4). https://doi.org/10.1177/1087054708320440

Ardizzi, M., Calbi, M., Tavaglione, S., Umiltà, M. A., & Gallese, V. (2020). Audience spontaneous entrainment during the collective enjoyment of live performances: physiological and behavioral measurements. *Scientific Reports*, *10*(1), 1–12. https://doi.org/10.1038/s41598-020-60832-7

Arico, P., Borghini, G., Di Flumeri, G., Sciaraffa, N., & Babiloni, F. (2018). Passive BCI beyond the lab: Current trends and future directions. *Physiological Measurement*, *39*(8). https://doi.org/10.1088/1361-6579/aad57e

Arons, B. (1992). A Review of The Cocktail Party Effect. *Journal of the American Voice I/O Society*, *12*.

Arvola, A., Lähteenmäki, L., & Tuorila, H. (1999). Predicting the intent to purchase unfamiliar and familiar cheeses: The effects of attitudes, expected liking and food neophobia. *Appetite*, *32*(1). https://doi.org/10.1006/appe.1998.0181

Attallah, B., & Ilagure, Z. (2018). Wearable Technology: Facilitating or Complexing Education? *International Journal of Information and Education Technology*, *8*(6), 433–436. https://doi.org/10.18178/ijiet.2018.8.6.1077

Attfield, S., Kazai, G., & Lalmas, M. (2011). Towards a science of user engagement (Position Paper). *WSDM Workshop on User Modelling for Web Applications*.

Awais, M., Badruddin, N., & Drieberg, M. (2017). A hybrid approach to detect driver drowsiness utilizing physiological signals to improve system performance and Wearability. *Sensors (Switzerland)*, *17*(9). https://doi.org/10.3390/s17091991

Ayrolles, A., Brun, F., Chen, P., Djalovski, A., Beauxis, Y., Delorme, R., Bourgeron, T., Dikker, S., & Dumas, G. (2021). HyPyP: A Hyperscanning Python Pipeline for inter-brain connectivity analysis. *Social Cognitive and Affective Neuroscience*, *16*(1–2). https://doi.org/10.1093/scan/nsaa141

Babiker, A., Faye, I., Mumtaz, W., Malik, A. S., & Sato, H. (2019). EEG in classroom: EMD features to detect situational interest of students during learning. *Multimedia Tools and Applications*, *78*(12). https://doi.org/10.1007/s11042-018-7016-z

Babiloni, F., & Astolfi, L. (2014). Social neuroscience and hyperscanning techniques: Past, present and future. In *Neuroscience and Biobehavioral Reviews* (Vol. 44). https://doi.org/10.1016/j.neubiorev.2012.07.006

Bäckström, A., Pirttilä-Backman, A. M., & Tuorila, H. (2004). Willingness to try new foods as predicted by social representations and attitude and trait scales. *Appetite*, *43*(1). https://doi.org/10.1016/j.appet.2004.03.004

Bailenson, J. N., Pontikakis, E. D., Mauss, I. B., Gross, J. J., Jabon, M. E., Hutcherson, C. A. C., Nass, C., & John, O. (2008). Real-time classification of evoked emotions using facial feature tracking and physiological responses. *International Journal of Human Computer Studies*, 66(5). https://doi.

org/10.1016/j.ijhcs.2007.10.011

Baron-Cohen, S. (1997). *Mindblindness: An essay on autism and theory of mind*. MIT Press.

Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences*, *12*(3). https://doi.org/10.1207/S15327809JLS1203_1

Basner, M., & Dinges, D. F. (2011). Maximizing sensitivity of the Psychomotor Vigilance Test (PVT) to sleep loss. *Sleep*, *34*(5). https://doi.org/10.1093/sleep/34.5.581

Bauckhage, C. (2015). *NumPy / SciPy Recipes for Data Science: k-Medoids Clustering*.

Bayle, D. J., Schoendorff, B., Hénaff, M. A., & Krolak-Salmon, P. (2011). Emotional facial expression detection in the peripheral visual field. *PLoS ONE*, *6*(6). https://doi.org/10.1371/journal.pone.0021584

Behrmann, M., Geng, J. J., & Shomstein, S. (2004). Parietal cortex and attention. In *Current Opinion in Neurobiology* (Vol. 14, Issue 2). https://doi.org/10.1016/j.conb.2004.03.012

Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, *7*(6). https://doi.org/10.1162/neco.1995.7.6.1129

Bellman, R. (1967). Dynamic programming. *Mathematics in Science and Engineering*, *40*(P1), 101–137. https://doi.org/10.1016/S0076-5392(08)61063-2

Ben-Shakhar, G. (1985). Standardization Within Individuals: A Simple Method to Neutralize Individual Differences in Skin Conductance. *Psychophysiology*, *22*(3). https://doi.org/10.1111/j.1469-8986.1985.tb01603.x

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, *190*(1), 80–91. https://doi.org/10.1016/j.jneumeth.2010.04.028

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*(1). https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Berger, H. (1929). “Über das Elektroenkephalogramm des Menschen.” *Archiv Für Psychiatrie Und Nervenkrankheiten*, *87*.

Berggren, N., & Eimer, M. (2021). The role of trait anxiety in attention and memory-related biases to threat: An event-related potential study. *Psychophysiology*, *58*(3). https://doi.org/10.1111/psyp.13742

Berndt, D., & Clifford, J. (1994). Using dynamic time warping to find patterns in time series. *Workshop on Knowledge Knowledge Discovery in Databases*, *398*.

Bevilacqua, D., Davidesco, I., Wan, L., Chaloner, K., Rowland, J., Ding, M., Poeppel, D., & Dikker, S. (2019). Brain-to-Brain Synchrony and Learning Outcomes Vary by Student–Teacher Dynamics: Evidence from a Real-world Classroom Electroencephalography Study. *Journal of Cognitive Neuroscience*, *31*(3), 401–411. https://doi.org/https://doi.org/10.1162/jocn_a_01274

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K. M., & Robbins, K. A. (2015). The PREP pipeline: Standardized preprocessing for large-scale EEG analysis. *Frontiers in Neuroinformatics*, *9*(JUNE). https://doi.org/10.3389/fninf.2015.00016

Borg, I., & Groenen, P. J. F. (2005). Modern multidimensional scaling: Theory and applications, 2nd ed. In *Modern multidimensional scaling: Theory and applications, 2nd ed.*

Borg, I., Groenen, P. J. F., & Mair, P. (2018). *Applied Multidimensional Scaling and Unfolding*. http://link.springer.com/10.1007/978-3-319-73471-2

Borovac, A., Stuldreher, I. V, & ... (2020). Validation of wearables for electrodermal activity (EdaMove) and heart rate (Wahoo Tickr). In *Proceedings of the 12th International Conference on Measurement and Behaviour and 6th International Seminar on Behavioral Methods, Krakow, Poland, 2021;* Volume 1, 18–24. doi: 10.6084/m9.figshare.13013717

Borthwick, A. C., Anderson, C. L., Finsness, E. S., & Foulger, T. S. (2015). Special Article Personal Wearable Technologies in Education: Value or Villain? *Journal of Digital Learning in Teacher Education*, *31*(3), 85–92. https://doi.org/10.1080/21532974.2015.1021982

Bottenheft, C., Hogenelst, K., Stuldreher, I. V., Kleemann, R., Groen, E. L., van Erp, J. B. F., Brouwer, A. M. (2023). Understanding the combined effects of sleep deprivation and acute social stress on cognitive performance using a comprehensive approach. *Brain, Behavior, & Immunity – Health, 34*, 100706. doi: 10.1016/j.bbih.2023.100706

Boucsein, W. (2012). Electrodermal Activity. In *Electrodermal Activity*. https://doi.org/10.1007/978-1-4614-1126-0

Bower, M., & Sturman, D. (2015). What are the educational affordances of wearable technologies? *Computers and Education*, *88*. https://doi.org/10.1016/j.compedu.2015.07.013

Bracken, B. K., Alexander, V., Zak, P. J., Romero, V., & Barraza, J. A. (2014). Physiological synchronization is associated with narrative emotionality and subsequent behavioral response. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *8534 LNAI*. https://doi.org/10.1007/978-3-319-07527-3_1

Bradley, M. M., Hamby, S., Löw, A., & Lang, P. J. (2007). Brain potentials in perception: Picture complexity and emotional arousal. *Psychophysiology*, *44*(3). https://doi.org/10.1111/j.1469-8986.2007.00520.x

Bradley, M. M., & Lang, P. J. (2000). Affective reactions to acoustic stimuli. *Psychophysiology*, *37*(2). https://doi.org/10.1017/S0048577200990012

Bradley, M. M., & Lang, P. J. (2007a). The International Affective Digitized Sounds (2nd Edition; IADS-2): Affective ratings of sounds and instruction manual. Technical report B-3. *University of Florida, Gainesville, Fl*.

Bradley, M. M., & Lang, P. J. (2007b). The International Affective Picture System (IAPS) in the study of emotion and attention. In *Handbook of emotion elicitation and assessement*.

Braithwaite, J., Watson, D., Robert, J., & Mickey, R. (2013). A Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological Experiments. ….

Brouwer, A. M. (2021). Challenges and Opportunities in Consumer Neuroergonomics. *Frontiers in Neuroergonomics*, *2*(March), 1–6. https://doi.org/10.3389/fnrgo.2021.606646

Brouwer, A. M., Stuldreher, I. V., & van Erp, J. B. F. (2023). Interpersonal physiological synchrony based BCI: a perspective. *Proceedings of the 10th International Brain-Computer Interface Meeting 2023*, 144470. https://doi.org/https://doi.org/10.3217/978-3-85125-962-9-87

Brouwer, A. M., & Hogervorst, M. A. (2014). A new paradigm to induce mental stress: The Sing-a-Song Stress Test (SSST). *Frontiers in Neuroscience*, *8*(8 JUL), 1–8. https://doi.org/10.3389/fnins.2014.00224

Brouwer, A. M., Hogervorst, M., Reuderink, B., van der Werf, Y., & van Erp, J. (2015a). Physiological signals distinguish between reading emotional and non-emotional sections in a novel. *Brain-Computer Interfaces*, *2*(2–3). https://doi.org/10.1080/2326263X.2015.1100037

Brouwer, A. M., Stuldreher, I. V., & Thammasan, N. (2019). Shared attention reflected in eeg, electrodermal activity and heart rate. *CEUR Workshop Proceedings*, *2474*, 27–31.

Brouwer, A. M., van Beers, J. J., Sabu, P., Stuldreher, I. V., Zech, H. G., & Kaneko, D. (2021). Measuring implicit approach–avoidance tendencies towards food using a mobile phone outside the lab. *Foods*, *10*(7). https://doi.org/10.3390/foods10071440

Brouwer, A. M., van Dam, E., van Erp, J. B. F., Spangler, D. P., & Brooks, J. R. (2018). Improving real-life estimates of emotion based on heart rate: A perspective on taking metabolic heart rate into account. *Frontiers in Human Neuroscience*, *12*. https://doi.org/10.3389/fnhum.2018.00284

Brouwer, A. M., Zander, T. O., van Erp, J. B. F., Korteling, J. E., & Bronkhorst, A. W. (2015b). Using neurophysiological signals that reflect cognitive or affective state: Six recommendations to avoid common pitfalls. *Frontiers in Neuroscience*, *9*(APR). https://doi.org/10.3389/fnins.2015.00136

Bruner, J. (1985). Child's Talk: Learning to Use Language. *Child Language Teaching and Therapy*, *1*(1). https://doi.org/10.1177/026565908500100113

Cacioppo, J. T., Sandman, C. A., & Walker, B. B. (1978). The Effects of Operant Heart Rate Conditioning on Cognitive Elaboration and Attitude Change. *Psychophysiology*, *15*(4). https://doi.org/10.1111/j.1469-8986.1978.tb01389.x

Cacioppo, J. T., Tassinary, L. G., & Berntson, G. G. (2000). Handbook of Psychophysiology. *Book*.

