

Shared responses and individual differences in the human brain during naturalistic stimulations

Edited by

Zhishan Hu, Xin Di and Zhi Yang

Published in

Frontiers in Human Neuroscience



FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-8325-2480-0
DOI 10.3389/978-2-8325-2480-0

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Shared responses and individual differences in the human brain during naturalistic stimulations

Topic editors

Zhishan Hu — Shanghai Jiao Tong University, China

Xin Di — New Jersey Institute of Technology, United States

Zhi Yang — Shanghai Key Laboratory of Psychotic Disorders, Shanghai Mental Health Center, China

Citation

Hu, Z., Di, X., Yang, Z., eds. (2023). *Shared responses and individual differences in the human brain during naturalistic stimulations*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-2480-0

Table of contents

- 04 **Editorial: Shared responses and individual differences in the human brain during naturalistic stimulations**
Zhishan Hu, Xin Di and Zhi Yang
- 07 **Interbrain Synchrony of Team Collaborative Decision-Making: An fNIRS Hyperscanning Study**
Mingming Zhang, Huibin Jia and Guanghai Wang
- 17 **Shared and Unshared Feature Extraction in Major Depression During Music Listening Using Constrained Tensor Factorization**
Xiulin Wang, Wenya Liu, Xiaoyu Wang, Zhen Mu, Jing Xu, Yi Chang, Qing Zhang, Jianlin Wu and Fengyu Cong
- 28 **Behavioral Experience-Sampling Methods in Neuroimaging Studies With Movie and Narrative Stimuli**
Iiro P. Jääskeläinen, Jyrki Ahveninen, Vasily Klucharev, Anna N. Shestakova and Jonathan Levy
- 35 **Influence of Auditory Cues on the Neuronal Response to Naturalistic Visual Stimuli in a Virtual Reality Setting**
George Al Boustani, Lennart Jakob Konstantin Weiß, Hongwei Li, Svea Marie Meyer, Lukas Hiendlmeier, Philipp Rinklin, Bjoern Menze, Werner Hemmert and Bernhard Wolfrum
- 48 **Type of bilingualism conditions individual differences in the oscillatory dynamics of inhibitory control**
Sergio Miguel Pereira Soares, Yanina Prystauka, Vincent DeLuca and Jason Rothman
- 68 **Auditory and cross-modal attentional bias toward positive natural sounds: Behavioral and ERP evidence**
Yanmei Wang, Zhenwei Tang, Xiaoxuan Zhang and Libing Yang
- 88 **Applying functional near-infrared spectroscopy and eye-tracking in a naturalistic educational environment to investigate physiological aspects that underlie the cognitive effort of children during mental rotation tests**
Raimundo da Silva Soares Jr., Amanda Yumi Ambriola Oku, Cândida S. F. Barreto and João Ricardo Sato
- 102 **The specificity, situational modulations, and behavioral correlates of parent-child neural synchrony**
Yi Liu, Jiaxin Li, Qi Wang and Yarong Li
- 117 **High-resolution surface electromyographic activities of facial muscles during mimic movements in healthy adults: A prospective observational study**
Nadiya Mueller, Vanessa Trentzsch, Roland Grassme, Orlando Guntinas-Lichius, Gerd Fabian Volk and Christoph Anders



OPEN ACCESS

EDITED AND REVIEWED BY
László Négyessy,
Hungarian Academy of Sciences, Hungary

*CORRESPONDENCE
Zhishan Hu
✉ huzhishan@me.com

RECEIVED 07 April 2023
ACCEPTED 20 April 2023
PUBLISHED 05 May 2023

CITATION
Hu Z, Di X and Yang Z (2023) Editorial: Shared responses and individual differences in the human brain during naturalistic stimulations. *Front. Hum. Neurosci.* 17:1201728. doi: 10.3389/fnhum.2023.1201728

COPYRIGHT
© 2023 Hu, Di and Yang. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Editorial: Shared responses and individual differences in the human brain during naturalistic stimulations

Zhishan Hu^{1*}, Xin Di² and Zhi Yang^{1,3,4}

¹Neuroimaging Core, Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai, China, ²Department of Biomedical Engineering, New Jersey Institute of Technology, Newark, NJ, United States, ³Beijing Key Laboratory of Mental Disorders, National Clinical Research Center for Mental Disorders and National Center for Mental Disorders, Beijing Anding Hospital, Capital Medical University, Beijing, China, ⁴Advanced Innovation Center for Human Brain Protection, Capital Medical University, Beijing, China

KEYWORDS

shared response, individual difference, naturalistic stimulation, psychiatry, education

Editorial on the Research Topic

Shared responses and individual differences in the human brain during naturalistic stimulations

Conventional functional brain imaging experiments present stimuli in a controlled and repetitive manner based on predefined protocols. Although offering crucial insights into brain function, their real-world relevance is uncertain. Investigating brain activity in naturalistic settings can help fill the gaps (Nastase et al., 2019; Finn et al., 2020).