Calvo, M. G., Beltrán, D., & Fernández-Martín, A. (2014). Processing of facial expressions in peripheral vision: Neurophysiological evidence. *Biological Psychology*, *100*(1). https://doi.org/10.1016/j.biopsycho.2014.05.007

Calvo, M. G., Fernández-Martín, A., & Nummenmaa, L. (2014). Facial expression recognition in peripheral versus central vision: Role of the eyes and the mouth. *Psychological Research*, *78*(2). https://

doi.org/10.1007/s00426-013-0492-x

Carini, R. M., Kuh, G. D., & Klein, S. P. (2006). Student engagement and student learning: Testing the linkages. *Research in Higher Education*, *47*(1). https://doi.org/10.1007/s11162-005-8150-9

Carreiras, C., Lourenço, A., Aidos, H., da Silva, H. P., & Fred, A. L. N. (2016). Unsupervised analysis of morphological ecg features for attention detection. In *Studies in Computational Intelligence* (Vol. 613). https://doi.org/10.1007/978-3-319-23392-5_24

Casson, A. J. (2019). Wearable EEG and beyond. In *Biomedical Engineering Letters* (Vol. 9, Issue 1). https://doi.org/10.1007/s13534-018-00093-6

Cherry, E. C. (1953). Some Experiments on the Recognition of Speech, with One and with Two Ears. *Journal of the Acoustical Society of America*, *25*(5). https://doi.org/10.1121/1.1907229

Cheung, M. N. (1981). Detection of and Recovery from Errors in Cardiac Interbeat Intervals. *Psychophysiology*, *18*(3). https://doi.org/10.1111/j.1469-8986.1981.tb03045.x

Choe, J. Y., & Cho, M. S. (2011). Food neophobia and willingness to try non-traditional foods for Koreans. *Food Quality and Preference*, *22*(7). https://doi.org/10.1016/j.foodqual.2011.05.002

Chung, L., Chung, S. J., Kim, J. Y., Kim, K. O., O'Mahony, M., Vickers, Z., Cha, S. M., Ishii, R., Baures, K., & Kim, H. R. (2012). Comparing the liking for Korean style salad dressings and beverages between US and Korean consumers: Effects of sensory and non-sensory factors. *Food Quality and Preference*, *26*(1). https://doi.org/10.1016/j.foodqual.2012.03.011

Cohen, S. S., Henin, S., & Parra, L. C. (2017). Engaging narratives evoke similar neural activity and lead to similar time perception. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-04402-4

Cohen, S. S., Madsen, J., Touchan, G., Robles, D., Lima, S. F. A., Henin, S., & Parra, L. C. (2018). Neural engagement with online educational videos predicts learning performance for individual students. *Neurobiology of Learning and Memory*, *155*(January), 60–64. https://doi.org/10.1016/j.nlm.2018.06.011

Cohen, S. S., & Parra, L. C. (2016). Memorable audiovisual narratives synchronize sensory and supramodal neural responses. *ENeuro*, *3*(6), 1–11. https://doi.org/10.1523/ENEURO.0203-16.2016

Collins, K. J. (1999). Temperature regulation and the autonomic nervous system. In M. C.J. & R. Bannister (Eds.), *Autonomic failure: a textbook of clinical disorders of the autonomic nervous system* (4th ed., pp. 169–195). Oxford University Press.

Comaniciu, D., & Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*(5). https://doi.org/10.1109/34.1000236

Compton, R. J. (2003). The interface between emotion and attention: a review of evidence from psychology and neuroscience. In *Behavioral and cognitive neuroscience reviews* (Vol. 2, Issue 2). https://doi.org/10.1177/1534582303002002003

Constable, R. T. (2012). Challenges in fMRI and its limitations. In *Functional Neuroradiology: Principles and Clinical Applications*. https://doi.org/10.1007/978-1-4419-0345-7_19

Cooper, N. R., Croft, R. J., Dominey, S. J. J., Burgess, A. P., & Gruzelier, J. H. (2003). Paradox lost? Exploring the role of alpha oscillations during externally vs. internally directed attention and the implications for idling and inhibition hypotheses. *International Journal of Psychophysiology*, *47*(1). https://doi.org/10.1016/S0167-8760(02)00107-1

Coull, J. T. (1998). Neural correlates of attention and arousal: Insights from electrophysiology, functional neuroimaging and psychopharmacology. *Progress in Neurobiology*, *55*(4). https://doi.org/10.1016/S0301-0082(98)00011-2

Critchley, H. D. (2002). Electrodermal responses: What happens in the brain. *Neuroscientist*, *8*(2), 132–142. https://doi.org/10.1177/107385840200800209

Critchley, H. D., Eccles, J., & Garfinkel, S. N. (2013). Interaction between cognition, emotion, and the autonomic nervous system. In *Handbook of Clinical Neurology* (Vol. 117). https://doi.org/10.1016/B978-0-444-53491-0.00006-7

Czeszumski, A., Eustergerling, S., Lang, A., Menrath, D., Gerstenberger, M., Schuberth, S., Schreiber, F., Rendon, Z. Z., & König, P. (2020). Hyperscanning: A Valid Method to Study Neural Inter-brain Underpinnings of Social Interaction. *Frontiers in Human Neuroscience*, *14*(February), 1–17. https://doi.org/10.3389/fnhum.2020.00039

Davidesco, I., Laurent, E., Valk, H., West, T., Dikker, S., Milne, C., & Poeppel, D. (2019). Brain-to-brain synchrony predicts long-term memory retention more accurately than individual brain measures. *BioRxiv*, 644047. http://biorxiv.org/content/early/2019/05/21/644047.abstract

Davidesco, I., Laurent, E., Valk, H., West, T., Milne, C., Poeppel, D., & Dikker, S. (2023). The Temporal Dynamics of Brain-to-Brain Synchrony Between Students and Teachers Predict Learning Outcomes. *Psychological Science*. https://doi.org/10.1177/09567976231163872

De Dieuleveult, A. L., Brouwer, A. M., Siemonsma, P. C., Van Erp, J. B. F., & Brenner, E. (2018). Aging and sensitivity to illusory target motion with or without secondary tasks. *Multisensory Research*, *31*(3–4). https://doi.org/10.1163/22134808-00002596

De Gelder, B., Snyder, J., Greve, D., Gerard, G., & Hadjikhani, N. (2004). Fear fosters flight: A mechanism for fear contagion when perceiving emotion expressed by a whole body. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(47). https://doi.org/10.1073/pnas.0407042101

de Groot, J. H. B., Smeets, M. A. M., Kaldewaij, A., Duijndam, M. J. A., & Semin, G. R. (2012). Chemosignals Communicate Human Emotions. *Psychological Science*, *23*(11). https://doi.org/10.1177/0956797612445317

de Groot, J. H. B., Smeets, M. A. M., Rowson, M. J., Bulsing, P. J., Blonk, C. G., Wilkinson, J. E., & Semin, G. R. (2015). A Sniff of Happiness. *Psychological Science*, *26*(6). https://doi.org/10.1177/0956797614566318

Debener, S., Emkes, R., De Vos, M., & Bleichner, M. (2015). Unobtrusive ambulatory EEG using a smartphone and flexible printed electrodes around the ear. *Scientific Reports*, *5*. https://doi.org/10.1038/srep16743

Dehais, F., Karwowski, W., & Ayaz, H. (2020). Brain at Work and in Everyday Life as the Next Frontier: Grand Field Challenges for Neuroergonomics. *Frontiers in Neuroergonomics*, *1*. https://doi.org/10.3389/fnrgo.2020.583733

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1). https://doi.org/10.1016/j.jneumeth.2003.10.009

Demir, E. B. K., & Demir, K. (2016). Enhancing learning with wearable technologies in and out of educational settings. In *Digital Tools for Seamless Learning*. https://doi.org/10.4018/978-1-5225-1692-7.ch006

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. In *Annual Review of Neuroscience* (Vol. 18). https://doi.org/10.1146/annurev.ne.18.030195.001205

Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., Rowland, J., Michalareas, G., Van Bavel, J. J., Ding, M., & Poeppel, D. (2017). Brain-to-Brain Synchrony Tracks Real-World Dynamic Group Interactions in the Classroom. *Current Biology*, *27*(9), 1375–1380. https://doi.org/10.1016/j.cub.2017.04.002

Dmochowski, J. P., Bezdek, M. A., Abelson, B. P., Johnson, J. S., Schumacher, E. H., & Parra, L. C. (2014). Audience preferences are predicted by temporal reliability of neural processing. *Nature Communications*, *5*, 1–9. https://doi.org/10.1038/ncomms5567

Dmochowski, J. P., Sajda, P., Dias, J., & Parra, L. C. (2012). Correlated components of ongoing EEG point to emotionally laden attention - a possible marker of engagement? *Frontiers in Human Neuroscience*, *6*(MAY 2012), 1–9. https://doi.org/10.3389/fnhum.2012.00112

Donald, M. W., & Little, R. (1981). The analysis of stimulus probability inside and outside the focus of attention, as reflected by the auditory N1 and P3 components. *Canadian Journal of Psychology*, *35*(2). https://doi.org/10.1037/h0081153

Doran, S. M., Van Dongen, H. P. A., & Dinges, D. F. (2001). Sustained attention performance during sleep deprivation: Evidence of state instability. *Archives Italiennes de Biologie*, *139*(3).

Dorrian, J., & Dinges, D. F. (2005). Sleep deprivation and its effects on cognitive performance. In *Encyclopedia of Sleep Medicine* (pp. 139–144). John Wiley and Sons, NJ.

Dumas, G., Nadel, J., Soussignan, R., Martinerie, J., & Garnero, L. (2010). Inter-brain synchronization during social interaction. *PLoS ONE*, *5*(8). https://doi.org/10.1371/journal.pone.0012166

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, *66*(3). https://doi.org/10.1037/h0047707

Eertmans, A., Victoir, A., Vansant, G., & Van den Bergh, O. (2005). Food-related personality traits, food choice motives and food intake: Mediator and moderator relationships. *Food Quality and Preference*, *16*(8). https://doi.org/10.1016/j.foodqual.2005.04.007

Elkins, A. N., Muth, E. R., Hoover, A. W., Walker, A. D., Carpenter, T. L., & Switzer, F. S. (2009). Physiological compliance and team performance. *Applied Ergonomics*, *40*(6). https://doi.org/10.1016/j.apergo.2009.02.002

Elliott, R., & Dolan, R. J. (1998). Activation of different anterior cingulate foci in association with hypothesis testing and response selection. *NeuroImage*, *8*(1). https://doi.org/10.1006/nimg.1998.0344

Engen, B. K., Giæver, T. H., & Mifsud, L. (2017). Teaching and Learning with Wearable Technologies. *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*, *2014*, 1057–1067.

Fairclough, S. H., & Lotte, F. (2020). Grand Challenges in Neurotechnology and System Neuroergonomics. *Frontiers in Neuroergonomics*, *1*. https://doi.org/10.3389/fnrgo.2020.602504

Falciglia, G. A., Couch, S. C., Gribble, L. S., Pabst, S. M., & Frank, R. (2000). Food neophobia in childhood affects dietary variety. *Journal of the American Dietetic Association*, *100*(12). https://doi.org/10.1016/S0002-8223(00)00412-0

Feldman, R., Magori-Cohen, R., Galili, G., Singer, M., & Louzoun, Y. (2011). Mother and infant coordinate heart rhythms through episodes of interaction synchrony. *Infant Behavior and Development*, *34*(4). https://doi.org/10.1016/j.infbeh.2011.06.008

Fellous, J. M., Sapiro, G., Rossi, A., Mayberg, H., & Ferrante, M. (2019). Explainable Artificial Intelligence for Neuroscience: Behavioral Neurostimulation. *Frontiers in Neuroscience*, *13*. https://doi.org/10.3389/fnins.2019.01346

Ferrer, E., & Helm, J. L. (2013). Dynamical systems modeling of physiological coregulation in dyadic interactions. *International Journal of Psychophysiology*, *88*(3). https://doi.org/10.1016/j.ijpsycho.2012.10.013

Field, A. (2013). Discovering statistics using IBM SPSS statistics. In *Statistics* (Vol. 58).

Filippo D'Antuono, L., & Bignami, C. (2012). Perception of typical Ukrainian foods among an Italian population. *Food Quality and Preference*, *25*(1). https://doi.org/10.1016/j.foodqual.2011.12.003

Finn, E. S., Glerean, E., Khojandi, A. Y., Nielson, D., Molfese, P. J., Handwerker, D. A., & Bandettini, P. A. (2020). Idiosynchrony: From shared responses to individual differences during naturalistic neuroimaging. *NeuroImage*, *215*(April), 116828. https://doi.org/10.1016/j.neuroimage.2020.116828

FIT SDK RELEASE NOTES. (2020). https://www.thisisant.com/developer/fit-sdk-release-notes

Fleureau, J., Guillotel, P., & Quan, H. T. (2012). Physiological-based affect event detector for entertainment video applications. *IEEE Transactions on Affective Computing*, *3*(3). https://doi.org/10.1109/T-AFFC.2012.2

Foody, G. M. (2009). Classification accuracy comparison: Hypothesis tests and the use of confidence intervals in evaluations of difference, equivalence and non-inferiority. *Remote Sensing of Environment*, *113*(8). https://doi.org/10.1016/j.rse.2009.03.014

Fox, E., Russo, R., Bowles, R., & Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *Journal of Experimental Psychology: General*, *130*(4). https://doi.org/10.1037/0096-3445.130.4.681

Fuller, D., Colwell, E., Low, J., Orychock, K., Ann Tobin, M., Simango, B., Buote, R., van Heerden, D., Luan, H., Cullen, K., Slade, L., & Taylor, N. G. A. (2020). Reliability and Validity of Commercially Available Wearable Devices for Measuring Steps, Energy Expenditure, and Heart Rate: Systematic Review. In *JMIR mHealth and uHealth* (Vol. 8, Issue 9). https://doi.org/10.2196/18694

Funane, T., Kiguchi, M., Atsumori, H., Sato, H., Kubota, K., & Koizumi, H. (2011). Synchronous activity of two people's prefrontal cortices during a cooperative task measured by simultaneous near-infrared spectroscopy. *Journal of Biomedical Optics*, *16*(7). https://doi.org/10.1117/1.3602853

Furman, O., Dorfman, N., Hasson, U., Davachi, L., & Dudai, Y. (2007). They saw a movie: long-term memory for an extended audiovisual narrative. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *14*(6).

https://doi.org/10.1101/lm.550407

Galloway, A. T., Lee, Y., & Birch, L. L. (2003). Predictors and consequences of food neophobia and pickiness in young girls. *Journal of the American Dietetic Association*, *103*(6). https://doi.org/10.1053/jada.2003.50134

Garbarino, M., Lai, M., Tognetti, S., Picard, R., & Bender, D. (2014). *Empatica E3 - A wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition*. https://doi.org/10.4108/icst.mobihealth.2014.257418

Gashi, S., Di Lascio, E., & Santini, S. (2019). Using Unobtrusive Wearable Sensors to Measure the Physiological Synchrony Between Presenters and Audience Members. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *3*(1). https://doi.org/10.1145/3314400

Gashi, S., Lascio, E. Di, & Santini, S. (2018). Using students' physiological synchrony to quantify the classroom emotional climate. *UbiComp/ISWC 2018 - Adjunct Proceedings of the 2018 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2018 ACM International Symposium on Wearable Computers*. https://doi.org/10.1145/3267305.3267693

Goense, J., Bohraus, Y., & Logothetis, N. K. (2016). fMRI at high spatial resolution implications for BOLD-models. *Frontiers in Computational Neuroscience*, *10*(Jun). https://doi.org/10.3389/fncom.2016.00066

Golland, Y., Arzouan, Y., & Levit-Binnun, N. (2015). The mere Co-presence: Synchronization of autonomic signals and emotional responses across Co-present individuals not engaged in direct interaction. *PLoS ONE*, *10*(5). https://doi.org/10.1371/journal.pone.0125804

Golland, Y., Keissar, K., & Levit-Binnun, N. (2014a). Studying the dynamics of autonomic activity during emotional experience. *Psychophysiology*, *51*(11), 1101–1111. https://doi.org/10.1111/psyp.12261

Gordon, I., Gilboa, A., Cohen, S., & Kleinfeld, T. (2020). The relationship between physiological synchrony and motion energy synchrony during a joint group drumming task. *Physiology and Behavior*, *224*(July), 113074. https://doi.org/10.1016/j.physbeh.2020.113074

Goverdovsky, V., Looney, D., Kidmose, P., & Mandic, D. P. (2016). In-Ear EEG From Viscoelastic Generic Earpieces: Robust and Unobtrusive 24/7 Monitoring. *IEEE Sensors Journal*, *16*(1). https://doi.org/10.1109/JSEN.2015.2471183

Gower, J. C., & Legendre, P. (1986). Metric and Euclidean properties of dissimilarity coefficients. *Journal of Classification*, *3*(1). https://doi.org/10.1007/BF01896809

Graham, F. K., & Clifton, R. K. (1966). Heart-rate change as a component of the orienting response. *Psychological Bulletin*, *65*(5). https://doi.org/10.1037/h0023258

Grinsted, A., Moore, J. C., & Jevrejeva, S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics*, *11*(5/6). https://doi.org/10.5194/npg-11-561-2004

Grira, N., Crucianu, M., & Boujemaa, N. (2004). Unsupervised and Semi-supervised Clustering: A Brief Survey. *A Review of Machine Learning Techniques for Processing Multimedia Content, Report of the MUSCLE European Network of Excellence (6th Framework Programme)*.

Groenen, P. J. F., & Borg, I. (2014). Past, Present, and Future of Multidimensional Scaling. In J. Blasius & M. Greenacre (Eds.), *Visualization and Verbalization of Data* (pp. 95–117). Chapman and Hall/CRC. https://doi.org/10.1201/b16741-13

Groenen, P. J. F., & van de Velden, M. (2005). Multidimensional Scaling. In B. S. Everitt & D. Howel (Eds.), *Encyclopedia of Statistics in Behavioral Science* (pp. 779–784). John Wiley and Sons. https://doi.org/10.1002/0470013192.bsa415

Guo, C. C., Nguyen, V. T., Hyett, M. P., Parker, G. B., & Breakspear, M. J. (2015). Out-of-sync: Disrupted neural activity in emotional circuitry during film viewing in melancholic depression. *Scientific Reports*, *5*. https://doi.org/10.1038/srep11605

Gupta, S., Kumar, R., Lu, K., Moseley, B., & Vassilvitskii, S. (2017). Local search methods for k-means with outliers. *Proceedings of the VLDB Endowment*, *10*(7). https://doi.org/10.14778/3067421.3067425

Hahn, S., & Gronlund, S. D. (2007). Top-down guidance in visual search for facial expressions. *Psychonomic Bulletin and Review*, *14*(1). https://doi.org/10.3758/BF03194044

Hamadicharef, B., Zhang, H., Guan, C., Wang, C., Kok, S. P., Keng, P. T., & Kai, K. A. (2009). Learning EEG-based spectral-spatial patterns for attention level measurement. *Proceedings - IEEE International Symposium on Circuits and Systems*. https://doi.org/10.1109/ISCAS.2009.5118043

Hansen, A.L., Johnsen, B.H., & Thayer, J.F. (2003). Vagal influence on working memory and attention. *International Journal of Psychophysiology*, *48*(3). https://doi.org/10.1016/S0167-8760(03)00073-4

Hansen, C. H., & Hansen, R. D. (1988). Finding the Face in the Crowd: An Anger Superiority Effect. *Journal of Personality and Social Psychology*, *54*(6). https://doi.org/10.1037/0022-3514.54.6.917

Hanson, S. J., Gagliardi, A. D., & Hanson, C. (2009). Solving the brain synchrony eigenvalue problem: Conservation of temporal dynamics (fMRI) over subjects doing the same task. *Journal of Computational Neuroscience*, *27*(1). https://doi.org/10.1007/s10827-008-0129-z

Hasson, U., Avidan, G., Gelbard, H., Vallines, I., Harel, M., Minshew, N., & Behrmann, M. (2009). Shared and idiosyncratic cortical activation patterns in autism revealed under continuous real-life viewing conditions. *Autism Research*, *2*(4). https://doi.org/10.1002/aur.89

Hasson, U., Landesman, O., Knappmeyer, B., Vallines, I., Rubin, N., & Heeger, D. J. (2008). Neurocinematics: The Neuroscience of Film. *Projections*, *2*(1). https://doi.org/10.3167/proj.2008.020102

Hasson, U., Malach, R., & Heeger, D. J. (2010). Reliability of cortical activity during natural stimulation. In *Trends in Cognitive Sciences* (Vol. 14, Issue 1). https://doi.org/10.1016/j.tics.2009.10.011

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject Synchronization of Cortical Activity during Natural Vision. *Science*, *303*(5664), 1634–1640. https://doi.org/10.1126/science.1089506

He, Z., Xu, X., & Deng, S. (2003). Discovering cluster-based local outliers. *Pattern Recognition Letters*, *24*(9–10). https://doi.org/10.1016/S0167-8655(03)00003-5

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *50*(3). https://doi.org/10.3758/s13428-017-0935-1

Helm, J. L., Miller, J. G., Kahle, S., Troxel, N. R., & Hastings, P. D. (2018). On Measuring and Modeling Physiological Synchrony in Dyads. *Multivariate Behavioral Research*, *53*(4), 521–543. https://doi.org/10.1080/00273171.2018.1459292

Henriques, A. S., King, S. C., & Meiselman, H. L. (2009). Consumer segmentation based on food neophobia and its application to product development. *Food Quality and Preference*, *20*(2). https://doi.org/10.1016/j.foodqual.2008.01.003

Hettich, D. T., Bolinger, E., Matuz, T., Birbaumer, N., Rosenstiel, W., & Spüler, M. (2016). EEG responses to auditory stimuli for automatic affect recognition. *Frontiers in Neuroscience*, *10*(JUN). https://doi.org/10.3389/fnins.2016.00244

Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*(4108). https://doi.org/10.1126/science.182.4108.177

Hogervorst, M. A., Brouwer, A. M., & van Erp, J. B. F. (2014). Combining and comparing EEG, peripheral physiology and eye-related measures for the assessment of mental workload. *Frontiers in Neuroscience*, *8*(OCT). https://doi.org/10.3389/fnins.2014.00322

Holmes, A., Nielsen, M. K., & Green, S. (2008). Effects of anxiety on the processing of fearful and happy faces: An event-related potential study. *Biological Psychology*, *77*(2). https://doi.org/10.1016/j.biopsycho.2007.10.003

Holper, L., Scholkmann, F., & Wolf, M. (2012). Between-brain connectivity during imitation measured by fNIRS. *NeuroImage*, *63*(1). https://doi.org/10.1016/j.neuroimage.2012.06.028

Holtze, B., Rosenkranz, M., Jaeger, M., Debener, S., & Mirkovic, B. (2022). Ear-EEG Measures of Auditory Attention to Continuous Speech. *Frontiers in Neuroscience*, *16*(May), 1–14. https://doi.org/10.3389/fnins.2022.869426

Horton, C., Srinivasan, R., & D'Zmura, M. (2014). Envelope responses in single-trial EEG indicate attended speaker in a "cocktail party." *Journal of Neural Engineering*, *11*(4). https://doi.org/10.1088/1741-2560/11/4/046015