Following Hasson et al.'s pioneering work on shared brain activity during naturalistic movie viewing (Hasson et al., 2004), researchers have made substantial progress in this field. On one hand, researchers have found that shared brain activity can reflect the multifaceted cognitive and emotional information processing properties of the human brain, such as shared understanding, empathy, mathematical ability (Cantlon and Li, 2013), cognitive development (Di and Biswal, 2022), education (Hu et al., 2019; Meshulam et al., 2021), interpersonal relationships (Parkinson et al., 2018), and even brain characteristics of mental illness (Hasson et al., 2009; Zhang et al., 2022; Jin et al., 2023). On the other hand, shared brain activity has emerged as a valuable tool for identifying individual differences (Finn et al., 2020; Di et al., 2023). By detecting the spatiotemporal characteristics of the emergence of shared brain activity in healthy groups or reference groups, researchers can construct reference or normative models that enable them to make inferences about the degree of deviation from the characteristics of brain activity in individuals (Yang et al., 2020). Naturalistic stimuli with rich social-emotional information are able to express individual differences stably in many brain regions (Gao et al., 2020), providing important support for individualized inferences based on brain responses.

The current Research Topic encompasses studies that employ a naturalistic paradigm and a variety of brain imaging and behavioral measuring devices, including functional near-infrared spectroscopy (fNIRS), electroencephalograph (EEG), surface electromyography (sEMG), virtual reality device, and eye-tracking glasses. By leveraging the signals from these devices and utilizing cutting edge data analysis approaches, their studies have significantly deepened our understanding of how the brain responds to stimuli in real-world settings.

This collection includes two review articles. [Jääskeläinen et al.](#) provided an overview of inner experiences sampling techniques and the joint analyses of these measurements with neuroimaging data. Their review highlighted the potential for developing neurophenomenological frameworks to understand the relationship between inner experiences and their neural underpinnings. [Liu et al.](#)'s review focused on the parent-child inter-brain neural synchrony (INS) and its impact on children's social development. Their review demonstrated stronger parent-child INS as compared to the stranger-child dyads, discussed situational factors that influence the parent-child INS, and explored how INS relates to behavioral tendencies and individual features of both parents and children.

The studies featured in this collection showed that stimuli with more naturalistic properties can effectively elicit shared response. Using the sound from the natural environment, for example, [Wang Y. et al.](#) found positive emotional naturalistic sounds could modulate visual spatial attention. The event-related potential (ERP) revealed enlarged N1 amplitudes after positive sounds than neutral sounds. Similarly, in a naturalistic setting, [Mueller et al.](#) used sEMG to measure the facial muscles during mimic movements, and found that facial movement tasks evoked shared response in the complex network of facial muscles rather than activating a single muscle. Further, [Boustani et al.](#) employed virtual reality technique to study the brain-computer interface (BCI) performance in a real-world environment. When the additional auditory cues were presented, they observed increased classification accuracy for a detection task using brain activity features, compared to the visual-only condition, highlighting the importance of multisensory integration in BCI performance.

The naturalistic paradigm can be valuable in educational settings. For instance, [Zhang et al.](#) used a prisoner's dilemma game combined with the fNIRS-based hyper-scanning technique to explore the neural basis underlying team collaborative decision-making. Higher INS was identified in the right inferior frontal gyrus during team collaborative decision-making. Similarly, [Soares et al.](#) used fNIRS and eye-tracking to examine the physiological responses of children during a mental rotation test in a naturalistic education environment. They observed significant activation on the dorsolateral prefrontal cortex and increased pupil diameter during the task. The brain and physiological indicators discovered in these studies can be valuable for neurocomputational modeling of cognition in a naturalistic environment. In addition, [Pereira Soares et al.](#) found that individual differences in real life bilingual language experience affect oscillatory dynamics of inhibition and cognitive control.

The investigation of shared response and individual differences in the human brain in naturalistic environments holds potential for research in psychiatry. [Wang K. et al.](#) analyzed EEG signals in patients with major depression disorder (MDD) and healthy controls (HC) while freely listening to music. They found that both groups triggered similar brain dynamics when listening to music, but MDD patients also exhibited some alterations in brain

oscillatory network characteristics during music perception. These findings could potentially aid in the clinical diagnosis and treatment of MDD patients.