Hudson, A. N., Van Dongen, H. P. A., & Honn, K. A. (2020). Sleep deprivation, vigilant attention, and brain function: a review. In *Neuropsychopharmacology* (Vol. 45, Issue 1). https://doi.org/10.1038/

s41386-019-0432-6
Huettel, S. A., Song, A. W., & McCarthy, G. (2014). chapter 4 Functional Magnetic Resonance Imaging, Third Edition. *Functional Magnetic Resonance Imaging, Third Edition*.
Hussain, M. S., Alzoubi, O., Calvo, R. A., & D'Mello, S. K. (2011). Affect detection from multichannel physiology during learning sessions with autotutor. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6738 LNAI*. https://doi.org/10.1007/978-3-642-21869-9_19
Ienca, M., Haselager, P., & Emanuel, E. J. (2018). Brain leaks and consumer neurotechnology. In *Nature Biotechnology* (Vol. 36, Issue 9). https://doi.org/10.1038/nbt.4240
Ikeda, S., Nozawa, T., Yokoyama, R., Miyazaki, A., Sasaki, Y., Sakaki, K., & Kawashima, R. (2017). Steady beat sound facilitates both coordinated group walking and inter-subject neural synchrony. *Frontiers in Human Neuroscience*, *11*. https://doi.org/10.3389/fnhum.2017.00147
Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, *2*(3). https://doi.org/10.1038/35058500
Jääskeläinen, I. P., Koskentalo, K., Balk, M. H., Autti, T., Kauramäki, J., Pomren, C., & Sams, M. (2008). Inter-Subject Synchronization of Prefrontal Cortex Hemodynamic Activity During Natural Viewing. *The Open Neuroimaging Journal*, *2*(1). https://doi.org/10.2174/1874440000802010014
Jaeger, S. R., Chheang, S. L., Jin, D., Ryan, G., & Worch, T. (2021). The negative influence of food neophobia on food and beverage liking: Time to look beyond extreme groups analysis? *Food Quality and Preference*, *92*. https://doi.org/10.1016/j.foodqual.2021.104217
Jaeger, S. R., Rasmussen, M. A., & Prescott, J. (2017). Relationships between food neophobia and food intake and preferences: Findings from a sample of New Zealand adults. *Appetite*, *116*. https://doi.org/10.1016/j.appet.2017.05.030
James, W. (1890). The Principles Of Psychology Volume I By William James (1890). *The Principles of Psychology*, *I*(1890).
Jamet, E., Gavota, M., & Quaireau, C. (2008). Attention guiding in multimedia learning. *Learning and Instruction*, *18*(2). https://doi.org/10.1016/j.learninstruc.2007.01.011
Janssen, T. W. P., Grammer, J. K., Bleichner, M. G., Bulgarelli, C., Davidesco, I., Dikker, S., Jasińska, K. K., Siugzdaite, R., Vassena, E., Vatakis, A., Zion-Golumbic, E., & van Atteveldt, N. (2021). Opportunities and Limitations of Mobile Neuroimaging Technologies in Educational Neuroscience. *Mind, Brain, and Education*, *15*(4), 354–370. https://doi.org/10.1111/mbe.12302
Jiang, Y., & Chun, M. M. (2001). Selective attention modulates implicit learning. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *54*(4). https://doi.org/10.1080/713756001
Kaada, B. R. (1951). Somato-motor, autonomic and electrocorticographic responses to electrical stimulation of rhinencephalic and other structures in primates, cat, and dog; a study of responses from the limbic, subcallosal, orbito-insular, piriform and temporal cortex, hippoc. *Acta Physiologica Scandinavica. Supplementum*, *24*(83).
Kanaan-Izquierdo, S., Ziyatdinov, A., & Perera-Lluna, A. (2018). Multiview and multifeature spectral clustering using common eigenvectors. *Pattern Recognition Letters*, *102*. https://doi.org/10.1016/j.patrec.2017.12.011
Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2000). Principles of Neural Science, fourth addition. *McGraw-Hill Companies*, *4*.
Kaneko, D., Hogervorst, M., Toet, A., van Erp, J. B. F., Kallen, V., & Brouwer, A. M. (2019). Explicit and implicit responses to tasting drinks associated with different tasting experiences. *Sensors (Switzerland)*, *19*(20). https://doi.org/10.3390/s19204397
Kaneko, D., Stuldreher, I., Reuten, A. J. C., Toet, A., van Erp, J. B. F., & Brouwer, A. M. (2021). Comparing Explicit and Implicit Measures for Assessing Cross-Cultural Food Experience. *Frontiers in Neuroergonomics*, *2*(March), 1–16. https://doi.org/10.3389/fnrgo.2021.646280
Kaneko, D., Toet, A., Brouwer, A. M., Kallen, V., & van Erp, J. B. F. (2018). Methods for evaluating emotions evoked by food experiences: A literature review. In *Frontiers in Psychology* (Vol. 9, Issue JUN). https://doi.org/10.3389/fpsyg.2018.00911

Kaneko, D., Toet, A., Ushiama, S., Brouwer, A. M., Kallen, V., & van Erp, J. B. F. (2019). EmojiGrid: A 2D pictorial scale for cross-cultural emotion assessment of negatively and positively valenced food. In *Food Research International* (Vol. 115). https://doi.org/10.1016/j.foodres.2018.09.049

Kanwisher, N., & Wojciulik, E. (2000). Visual attention: Insights from brain imaging. *Nature Reviews Neuroscience*, *1*(2). https://doi.org/10.1038/35039043

Kastner, S., & Ungerleider, L. G. (2001). The neural basis of biased competition in human visual cortex. In *Neuropsychologia* (Vol. 39, Issue 12). https://doi.org/10.1016/S0028-3932(01)00116-6

Ki, J. J., Kelly, S. P., & Parra, L. C. (2016). Attention strongly modulates reliability of neural responses to naturalistic narrative stimuli. *Journal of Neuroscience*, *36*(10), 3092–3101. https://doi.org/10.1523/JNEUROSCI.2942-15.2016

Kjellberg, A. (1977). Sleep deprivation and some aspects of performance: I. Problems of arousal changes. *Waking & Sleeping*, *1*(2).

Knaapila, A., Silventoinen, K., Broms, U., Rose, R. J., Perola, M., Kaprio, J., & Tuorila, H. M. (2011). Food neophobia in young adults: Genetic architecture and relation to personality, pleasantness and use frequency of foods, and body mass index-A twin study. *Behavior Genetics*, *41*(4). https://doi.org/10.1007/s10519-010-9403-8

Konvalinka, I., Xygalatas, D., Bulbulia, J., Schjødt, U., Jegindø, E. M., Wallot, S., Van Orden, G., & Roepstorff, A. (2011). Synchronized arousal between performers and related spectators in a fire-walking ritual. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(20). https://doi.org/10.1073/pnas.1016955108

Koole, S. L., Atzil-Slonim, D., Butler, E., Dikker, S., Tschacher, W., & Wilderjans, T. (2020). In Sync with Your Shrink. In *Applications of Social Psychology*. https://doi.org/10.4324/9780367816407-9

Kruskal, J. B. (1964). Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, *29*(2). https://doi.org/10.1007/BF02289694

Kuppens, P., Tuerlinckx, F., Russell, J. A., & Barrett, L. F. (2013). The relation between valence and arousal in subjective experience. *Psychological Bulletin*, *139*(4). https://doi.org/10.1037/a0030811

Lane, R. D., Reiman, E. M., Bradley, M. M., Lang, P. J., Ahern, G. L., Davidson, R. J., & Schwartz, G. E. (1997). Neuroanatomical correlates of pleasant and unpleasant emotion. *Neuropsychologia*, *35*(11). https://doi.org/10.1016/S0028-3932(97)00070-5

Lang, P. J. (1995). The Emotion Probe: Studies of Motivation and Attention. *American Psychologist*, *50*(5). https://doi.org/10.1037/0003-066X.50.5.372

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). Motivated attention: Affect, activation, and action. In *Attention and Orienting: Sensory and Motivational Processes*.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, motivation, and anxiety: Brain mechanisms and psychophysiology. *Biological Psychiatry*, *44*(12). https://doi.org/10.1016/S0006-3223(98)00275-3

Lankinen, K., Saari, J., Hari, R., & Koskinen, M. (2014). Intersubject consistency of cortical MEG signals during movie viewing. *NeuroImage*, *92*. https://doi.org/10.1016/j.neuroimage.2014.02.004

Levenson, R. W., & Gottman, J. M. (1983). Marital interaction: Physiological linkage and affective exchange. *Journal of Personality and Social Psychology*, *45*(3). https://doi.org/10.1037/0022-3514.45.3.587

Li, X., Li, X., & Luo, Y. J. (2005). Anxiety and attentional bias for threat: An event-related potential study. *NeuroReport*, *16*(13). https://doi.org/10.1097/01.wnr.0000176522.26971.83

Lim, J., & Dinges, D. F. (2008). Sleep deprivation and vigilant attention. *Annals of the New York Academy of Sciences*, *1129*. https://doi.org/10.1196/annals.1417.002

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: A meta-analytic review. In *Behavioral and Brain Sciences* (Vol. 35, Issue 3). https://doi.org/10.1017/S0140525X11000446

Linssen, L., Landman, A., van Baardewijk, J. U., Bottenheft, C., & Binsch, O. (2022). Using accelerometry and heart rate data for real-time monitoring of soldiers' stress in a dynamic military virtual reality scenario. *Multimedia Tools and Applications*, *81*(17). https://doi.org/10.1007/s11042-022-12705-6

Liu, J., Zhang, R., Geng, B., Zhang, T., Yuan, D., Otani, S., & Li, X. (2019). Interplay between prior knowledge and communication mode on teaching effectiveness: Interpersonal neural synchronization as a neural marker. *NeuroImage*, *193*. https://doi.org/10.1016/j.neuroimage.2019.03.004

Liu, N. H., Chiang, C. Y., & Chu, H. C. (2013). Recognizing the degree of human attention using EEG signals from mobile sensors. *Sensors (Switzerland)*, *13*(8). https://doi.org/10.3390/s130810273

Liu, S., Zhou, Y., Palumbo, R., & Wang, J. L. (2016). Dynamical correlation: A new method for quantifying synchrony with multivariate intensive longitudinal data. *Psychological Methods*, *21*(3), 291–308. https://doi.org/10.1037/met0000071

Liu, Y., Ayaz, H., & Shewokis, P. A. (2017). Multisubject "learning" for mental workload classification using concurrent EEG, fNIRS, and physiological measures. *Frontiers in Human Neuroscience*, *11*(July). https://doi.org/10.3389/fnhum.2017.00389

Liu, Y., Wang, T., Wang, K., & Zhang, Y. (2021). Collaborative Learning Quality Classification Through Physiological Synchrony Recorded by Wearable Biosensors. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.674369

Lloyd, S. P. (1982). Least Squares Quantization in PCM. *IEEE Transactions on Information Theory*, *28*(2). https://doi.org/10.1109/TIT.1982.1056489

Loftus, E. F., Loftus, G. R., & Messo, J. (1987). Some facts about "weapon focus." *Law and Human Behavior*, *11*(1). https://doi.org/10.1007/BF01044839

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*(6843). https://doi.org/10.1038/35084005

Looney, D., Kidmose, P., Park, C., Ungstrup, M., Rank, M., Rosenkranz, K., & Mandic, D. (2012). The in-the-ear recording concept: User-centered and wearable brain monitoring. *IEEE Pulse*, *3*(6). https://doi.org/10.1109/MPUL.2012.2216717

Lotte, F., Bougrain, L., Cichocki, A., Clerc, M., Congedo, M., Rakotomamonjy, A., & Yger, F. (2018). A review of classification algorithms for EEG-based brain-computer interfaces: A 10 year update. In *Journal of Neural Engineering* (Vol. 15, Issue 3). https://doi.org/10.1088/1741-2552/aab2f2

Maaoui, C., & Pruski, A. (2018). Unsupervised stress detection from remote physiological signal. *Proceedings of the IEEE International Conference on Industrial Technology*, *2018-February*. https://doi.org/10.1109/ICIT.2018.8352409

Madsen, J., Margulis, E. H., Simchy-Gross, R., & Parra, L. C. (2019). Music synchronizes brainwaves across listeners with strong effects of repetition, familiarity and training. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-40254-w

Madsen, J., & Parra, L. C. (2021). Similar cognitive processing synchronizes brains, hearts, and eyes. *BioRxiv*. https://www.biorxiv.org/content/early/2021/09/20/2021.09.16.460722

Madsen, J., & Parra, L. C. (2022). Cognitive processing of a common stimulus synchronizes brains, hearts, and eyes. *PNAS Nexus*, *1*(1), 1–14. https://doi.org/10.1093/pnasnexus/pgac020

Madsen, J., & Parra, L. C. (2023). Brain-body interaction during natural story listening. *BioRxiv*, 2023.01.31.526511. https://www.biorxiv.org/content/10.1101/2023.01.31.526511v1%0Ahttps://www.biorxiv.org/content/10.1101/2023.01.31.526511v1.abstract

Malmberg, J., Järvelä, S., Holappa, J., Haataja, E., Huang, X., & Siipo, A. (2019). Going beyond what is visible: What multichannel data can reveal about interaction in the context of collaborative learning? *Computers in Human Behavior*, *96*. https://doi.org/10.1016/j.chb.2018.06.030

Mandernach, B. J. (2015). Assessment of student engagement in higher education: A synthesis of literature and assessment tools. *International Journal of Learning, Teaching and Educational Research*, *12*(2).

Mäntylä, T., Nummenmaa, L., Rikandi, E., Lindgren, M., Kieseppä, T., Hari, R., Suvisaari, J., & Raij, T. T. (2018). Aberrant Cortical Integration in First-Episode Psychosis During Natural Audiovisual Processing. *Biological Psychiatry*, *84*(9). https://doi.org/10.1016/j.biopsych.2018.04.014

Maratos, F. A., & Staples, P. (2015). Attentional biases towards familiar and unfamiliar foods in children. The role of food neophobia. *Appetite*, *91*. https://doi.org/10.1016/j.appet.2015.04.003

Marchegiani, L., Karadoğan, S. G., Andersen, T., Larsen, J., & Hansen, L. K. (2011). The role of top-down attention in the cocktail party: Revisiting Cherry's experiment after sixty years. *Proceedings -*

10th International Conference on Machine Learning and Applications, ICMLA 2011, *1*. https://doi.org/10.1109/ICMLA.2011.143

Marci, C. D., Ham, J., Moran, E., & Orr, S. P. (2007). Physiologic correlates of perceived therapist empathy and social-emotional process during psychotherapy. *Journal of Nervous and Mental Disease*, *195*(2), 103–111. https://doi.org/10.1097/01.nmd.0000253731.71025.fc

Marci, C. D., & Orr, S. P. (2006). The effect of emotional distance on psychophysiologic concordance and perceived empathy between patient and interviewer. *Applied Psychophysiology Biofeedback*, *31*(2). https://doi.org/10.1007/s10484-006-9008-4

Marcus, E. M., Jacobson, S., & Sabin, T. D. (2015). Hypothalamus, Neuroendocrine System, and Autonomic Nervous System. In *Integrated Neuroscience and Neurology*. https://doi.org/10.1093/med/9780199744435.003.0016

McAssey, M. P., Helm, J., Hsieh, F., Sbarra, D. A., & Ferrer, E. (2013). Methodological advances for detecting physiological synchrony during dyadic interactions. *Methodology*, *9*(2), 41–53. https://doi.org/10.1027/1614-2241/a000053

McInnes, L., Healy, J., Saul, N., & Großberger, L. (2018). UMAP: Uniform Manifold Approximation and Projection. *Journal of Open Source Software*, *3*(29), 861. https://doi.org/10.21105/joss.00861

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2009). Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies. In *Learning Unbound: Select Research and Analyses of Distance Education and Online Learning*.

Mecacci, G., & Haselager, P. (2019). Identifying Criteria for the Evaluation of the Implications of Brain Reading for Mental Privacy. *Science and Engineering Ethics*, *25*(2). https://doi.org/10.1007/s11948-017-0003-3

Miró, E., Cano-Lozano, M. C., & Buela-Casal, G. (2002). Electrodermal activity during total sleep deprivation and its relationship with other activation and performance measures. *Journal of Sleep Research*, *11*(2). https://doi.org/10.1046/j.1365-2869.2002.00286.x

Mitkidis, P., McGraw, J. J., Roepstorff, A., & Wallot, S. (2015). Building trust: Heart rate synchrony and arousal during joint action increased by public goods game. *Physiology and Behavior*, *149*. https://doi.org/10.1016/j.physbeh.2015.05.033

Mobbs, D., Yu, R., Meyer, M., Passamonti, L., Seymour, B., Calder, A. J., Schweizer, S., Frith, C. D., & Dalgleish, T. (2009). A key role for similarity in vicarious reward. *Science*, *324*(5929). https://doi.org/10.1126/science.1170539

Mohanty, A., & Sussman, T. J. (2013). Top-down modulation of attention by emotion. *Frontiers in Human Neuroscience*, *MAR*. https://doi.org/10.3389/fnhum.2013.00102

Moser, J. S., Huppert, J. D., Duval, E., & Simons, R. F. (2008). Face processing biases in social anxiety: An electrophysiological study. *Biological Psychology*, *78*(1). https://doi.org/10.1016/j.biopsycho.2008.01.005

Mühlberger, A., Wieser, M. J., Herrmann, M. J., Weyers, P., Tröger, C., & Pauli, P. (2009). Early cortical processing of natural and artificial emotional faces differs between lower and higher socially anxious persons. *Journal of Neural Transmission*, *116*(6). https://doi.org/10.1007/s00702-008-0108-6

Mukhopadhyay, S. C. (2015). Wearable sensors for human activity monitoring: A review. In *IEEE Sensors Journal* (Vol. 15, Issue 3). https://doi.org/10.1109/JSEN.2014.2370945

Müller, V., & Lindenberger, U. (2011). Cardiac and respiratory patterns synchronize between persons during choir singing. *PLoS ONE*, *6*(9). https://doi.org/10.1371/journal.pone.0024893

Müller, V., Sänger, J., & Lindenberger, U. (2013). Intra- and Inter-Brain Synchronization during Musical Improvisation on the Guitar. *PLoS ONE*, *8*(9). https://doi.org/10.1371/journal.pone.0073852

Näätänen, R. (1988). Implications of ERP data for psychological theories of attention. *Biological Psychology*, *26*(1–3). https://doi.org/10.1016/0301-0511(88)90017-8

Najafi, T. A., Affanni, A., Rinaldo, R., & Zontone, P. (2023). Driver Attention Assessment Using Physiological Measures from EEG, ECG, and EDA Signals †. *Sensors*, *23*(4), 2039. https://doi.org/10.3390/s23042039

Näpflin, M., Wildi, M., & Sarnthein, J. (2007). Test-retest reliability of resting EEG spectra validates a statistical signature of persons. *Clinical Neurophysiology*, *118*(11). https://doi.org/10.1016/j.

clinph.2007.07.022

Neafsey, E. J. (1991). Chapter 7 Prefrontal cortical control of the autonomic nervous system: Anatomical and physiological observations. *Progress in Brain Research*, *85*(C). https://doi.org/10.1016/S0079-6123(08)62679-5

Newmann, F., Wehlage, G., & Lamborn, S. (1992). The significance and sources of student engagement. In *Student engagement and achievement in American secondary schools*.

Nijs, I. M. T., Franken, I. H. A., & Muris, P. (2009). Enhanced processing of food-related pictures in female external eaters. *Appetite*, *53*(3). https://doi.org/10.1016/j.appet.2009.07.022

Nummenmaa, L., Glerean, E., Viinikainen, M., Jääskeläinen, I. P., Hari, R., & Sams, M. (2012). Emotions promote social interaction by synchronizing brain activity across individuals. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(24). https://doi.org/10.1073/pnas.1206095109

O'Connell, R. G., Dockree, P. M., Robertson, I. H., Bellgrove, M. A., Foxe, J. J., & Kelly, S. P. (2009). Uncovering the neural signature of lapsing attention: Electrophysiological signals predict errors up to 20 s before they occur. *Journal of Neuroscience*, *29*(26). https://doi.org/10.1523/JNEUROSCI.5967-08.2009

O'Sullivan, J. A., Power, A. J., Mesgarani, N., Rajaram, S., Foxe, J. J., Shinn-Cunningham, B. G., Slaney, M., Shamma, S. A., & Lalor, E. C. (2015). Attentional Selection in a Cocktail Party Environment Can Be Decoded from Single-Trial EEG. *Cerebral Cortex*, *25*(7). https://doi.org/10.1093/cercor/bht355

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General*, *130*(3). https://doi.org/10.1037/0096-3445.130.3.466

Palumbo, R. V., Marraccini, M. E., Weyandt, L. L., Wilder-Smith, O., McGee, H. A., Liu, S., & Goodwin, M. S. (2017). Interpersonal Autonomic Physiology: A Systematic Review of the Literature. *Personality and Social Psychology Review*, *21*(2), 99–141. https://doi.org/10.1177/1088868316628405

Palva, S., & Palva, J. M. (2007). New vistas for α-frequency band oscillations. *Trends in Neurosciences*, *30*(4). https://doi.org/10.1016/j.tins.2007.02.001

Pan, J., & Tompkins, W. J. (1985). A Real-Time QRS Detection Algorithm. *IEEE Transactions on Biomedical Engineering*, *BME-32*(3). https://doi.org/10.1109/TBME.1985.325532

Pan, Y., Dikker, S., Goldstein, P., Zhu, Y., Yang, C., & Hu, Y. (2020). Instructor-learner brain coupling discriminates between instructional approaches and predicts learning. *NeuroImage*, *211*(August 2019). https://doi.org/10.1016/j.neuroimage.2020.116657

Parasuraman, R. (2003). Neuroergonomics: Research and practice. *Theoretical Issues in Ergonomics Science*, *4*(1–2). https://doi.org/10.1080/14639220210199753

Parra, L. C., Haufe, S., & Dmochowski, J. P. (2019). Correlated Components Analysis - Extracting Reliable Dimensions in Multivariate Data. *Neurons, Behavior, Data Analysis, and Theory*, *2*(1). https://doi.org/10.51628/001c.7125

Pashler, H., Johnston, J. C., & Ruthruff, E. (2001). Attention and Performance. *Annual Review of Psychology*, *52*, 629–651. https://doi.org/10.1146/annurev.psych.52.1.629

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1). https://doi.org/10.3758/s13428-018-01193-y

Pérez, P., Madsen, J., Banellis, L., Türker, B., Raimondo, F., Perlbarg, V., Valente, M., Niérat, M. C., Puybasset, L., Naccache, L., Similowski, T., Cruse, D., Parra, L. C., & Sitt, J. D. (2021). Conscious processing of narrative stimuli synchronizes heart rate between individuals. *Cell Reports*, *36*(11). https://doi.org/10.1016/j.celrep.2021.109692

Pessoa, L., Kastner, S., & Ungerleider, L. G. (2003). Neuroimaging studies of attention: From modulation of sensory processing to top-down control. In *Journal of Neuroscience* (Vol. 23, Issue 10). https://doi.org/10.1523/jneurosci.23-10-03990.2003

Picton, T. W. (1992). The P300 wave of the human event-related potential. *Journal of Clinical Neurophysiology*, *9*(4). https://doi.org/10.1097/00004691-199210000-00002

Pilcher, J. J., Band, D., Odle-Dusseau, H. N., & Muth, E. R. (2007). Human performance under sustained operations and 'acute sleep deprivation conditions: Toward a model of controlled attention.

Aviation Space and Environmental Medicine, *78*(5 II).

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, *198*. https://doi.org/10.1016/j.neuroimage.2019.05.026

Pliner, P., & Hobden, K. (1992). Development of a scale to measure the trait of food neophobia in humans. *Appetite*, *19*(2), 105–120. https://doi.org/10.1016/0195-6663(92)90014-W

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. In *Clinical Neurophysiology* (Vol. 118, Issue 10). https://doi.org/10.1016/j.clinph.2007.04.019

Polich, J., & Kok, A. (1995). Cognitive and biological determinants of P300: an integrative review. *Biological Psychology*, *41*(2). https://doi.org/10.1016/0301-0511(95)05130-9

Porges, S. W. (2001). The polyvagal theory: Phylogenetic substrates of a social nervous system. *International Journal of Psychophysiology*, *42*(2). https://doi.org/10.1016/S0167-8760(01)00162-3

Posada-Quintero, H. F., Bolkhovsky, J. B., Reljin, N., & Chon, K. H. (2017). Sleep deprivation in young and healthy subjects is more sensitively identified by higher frequencies of electrodermal activity than by skin conductance level evaluated in the time domain. *Frontiers in Physiology*, *8*(JUN). https://doi.org/10.3389/fphys.2017.00409

Poulsen, A. T., Kamronn, S., Dmochowski, J., Parra, L. C., & Hansen, L. K. (2017). EEG in the classroom: Synchronised neural recordings during video presentation. *Scientific Reports*, *7*, 1–9. https://doi.org/10.1038/srep43916

Prehn-Kristensen, A., Wiesner, C., Bergmann, T. O., Wolff, S., Jansen, O., Mehdorn, H. M., Ferstl, R., & Pause, B. M. (2009). Induction of empathy by the smell of anxiety. *PLoS ONE*, *4*(6). https://doi.org/10.1371/journal.pone.0005987

Pugh, K. R., Shaywitz, B. A., Shaywitz, S. E., Fulbright, R. K., Byrd, D., Skudlarski, P., Shankweiler, D. P., Katz, L., Constable, R. T., Fletcher, J., Lacadie, C., Marchione, K., & Gore, J. C. (1996). Auditory selective attention: An fMRI investigation. *NeuroImage*, *4*(3). https://doi.org/10.1006/nimg.1996.0067

Quer, G., Daftari, J., & Rao, R. R. (2016). Heart rate wavelet coherence analysis to investigate group entrainment. *Pervasive and Mobile Computing*, *28*, 21–34. https://doi.org/10.1016/j.pmcj.2015.09.008

Ragot, M., Martin, N., Em, S., Pallamin, N., & Diverrez, J. M. (2018). Emotion recognition using physiological signals: Laboratory vs. wearable sensors. *Advances in Intelligent Systems and Computing*, *608*. https://doi.org/10.1007/978-3-319-60639-2_2