In conclusion, the utilization of naturalistic paradigms using diverse brain imaging and behavioral measuring devices can provide valuable insights into shared responses and individual differences in the human brain. The research articles in this collection highlight the potential of neurophenomenological frameworks, the significance of multisensory integration in brain-computer interface performance, and the applicability of naturalistic paradigms in educational and psychiatric settings.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

ZH is supported by National Natural Science Foundation of China (62007002). XD is supported by (US) National Institutes of Mental Health (R15MH125332). ZY is supported by National Natural Science Foundation of China (81971682, 81571756, and 81270023), Natural Science Foundation of Shanghai (20ZR1472800), Shanghai Municipal Commission of Health (2018BR17 and 2019ZB0201), and Shanghai Clinical Research Center for Mental Health (19MC1911100).

Acknowledgments

We thank the authors for their contributions to this Research Topic.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Cantlon, J. F., and Li, R. (2013). Neural activity during natural viewing of sesame street statistically predicts test scores in early childhood. *PLoS Biology*. 11, e1001462. doi: 10.1371/journal.pbio.1001462
- Di, X., and Biswal, B. B. (2022). Principal component analysis reveals multiple consistent responses to naturalistic stimuli in children and adults. *Hum. Brain Mapp.* 43, 3332–3345. doi: 10.1002/hbm.25568
- Di, X., Xu, T., Uddin, L. Q., and Biswal, B. B. (2023). Individual differences in time-varying and stationary brain connectivity during movie watching from childhood to early adulthood: effects of age, sex, and behavioral associations. *bioRxiv*. (2023). doi: 10.1101/2023.01.30.526311
- Finn, E. S., Glerean, E., Khojandi, A. Y., Nielson, D., Molfese, P. J., Handwerker, D. A., et al. (2020). Idiosyncrony: from shared responses to individual differences during naturalistic neuroimaging. *NeuroImage*. 215, 116828. doi: 10.1016/j.neuroimage.2020.116828
- Gao, J., Chen, G., Wu, J., Wang, Y., Hu, Y., Xu, T., et al. (2020). Reliability map of individual differences reflected in inter-subject correlation in naturalistic imaging. *NeuroImage*. 223, 117277. doi: 10.1016/j.neuroimage.2020.117277
- Hasson, U., Avidan, G., Gelbard, H., Vallines, I., Harel, M., Minshew, N., et al. (2009). Shared and idiosyncratic cortical activation patterns in autism revealed under continuous real-life viewing conditions. *Autism Res.* 2, 220–231. doi: 10.1002/aur.89
- Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., and Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science*. 303, 1634. doi: 10.1126/science.1089506
- Hu, Z., Lam, K.-F., Xiang, Y.-T., and Yuan, Z. (2019). Causal cortical network for arithmetic problem-solving represents brain's planning rather than reasoning. *Int J Biol Sci.* 15, 1148–1160. doi: 10.7150/ijbs.33400
- Jin, S., Liu, W., Hu, Y., Liu, Z., Xia, Y., Zhang, X., et al. (2023). Aberrant functional connectivity of the bed nucleus of the stria terminalis and its age dependence in children and adolescents with social anxiety disorder. *Asian J. Psychiatry*. 103498. doi: 10.1016/j.ajp.2023.103498
- Meshulam, M., Hasenfratz, L., Hillman, H., Liu, Y.-F., Nguyen, M., Norman, K. A., et al. (2021). Neural alignment predicts learning outcomes in students taking an introduction to computer science course. *Nat Commun.* 12, 1922. doi: 10.1038/s41467-021-22202-3
- Nastase, S. A., Gazzola, V., Hasson, U., and Keysers, C. (2019). Measuring shared responses across subjects using intersubject correlation. *Social Cognitive and Affective Neurosci.* 14, 667–685. doi: 10.1093/scan/nsz037
- Parkinson, C., Kleinbaum, A. M., and Wheatley, T. (2018). Similar neural responses predict friendship. *Nat Commun.* 9, 332. doi: 10.1038/s41467-017-02722-7
- Yang, Z., Wu, J., Xu, L., Deng, Z., Tang, Y., Gao, J., et al. (2020). Individualized psychiatric imaging based on inter-subject neural synchronization in movie watching. *NeuroImage*. 216, 116227. doi: 10.1016/j.neuroimage.2019.116227
- Zhang, Q., Li, B., Jin, S., Liu, W., Liu, J., Xie, S., et al. (2022). Comparing the Effectiveness of Brain Structural Imaging, Resting-state fMRI, and Naturalistic fMRI in Recognizing Social Anxiety Disorder in Children and Adolescents. *Psychiat. Res.* 111485. doi: 10.1016/j.psychresns.2022.111485



Interbrain Synchrony of Team Collaborative Decision-Making: An fNIRS Hyperscanning Study