Raudenbush, B., & Capiola, A. (2012). Physiological responses of food neophobics and food neophilics to food and non-food stimuli. *Appetite*, *58*(3). https://doi.org/10.1016/j.appet.2012.02.042

Raudenbush, B., Corley, N., Flower, N. R., Kozlowski, A., & Meyer, B. (2003). Cephalic phase salivary response differences characterize level of food neophobia. *Appetite*, *41*(2). https://doi.org/10.1016/S0195-6663(03)00059-X

Raudenbush, B., & Frank, R. A. (1999). Assessing food neophobia: The role of stimulus familiarity. *Appetite*, *32*(2). https://doi.org/10.1006/appe.1999.0229

Ray, W. J., & Cole, H. W. (1985). EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes. *Science*, *228*(4700). https://doi.org/10.1126/science.3992243

Reeves, B., Thorson, E., Rothschild, M. L., Mcdonald, D., Hirsch, J., & Goldstein, R. (1985). Attention to television: Intrastimulus effects of movement and scene changes on alpha variation Over time. *International Journal of Neuroscience*, *27*(3–4). https://doi.org/10.3109/00207458509149770

Reinero, D. A., Dikker, S., & Van Bavel, J. J. (2021). Inter-brain synchrony in teams predicts collective performance. *Social Cognitive and Affective Neuroscience*, *16*(1–2), 43–57. https://doi.org/10.1093/scan/nsaa135

Richards, G. P., Samuels, S. J., Turnure, J. E., & Ysseldyke, J. E. (1990). Sustained and selective attention in children with learning disabilities. *Journal of Learning Disabilities*, *23*(2). https://doi.org/10.1177/002221949002300210

Rosenberg, M., Noonan, S., DeGutis, J., & Esterman, M. (2013). Sustaining visual attention in the face of distraction: A novel gradual-onset continuous performance task. *Attention, Perception, and Psychophysics*, *75*(3). https://doi.org/10.3758/s13414-012-0413-x

Rosenkranz, M., Holtze, B., Jaeger, M., & Debener, S. (2021). EEG-Based Intersubject Correlations Reflect

Selective Attention in a Competing Speaker Scenario. *Frontiers in Neuroscience*, *15*(June), 1–12. https://doi.org/10.3389/fnins.2021.685774

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*(C). https://doi.org/10.1016/0377-0427(87)90125-7

Rozin, P., & Fallon, A. (1980). The psychological categorization of foods and non-foods: A preliminary taxonomy of food rejections. *Appetite*, *1*(3). https://doi.org/10.1016/S0195-6663(80)80027-4

Sabu, P., Stuldreher, I. V., & Brouwer, A. M. (2022). An Attempt to Assess the Effects of Social Demand using Explicit and Implicit Measures of Food Experience. *Volume 2 of the Proceedings of the Joint 12th International Conference on Methods and Techniques in Behavioral Research and 6th Seminar on Behavioral Methods Held Online, May 18-20, 2022*, 143–146.

Salmi, J., Roine, U., Glerean, E., Lahnakoski, J., Nieminen-Von Wendt, T., Tani, P., Leppämäki, S., Nummenmaa, L., Jääskeläinen, I. P., Carlson, S., Rintahaka, P., & Sams, M. (2013). The brains of high functioning autistic individuals do not synchronize with those of others. *NeuroImage: Clinical*, *3*. https://doi.org/10.1016/j.nicl.2013.10.011

Sandall, B. K. (2016). Wearable technology and schools. Where are we and were do we go from here. *Journal of Curriculum, Teaching, Learning and Leadership in Education:*, *1*(1), 74–83.

Sbarra, D. A., & Hazan, C. (2008). Coregulation, dysregulation, self-regulation: An integrative analysis and empirical agenda for understanding adult attachment, separation, loss, and recovery. *Personality and Social Psychology Review*, *12*(2). https://doi.org/10.1177/1088868308315702

Schickenberg, B., Van Assema, P., Brug, J., & De Vries, N. K. (2008). Are the Dutch acquainted with and willing to try healthful food products? The role of food neophobia. *Public Health Nutrition*, *11*(5). https://doi.org/10.1017/S1368980007000778

Schmälzle, R., Häcker, F. E. K., Honey, C. J., & Hasson, U. (2014). Engaged listeners: Shared neural processing of powerful political speeches. *Social Cognitive and Affective Neuroscience*, *10*(8), 1137–1143. https://doi.org/10.1093/scan/nsu168

Schmidt, R. (1992). Awareness and Second Language Acquisition. *Annual Review of Applied Linguistics*, *13*. https://doi.org/10.1017/s0267190500002476

Schubert, E., Sander, J., Ester, M., Kriegel, H. P., & Xu, X. (2017). DBSCAN revisited, revisited: Why and how you should (still) use DBSCAN. *ACM Transactions on Database Systems*, *42*(3). https://doi.org/10.1145/3068335

Schultze-Kraft, M., Dahne, S., Gugler, M., Curio, G., & Blankertz, B. (2016). Unsupervised classification of operator workload from brain signals. *Journal of Neural Engineering*, *13*(3). https://doi.org/10.1088/1741-2560/13/3/036008

Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Cacioppo, J. T., Tiffany, I., & Lang, P. J. (2000). Affective picture processing: The late positive potential is modulated by motivational relevance. *Psychophysiology*, *37*(2). https://doi.org/10.1017/S0048577200001530

Schupp, H. T., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2003). Emotional facilitation of sensory processing in the visual cortex. *Psychological Science*, *14*(1). https://doi.org/10.1111/1467-9280.01411

Seshadri, D. R., Li, R. T., Voos, J. E., Rowbottom, J. R., Alfes, C. M., Zorman, C. A., & Drummond, C. K. (2019). Wearable sensors for monitoring the physiological and biochemical profile of the athlete. In *npj Digital Medicine* (Vol. 2, Issue 1). https://doi.org/10.1038/s41746-019-0150-9

Siegrist, M., Hartmann, C., & Keller, C. (2013). Antecedents of food neophobia and its association with eating behavior and food choices. *Food Quality and Preference*, *30*(2). https://doi.org/10.1016/j.foodqual.2013.06.013

Silver, R., & Parente, R. (2004). The psychological and physiological dynamics of a simple conversation. *Social Behavior and Personality*, *32*(5). https://doi.org/10.2224/sbp.2004.32.5.413

Simpson, J. R., Öngür, D., Akbudak, E., Conturo, T. E., Ollinger, J. M., Snyder, A. Z., Gusnard, D. A., & Raichle, M. E. (2000). The emotional modulation of cognitive processing: An fMRI study. *Journal of Cognitive Neuroscience*, *12*(SUPPL. 2). https://doi.org/10.1162/089892900564019

Sinaga, K. P., & Yang, M. S. (2020). Unsupervised K-means clustering algorithm. *IEEE Access*, *8*. https://doi.org/10.1109/ACCESS.2020.2988796

Smith, D. B. D., Donchin, E., Cohen, L., & Starr, A. (1970). Auditory averaged evoked potentials in man during selective binaural listening. *Electroencephalography and Clinical Neurophysiology*, *28*(2). https://doi.org/10.1016/0013-4694(70)90182-3

Smith, J. C., Bradley, M. M., Scott, R. P., & Lang, P. J. (2004). The Psychophysiology of Emotion. *Medicine & Science in Sports & Exercise*, *36*(Supplement). https://doi.org/10.1097/00005768-200405001-00432

Smith, R., Thayer, J. F., Khalsa, S. S., & Lane, R. D. (2017). The hierarchical basis of neurovisceral integration. In *Neuroscience and Biobehavioral Reviews* (Vol. 75). https://doi.org/10.1016/j.neubiorev.2017.02.003

Snoek, A., Stuldreher, I. V., Horstkötter, D., Haselager, P., & Brouwer, A. M. (2022). *Wearables that track mental states in the classroom: ethical reflections*.

Snyder, E., & Hillyard, S. A. (1976). Long-latency evoked potentials to irrelevant, deviant stimuli. *Behavioral Biology*, *16*(3). https://doi.org/10.1016/S0091-6773(76)91447-4

Soleymani, M., Asghari-Esfeden, S., Fu, Y., & Pantic, M. (2016). Analysis of EEG Signals and Facial Expressions for Continuous Emotion Detection. *IEEE Transactions on Affective Computing*, *7*(1). https://doi.org/10.1109/TAFFC.2015.2436926

Soleymani, M., Pantic, M., & Pun, T. (2012). Multimodal emotion recognition in response to videos. *IEEE Transactions on Affective Computing*, *3*(2). https://doi.org/10.1109/T-AFFC.2011.37

Steiger, B. K., Kegel, L. C., Spirig, E., & Jokeit, H. (2019). Dynamics and diversity of heart rate responses to a disaster motion picture. *International Journal of Psychophysiology*, *143*(May), 64–79. https://doi.org/10.1016/j.ijpsycho.2019.06.015

Strang, A. J., Funke, G. J., Russell, S. M., Dukes, A. W., & Middendorf, M. S. (2014). Physio-behavioral coupling in a cooperative team task: Contributors and relations. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(1). https://doi.org/10.1037/a0033125

Stuldreher, I. V., van Erp, J. B. F., & Brouwer, A. M. (2023a). Robustness of physiological synchrony in wearable electrodermal activity and heart rate as a measure of attentional engagement to movie clips. *Sensors*, *23*(6), 3006. https://doi.org/%0A10.3390/s23063006

Stuldreher, I. V., de Winter, J. C. F., Thammasan, N., & Brouwer, A. M. (2019). Analytic approaches for the combination of autonomic and neural activity in the assessment of physiological synchrony. *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, *2019-Octob*. https://doi.org/10.1109/SMC.2019.8913927

Stuldreher, I. V., Kaneko, D., Hiraguchi, H., van Erp, J. B. F., & Brouwer, A. M. (2023b). EEG measures of attention toward food-related stimuli vary with food neophobia. *Food Quality and Preference*, *106*(September 2022), 104805. https://doi.org/10.1016/j.foodqual.2022.104805

Stuldreher, I. V., Thammasan, N., van Erp, J. B. F., & Brouwer, A. M. (2020b). Physiological Synchrony in EEG, Electrodermal Activity and Heart Rate Detects Attentionally Relevant Events in Time. *Frontiers in Neuroscience*, *14*(December), 1–11. https://doi.org/10.3389/fnins.2020.575521

Stuldreher, I. V., Thammasan, N., Van Erp, J. B. F., & Brouwer, A. M. (2020a). Physiological synchrony in EEG, electrodermal activity and heart rate reflects shared selective auditory attention. *Journal of Neural Engineering*, *17*(4). https://doi.org/10.1088/1741-2552/aba87d

Suess, P. E., Porges, S. W., & Plude, D. J. (1994). Cardiac vagal tone and sustained attention in school-age children. *Psychophysiology*, *31*(1). https://doi.org/10.1111/j.1469-8986.1994.tb01020.x

Suveg, C., Shaffer, A., & Davis, M. (2016). Family stress moderates relations between physiological and behavioral synchrony and child self-regulation in mother-preschooler dyads. *Developmental Psychobiology*, *58*(1). https://doi.org/10.1002/dev.21358

Svaldi, J., Tuschen-Caffier, B., Peyk, P., & Blechert, J. (2010). Information processing of food pictures in binge eating disorder. *Appetite*, *55*(3). https://doi.org/10.1016/j.appet.2010.10.002

Thammasan, N., Stuldreher, I. V., Schreuders, E., Giletta, M., & Brouwer, A. M. (2020). A usability study of physiological measurement in school using wearable sensors. *Sensors (Switzerland)*, *20*(18), 1–24. https://doi.org/10.3390/s20185380

Thayer, J. F., Hansen, A. L., Saus-Rose, E., & Johnsen, B. H. (2009). Heart rate variability, prefrontal neural function, and cognitive performance: The neurovisceral integration perspective on self-regulation, adaptation, and health. In *Annals of Behavioral Medicine* (Vol. 37, Issue 2). https://doi.org/10.1007/s12160-009-9101-z

Thayer, J. F., & Lane, R. D. (2009). Claude Bernard and the heart-brain connection: Further elaboration of a model of neurovisceral integration. In *Neuroscience and Biobehavioral Reviews* (Vol. 33, Issue 2). https://doi.org/10.1016/j.neubiorev.2008.08.004

Thierry, G., & Roberts, M. V. (2007). Event-related potential study of attention capture by affective sounds. *NeuroReport*, *18*(3). https://doi.org/10.1097/WNR.0b013e328011dc95

Thomann, J., Baumann, C. R., Landolt, H., & Werth, E. (2014). Psychomotor Vigilance Task Demonstrates Impaired Vigilance. *Journal of Clincal Sleep Medicine*, *10*(9), 1019–1024.

Tickle-Degnen, L., & Rosenthal, R. (1990). The Nature of Rapport and Its Nonverbal Correlates. *Psychological Inquiry*, *1*(4), 285–293. https://doi.org/10.1207/s15327965pli0104_1

Toet, A., Kaneko, D., Kruijf, I. de, Ushiama, S., van Schaik, M. G., Brouwer, A. M., Kallen, V., & van Erp, J. B. F. (2019). CROCUFID: A cross-cultural food image database for research on food elicited affective responses. *Frontiers in Psychology*, *10*(JAN). https://doi.org/10.3389/fpsyg.2019.00058

Toet, A., Kaneko, D., Ushiama, S., Hoving, S., Kruijf, I. de, Brouwer, A. M., Kallen, V., & van Erp, J. B. F. (2018). EmojiGrid: A 2D pictorial scale for the assessment of food elicited emotions. *Frontiers in Psychology*, *9*(NOV). https://doi.org/10.3389/fpsyg.2018.02396

Tourunen, A., Kykyri, V. L., Seikkula, J., Kaartinen, J., Tolvanen, A., & Penttonen, M. (2020). Sympathetic nervous system synchrony: An exploratory study of its relationship with the therapeutic alliance and outcome in couple therapy. *Psychotherapy*, *57*(2). https://doi.org/10.1037/pst0000198

Trendafilov, N. T. (2010). Stepwise estimation of common principal components. *Computational Statistics and Data Analysis*, *54*(12). https://doi.org/10.1016/j.csda.2010.03.010

Tronstad, C., Amini, M., Bach, D. R., & Martinsen, Ø. G. (2022). Current trends and opportunities in the methodology of electrodermal activity measurement. In *Physiological Measurement* (Vol. 43, Issue 2). https://doi.org/10.1088/1361-6579/ac5007

Tuorila, H., Lähteenmäki, L., Pohjalainen, L., & Lotti, L. (2001). Food neophobia among the Finns and related responses to familiar and unfamiliar foods. *Food Quality and Preference*, *12*(1). https://doi.org/10.1016/S0950-3293(00)00025-2

Van Beers, J. J., Stuldreher, I. V., Thammasan, N., & Brouwer, A. M. (2020). A Comparison between Laboratory and Wearable Sensors in the Context of Physiological Synchrony. *ICMI 2020 - Proceedings of the 2020 International Conference on Multimodal Interaction*, 604–608. https://doi.org/10.1145/3382507.3418837

Van Den Berg, J., & Neely, G. (2006). Performance on a simple reaction time task while sleep deprived. *Perceptual and Motor Skills*, *102*(2). https://doi.org/10.2466/PMS.102.2.589-599

van Lier, H. G., Pieterse, M. E., Garde, A., Postel, M. G., de Haan, H. A., Vollenbroek-Hutten, M. M. R., Schraagen, J. M., & Noordzij, M. L. (2020). A standardized validity assessment protocol for physiological signals from wearable technology: Methodological underpinnings and an application to the E4 biosensor. *Behavior Research Methods*, *52*(2). https://doi.org/10.3758/s13428-019-01263-9

Verdiere, K. J., Albert, M., Dehais, F., & Roy, R. N. (2020). Physiological Synchrony Revealed by Delayed Coincidence Count: Application to a Cooperative Complex Environment. *IEEE Transactions on Human-Machine Systems*, *50*(5). https://doi.org/10.1109/THMS.2020.2986417

Verma, G. K., & Tiwary, U. S. (2014). Multimodal fusion framework: A multiresolution approach for emotion classification and recognition from physiological signals. In *NeuroImage* (Vol. 102, Issue P1). https://doi.org/10.1016/j.neuroimage.2013.11.007

Von Luxburg, U. (2007). A tutorial on spectral clustering. *Statistics and Computing*, *17*(4). https://doi.org/10.1007/s11222-007-9033-z

Vortmann, L. M., Kroll, F., & Putze, F. (2019). EEG-Based Classification of Internally- and Externally-Directed Attention in an Augmented Reality Paradigm. *Frontiers in Human Neuroscience*, *13*. https://doi.org/10.3389/fnhum.2019.00348

Vuilleumier, P., & Driver, J. (2007). Modulation of visual processing by attention and emotion: Windows on causal interactions between human brain regions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481). https://doi.org/10.1098/rstb.2007.2092

Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, *58*(301). https://doi.org/10.1080/01621459.1963.10500845

Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in My insula: The common neural basis of seeing and feeling disgust. *Neuron*, *40*(3). https://doi.org/10.1016/S0896-6273(03)00679-2

Williams, H. L., Lubin, A., & Goodnow, J. J. (1959). Impaired performance with acute sleep loss. *Psychological Monographs: General and Applied*, *73*(14). https://doi.org/10.1037/h0093749

Williams, M. A., Moss, S. A., Bradshaw, J. L., & Mattingley, J. B. (2005). Look at me, I'm smiling: Visual search for threatening and nonthreatening facial expressions. *Visual Cognition*, *12*(1). https://doi.org/10.1080/13506280444000193

Wilson, S. J., Bailey, B. E., Jaremka, L. M., Fagundes, C. P., Andridge, R., Malarkey, W. B., Gates, K. M., & Kiecolt-Glaser, J. K. (2018). When couples' hearts beat together: Synchrony in heart rate variability during conflict predicts heightened inflammation throughout the day. *Psychoneuroendocrinology*, *93*. https://doi.org/10.1016/j.psyneuen.2018.04.017

Wilson, S. M., Molnar-Szakacs, I., & Iacoboni, M. (2008). Beyond superior temporal cortex: Intersubject correlations in narrative speech comprehension. *Cerebral Cortex*, *18*(1). https://doi.org/10.1093/cercor/bhm049

Winkler, I., Debener, S., Muller, K. R., & Tangermann, M. (2015). On the influence of high-pass filtering on ICA-based artifact reduction in EEG-ERP. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, *2015-November*. https://doi.org/10.1109/EMBC.2015.7319296

Winkler, I., Haufe, S., & Tangermann, M. (2011). Automatic Classification of Artifactual ICA-Components for Artifact Removal in EEG Signals. *Behavioral and Brain Functions*, *7*. https://doi.org/10.1186/1744-9081-7-30

Wollmer, M., Weninger, F., Knaup, T., Schuller, B., Sun, C., Sagae, K., & Morency, L. P. (2013). You tube movie reviews: Sentiment analysis in an audio-visual context. *IEEE Intelligent Systems*, *28*(3), 46–53. https://doi.org/10.1109/MIS.2013.34

Woltering, S., Lishak, V., Elliott, B., Ferraro, L., & Granic, I. (2015). Dyadic Attunement and Physiological Synchrony During Mother-Child Interactions: An Exploratory Study in Children With and Without Externalizing Behavior Problems. *Journal of Psychopathology and Behavioral Assessment*, *37*(4). https://doi.org/10.1007/s10862-015-9480-3

Yu, S. S., Chu, S. W., Wang, C. M., Chan, Y. K., & Chang, T. C. (2018). Two improved k-means algorithms. *Applied Soft Computing Journal*, *68*. https://doi.org/10.1016/j.asoc.2017.08.032

Yun, K., Watanabe, K., & Shimojo, S. (2012). Interpersonal body and neural synchronization as a marker of implicit social interaction. *Scientific Reports*, *2*. https://doi.org/10.1038/srep00959

Zander, T. O., & Kothe, C. (2011). Towards passive brain-computer interfaces: Applying brain-computer interface technology to human-machine systems in general. *Journal of Neural Engineering*, *8*(2). https://doi.org/10.1088/1741-2560/8/2/025005

Zehetleitner, M., Goschy, H., & Müller, H. J. (2012). Top-down control of attention: It's gradual, practice-dependent, and hierarchically organized. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(4). https://doi.org/10.1037/a0027629

Zhang, J., Wang, K., & Zhang, Y. (2021). Physiological Characterization of Student Engagement in the Naturalistic Classroom: A Mixed-Methods Approach. *Mind, Brain, and Education*, *15*(4). https://doi.org/10.1111/mbe.12300

Zheng, L., Chen, C., Liu, W., Long, Y., Zhao, H., Bai, X., Zhang, Z., Han, Z., Liu, L., Guo, T., Chen, B., Ding, G., & Lu, C. (2018). Enhancement of teaching outcome through neural prediction of the students' knowledge state. *Human Brain Mapping*, *39*(7). https://doi.org/10.1002/hbm.24059

Zotev, V., Yuan, H., Misaki, M., Phillips, R., Young, K. D., Feldner, M. T., & Bodurka, J. (2016). Correlation between amygdala BOLD activity and frontal EEG asymmetry during real-time fMRI neurofeedback training in patients with depression. *NeuroImage: Clinical*, *11*. https://doi.org/10.1016/j.nicl.2016.02.003

About the author

Ivo Stuldreher was born on the 26th of October 1996 in Alphen aan den Rijn, the Netherlands. In 2019 he obtained his MSc in biomechanical engineering (cum laude) from the Delft University of Technology. He completed his master's program with an internship at the Netherlands Organisation for Applied Scientific Research (TNO) in Soesterberg, the Netherlands.

After his graduation, Ivo started his professional career at the Perceptual and Cognitive Systems department, currently the Human Performance department, of TNO in Soesterberg. His research at TNO is focused on obtaining implicit measures of individuals' mental states, mainly through the use of physiological measurements. At TNO, scientists from different backgrounds, such as engineering, computer science, psychology and neuroscience, work together on evolving applied scientific research. Ivo's engineering background allows him to apply technical knowledge, for instance about signal processing, to psychophysiological research. In his research Ivo aims to find innovative ways to extract information on mental states from physiological measures. Physiological synchrony, presented in this thesis, is one of those approaches. Ivo is a strong believer in pursuing research together in teams. He is also an advocate of open research. Many of the data that have been collected and scripts that have been made in the context of this dissertation have therefore also been published online for other researchers to use.

Ivo lives together with his partner Fenna in Utrecht, the Netherlands. In his spare time Ivo likes to be outside on one of his bicycles.

List of publications

Journal articles

Stuldreher, I. V., Thammasan, N., van Erp, J. B. F., Brouwer, A. M. (2020). Physiological synchrony in EEG, electrodermal activity and heart rate reflects shared selective auditory attention. *Journal of Neural Engineering, 17*(4), 046028. doi: 10.1088/1741-2552/aba87d

Thammasan, N., **Stuldreher, I. V.**, Schreuders, E., Giletta, M., Brouwer, A. M. (2020). A usability study of physiological measurement in school using wearable sensors. *Sensors, 20*(18), 5380. doi: 10.3390/s20185380

Stuldreher, I. V., Thammasan, N., Van Erp, J. B. F., Brouwer, A. M. (2020). Physiological synchrony in EEG, electrodermal activity and heart rate detects attentionally relevant events in time. *Frontiers in Neuroscience, 14*, 1257. doi: 10.3389/fnins.2020.575521

Bottenheft, C., Brouwer, A. M., **Stuldreher, I. V.**, Groen, E. L., van Erp, J. B. F. (2020). Cognitive task performance under (combined) conditions of a metabolic and sensory stressor. *Cognition, Technology & Work, 23*, 805-817. doi: 10.1007/s10111-020-00653-w

Brouwer, A. M., van Beers, J. J., Sabu, P., **Stuldreher, I. V.**, Zech, H. G., Kaneko, D. (2021). Measuring Implicit Approach–Avoidance Tendencies towards *Food Using a Mobile Phone Outside the Lab. Foods, 10*(7), 1440. doi: 10.3390/foods10071440

Kaneko, D., **Stuldreher, I. V.**, Reuten, A. J., Toet, A., Van Erp, J. B. F., Brouwer, A. M. (2021). Comparing explicit and implicit measures for assessing cross-cultural food experience. *Frontiers in Neuroergonomics, 2*, 646280. doi: 10.3389/fnrgo.2021.646280

Stuldreher, I. V., Merasli, A., Thammasan, N., Van Erp, J. B. F., Brouwer, A. M. (2021). Unsupervised clustering of individuals sharing selective attentional focus using physiological synchrony. *Frontiers in Neuroergonomics, 2*, 750248. doi: 10.3389/fnins.2020.575521

Sabu, P., **Stuldreher, I. V.**, Kaneko, D., Brouwer, A. M. (2022). A review on the role of affective stimuli in event-related frontal alpha asymmetry. *Frontiers in Computer Science, 4*, 869123. doi: 10.3389/fcomp.2022.869123

Stuldreher, I. V., van Erp, J. B. F., & Brouwer, A. M. (2023). Robustness of physiological synchrony in wearable electrodermal activity and heart rate as a measure of attentional engagement to movie clips. *Sensors, 23*(6), 3006. doi: 10.3390/s23063006

Stuldreher, I. V., Kaneko, D., Hiraguchi, H., van Erp, J. B. F., & Brouwer, A. M. (2023). EEG measures of attention toward food-related stimuli vary with food neophobia. *Food Quality and Preference, 106*, 104805. doi: 10.1016/j.foodqual.2022.104805

Stuldreher, I. V., Maasland, E., Bottenheft, C., Van Erp, J. B. F., Brouwer, A. M. (2023). Physiological synchrony in electrodermal activity predicts decreased vigilant attention induced by sleep deprivation. *Frontiers in Neuroergonomics, 4*, 1199347. doi: 10.3389/fnrgo.2023.1199347

Bottenheft, C., Hogenelst, K., **Stuldreher, I. V.**, Kleemann, R., Groen, E. L., van Erp, J. B. F., Brouwer, A. M. (2023). Understanding the combined effects of sleep deprivation and acute social stress on cognitive performance using a comprehensive approach. *Brain, Behavior, & Immunity – Health, 34*, 100706. doi: 10.1016/j.bbih.2023.100706

Lorenz, T. I., Schreuders, E., **Stuldreher, I. V.**, Thammasan, N., Brouwer, A. M., & Giletta, M. (2023). The interplay of peer victimization and parasympathetic nervous system activity on acute inflammatory stress responses in adolescence. *Research on Child and Adolescent Psychopathology*. doi: 10.1007/s10802-023-01156-8

Stuldreher I. V., van der Burg, E., Velut, S., Toet, A., Hiraguchi, H., Hogervorst, M. A., Zandstra, E. H., van Erp, J. B. F., Brouwer, A. M. (in press). Electrodermal activity as an index of food neophobia outside the lab. *Frontiers in Neuroergonomics.*

Houben, M. M. J., **Stuldreher I. V.**, Forbes, P. A., Groen, E. L. (in press). Using galvanic vestibular stimulation to induce post-roll illusion in a fixed-base flight simulator. *Aerospace Medicine and Human Performance.*

Conference proceedings and presentations

Stuldreher, I. V., de Winter, J. C. F., Thammasan, N., Brouwer, A. M. (2019). Analytic approaches for the combination of autonomic and neural activity in the assessment of physiological synchrony. In *Proceedings of 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC).* 4143 – 4148.

Brouwer, A. M., **Stuldreher, I. V.**, Thammasan, N. (2019). Shared attention reflected in EEG, electrodermal activity and heart rate. In *Proceedings of the Workshop Socio-Affective Technologies: an interdisciplinary approach co-located with IEEE SMC 2019.* 27-31.

Thammasan, N., **Stuldreher, I. V.**, Wismeijer, D., Poel, M., van Erp, J. B. F., Brouwer, A. M. (2019). A novel, simple and objective method to detect movement artefacts in electrodermal activity. In *Proceedings of 2019 International Conference on Affective Computing and Intelligent Interaction (ACII).* 371 – 377

Brouwer, A. M. **Stuldreher, I. V.**, Penen, S. H., Lingelbach, K., Vukelić, M. (2020). Combining eye tracking and physiology for detection of emotion and workload. In *Proceedings of the 12th International Conference on Measurement and Behaviour and 6th International Seminar on Behavioral Methods, Krakow, Poland, 2021;* Volume 1, 2–11. doi: 10.6084/m9.figshare.13013717

Borovac, A., **Stuldreher, I. V.**, Thammasan, N., Brouwer, A. M. (2020). Validation of wearables for electrodermal activity (EdaMove) and heart rate (Wahoo Tickr). In *Proceedings of the 12th International Conference on Measurement and Behaviour and 6th International Seminar on Behavioral Methods, Krakow, Poland, 2021;* Volume 1, 18–24. doi: 10.6084/m9.figshare.13013717

Stuldreher, I. V., Thammasan, N., van Erp, J. B. F., Brouwer, A. M. (2020). Physiological synchrony in EEG, electrodermal activity and heart rate for the assessment of shared attention. In *Proceedings of the 12th International Conference on Measurement and Behaviour and 6th International Seminar on Behavioral Methods, Krakow, Poland, 2021;* Volume 1, 12–17. doi: 10.6084/m9.figshare.13013717

Merasli A., **Stuldreher, I. V.**, Brouwer, A. M. (2020). Unsupervised clustering of groups with different selective attentional instructions using physiological synchrony. *In Proceedings of the 32nd Benelux Conference on Artificial Intelligence, BNAIC 2020 and 29th Annual Belgian-Dutch Conference on Machine Learning, BeneLearn 2020*, Leiden, The Netherlands. 428-429

van Beers, J. J., Kaneko, D., **Stuldreher, I. V.**, Zech, H. G., Brouwer, A. M. (2020). An accessible tool to measure implicit approach-avoidance tendencies towards food outside the lab. In *Proceedings of the 2020 International Conference on Multimodal Interaction*, 307-311. doi: 10.1145/3382507.3418837

Stuldreher, I. V. (2020). Multimodal Physiological Synchrony as Measure of Attentional Engagement. In *Proceedings of the 2020 International Conference on Multimodal Interaction*, 718-722. doi: 10.1145/3382507.3421152

Verhagen, P., Kuling, I., Gijsbertse, K., **Stuldreher, I. V.**, Overvliet, K., Falcone, S., Van Erp, J. B. F., Brouwer, A. M. (2020). The cross-modal congruency effect as an objective measure of embodiment. In *Proceedings of the 2020 International Conference on Multimodal Interaction*, 107-111. doi: 10.1145/3395035.3425264

Stuldreher, I. V., Thammasan, N., Schreuders, E., Giletta, M., Tjew-A-Sin, M., de Geus, E. J., Van Erp, J. B. F., Brouwer, A. M. (2021). Physiological synchrony in the classroom. In *Proceedings of the Neuroergonomics Conference 2021 September 11-16, München*.

Lingelbach, K., Piechnik, D., Brouwer, A. M., **Stuldreher, I. V.**, Vukelić, M. (2021). Temporal Decoding of Emotion and Workload from Fixation-Related EEG Recordings. In *Proceedings of the Neuroergonomics Conference 2021 September 11-16, München*.

Falcone, S., Brouwer, A. M., Heylen, D., Van Erp, J. B. F., Zhang, L., Pradhan, S. S., **Stuldreher, I. V.**, Cocu, I., Heuvel, M., de Vries, P. S., Gijsbertse, K., Englebienne, G. (2022). Pupil diameter as implicit measure to estimate sense of embodiment. In *Proceedings of the Annual Meeting of the Cognitive Science Society, 44*(44). 2446-2453.

Tolston, M. T., **Stuldreher, I. V.**, Funke, G. J., Brouwer, A. M. (2021). Using interpersonal similarity in complex networks from physiological data to assess attentional focus. In *Proceedings of the HFM-334 Symposium on "Applying Neuroscience to Performance: From Rehabilitation to Human Cognitive Augmentation.*

Falcone, S., Englebienne, G., Brouwer, A. M., Zhang, L., Pradhan, S., **Stuldreher, I. V.**, Cocu, I., Heuvel, M. de Vries, P. S., Gijsbertse, K., Heylen, D., van Erp, J. B. F. (2022). Assessing the Pupil Dilation as Implicit Measure of the Sense of Embodiment in Two User Studies. In *Volume 2 of the Proceedings of the joint 12th International Conference on Methods and Techniques in Behavioral Research and 6th Seminar of Behavioural Methods, online*. 89-94. doi: 10.6084/m9.figshare.20066849

Sabu, P., Kaneko, D., **Stuldreher, I. V.**, Brouwer, A. M. (2022). An Attempt to Assess the Effects of Social Demand using Explicit and Implicit Measures of Food Experience. In *Volume 2 of the Proceedings of the joint 12th International Conference on Methods and Techniques in Behavioral Research and 6th Seminar of Behavioural Methods, online*. 143-146. doi: 10.6084/m9.figshare.20066849

Bottenheft, C., **Stuldreher, I. V.**, Hogenelst, K., Groen, E., van Erp, J., Kleemann, R., Brouwer, A. M. (2022). Understanding the effects of sleep deprivation and acute social stress on cognitive performance using a comprehensive approach. In *Volume 2 of the Proceedings of the joint 12th International Conference on Methods and Techniques in Behavioral Research and 6th Seminar of Behavioural Methods, online*. 114-115. doi: 10.6084/m9.figshare.20066849

Stuldreher, I. V., van Erp, J. B. F., & Brouwer, A. M. (2022). The Effects of Stimulus Duration and Group Size on Wearable Physiological Synchrony. In *Volume 2 of the Proceedings of the joint 12th International Conference on Methods and Techniques in Behavioral Research and 6th Seminar of Behavioural Methods, online*. 44-46. doi: 10.6084/m9.figshare.20066849

Stuldreher, I. V., van Erp, J. B. F., Brouwer, A. M. (2022). *Interpersonal synchrony in electrodermal activity predicts decreased performance in a vigilance task induced by sleep deprivation*. The Third Neuroadaptive Technology Conference, Lübbenau, Germany. 20-22

Stuldreher, I. V., Roijendijk, L., Michel, M. Toet, A. (2022). Gaze behavior as an objective measure to assess social presence during immersive mediated communication. In *Proceedings of the 8th international conference on human interaction and emerging technologies (IHIET), Nice, France*. 300-307. doi: 10.54941/ahfe1002746

Ledegang, W. D., van der Burg, E., **Stuldreher, I. V.**, Houben, M. M. J., Groen, E. J., van der Horst, D., Starmans, E. A. M., Almekinders, G. (2023). Acquiring manual flying skills in a virtual reality flight simulator. In *Proceedings of the 22nd International Symposium of Aviation Psychology, Rochester, NY, United States of America*. 275-281

Stuldreher, I. V., van der Burg, E., Ledegang, W. D., Houben, M. M. J., Groen, E. J., van der Horst, D., Starmans, E. A. M., Almekinders, G. (2023). Measuring the lookout behavior of student pilots in a virtual reality flight simulator. In *Proceedings of the 22nd International Symposium of Aviation Psychology, Rochester, NY, United States of America*. 288-293

Stuldreher I. V., van Erp, J. B. F., Brouwer, A. M. (2023). *Inter-individual synchrony reflects natural variations in attention*. 14th International Conference on Applied Human Factors and Ergonomics, San Francisco, CA, United States of America.

Stuldreher I. V., van der Burg, E., Velut, S., van Erp, J. B. F., Zandstra, E. H., Hiraguchi, H., Hogervorst, M. A., van Os, D., Brouwer, A. M. (2023). *Interpersonal physiological synchrony in skin conductance when attending to a pitch at a festival*. 14th International Conference on Applied Human Factors and Ergonomics, San Francisco, CA, United States of America.

Stuldreher I. V., De Groes, N., van der Stigchel, B., Reinten, J., Michel, M. (2023). *A mixed-reality environment for the assessment of situational awareness of the dismounted soldier*. The 6th International Conference on Soldiers' Physical Performance, London, United Kingdom.

Ziegveld, L., Michel, M., **Stuldreher, I. V.**, Dijkstra-Soudarissanane, S., Niamut, O. (2023) *A framework for assessing and enhancing presence in (a)symmetrical remote collaboration*. The 20th EuroXR international conference, Rotterdam, The Netherlands

Hiraguchi, H., van der Burg, E., **Stuldreher, I. V.**, Toet, A., Velut, S., Zandstra, E. H., van Os, D., Hogervorst, M. A., van Erp, J. B. F., Brouwer, A. M. (2023) Effects of Multisensory Contexts on Tofu and Soy Sauce Evaluation and Consumption. *Biological Life Science Forum 2023*, 26, 36. doi: 10.3390/Foods2023-15059

9 789036 559362