Mingming Zhang^{1*}, Huibin Jia^{2,3} and Guanghai Wang⁴

¹ Department of Psychology, College of Education, Shanghai Normal University, Shanghai, China, ² Department of Psychology, Henan University, Kaifeng, China, ³ Key Laboratory of Child Development and Learning Science, Ministry of Education, Southeast University, Nanjing, China, ⁴ Department of Developmental and Behavioral Pediatrics, Shanghai Children's Medical Center, Pediatric Translational Medicine Institute, Shanghai Jiao Tong University School of Medicine, Shanghai, China

OPEN ACCESS

Edited by:

Zhishan Hu,
Beijing Normal University, China

Reviewed by:

Qi Han Zhang,
Tianjin Normal University, China
Lian Duan,
Shenzhen University, China
Bin Xuan,
Anhui Normal University, China

*Correspondence:

Mingming Zhang
zmm2019@shnu.edu.cn

Specialty section:

This article was submitted to
Brain Imaging and Stimulation,
a section of the journal
Frontiers in Human Neuroscience

Received: 30 April 2021

Accepted: 21 June 2021

Published: 15 July 2021

Citation:

Zhang M, Jia H and Wang G
(2021) Interbrain Synchrony of Team
Collaborative Decision-Making: An
fNIRS Hyperscanning Study.
Front. Hum. Neurosci. 15:702959.
doi: 10.3389/fnhum.2021.702959

In many situations, decision-making behaviors are mostly composed of team patterns (i.e., more than two persons). However, brain-based models that inform how team interactions contribute and impact team collaborative decision-making (TCDM) behavior, is lacking. To examine the neural substrates activated during TCDM in realistic, interpersonal interaction contexts, dyads were asked to model TCDM toward their opponent, in a multi-person prisoner's dilemma game, while neural activity was measured using functional near infrared spectroscopy. These experiments resulted in two main findings. First, there are different neural substrates between TCDM and ISDM, which were modulated by social environmental cues. i.e., the low incentive reward yielded higher activation within the left inferior frontal gyrus (IFG), in individual separately decision-making (ISDM) stage while the dorsolateral prefrontal cortex (DLPFC) and the middle frontopolar area was activated in TCDM stage. The high incentive reward evoked a higher interbrain synchrony (IBS) value in the right IFG in TCDM stage. Second, males showed higher activation in the DLPFC and the middle frontopolar area during ISDM, while females evoked higher IBS in the right IFG during TCDM. These sex effects suggest that in individual social dilemma situations, males and females may separately depend on non-social and social cognitive ability to make decisions, while in the social interaction situations of TCDM, females may depend on both social and non-social cognitive abilities. This study provide a compelling basis and interesting perspective for future neuroscience work of TCDM behaviors.

Keywords: team collaborative decision-making, fNIRS, hyperscanning, sex effects, interbrain synchrony

INTRODUCTION

Cooperation refers to the behavior/intention of individuals or groups to collaborative in order to achieve a common goal and to promote an outcome that is both beneficial to themselves and to others (Decety et al., 2004). Previous studies have mainly focused on cooperative behavior in social dilemmas, e.g., "joint decision-making" behavior (hereafter JDM), where each individual receives a higher payoff for defecting than for cooperating, but as a whole, are better off if they choose to cooperate (Dawes and Messick, 2000). Nevertheless, in real life, social dilemmas are mostly

deliberated and resolved through team/group patterns where most of the decisions are made by more than two people. This type of decision has been defined as team/group collaborative decision-making (TCDM) (Bahrami et al., 2010). Although the goals of TCDM and JDM are to obtain the benefit of cooperation, TCDM does not have contradictions between individual interests and team interests (the goals and the interests of individuals in one team are the same), while in JDM behavior, individuals may give up cooperation within the team in order to pursue a higher payoff.

Previous research has paid extensive attention to JDM behavior. Researchers have unraveled the related cortical regions of cooperative behavior in social dilemmas situations based on abundance researches. In these studies, prefrontal regions, orbitofrontal cortex (OFC), ventromedial prefrontal cortex (VMPFC), dorsal medial prefrontal cortex (DMPFC), right dorsal lateral prefrontal cortex (DLPFC), left parietal operculum and anterior cingulate cortex (ACC), were most commonly reported (Rilling et al., 2002, 2008; Sanfey et al., 2003). However, previous study only focused on the exploration of JDM behavior in single brain frameworks, where solitary participants cooperate with a computer or a person via internet. The information processing across the two interacting brains is thus out of reach from the single-brain frame studies (Hasson et al., 2012; Liu and Pelowski, 2014).

With the development of the hyperscanning technique, the research tendency of cooperative behavior has shifted from experimental single-brain, to a natural multi-brain framework (Hasson et al., 2012; Schilbach et al., 2013). These studies have revealed increased synchronized activity in the right superior frontal cortices and the medial prefrontal region when performing joint actions (Funane et al., 2011; Cui et al., 2012; Dommer et al., 2012), the right temporo-parietal junction (TPJ) during face to face economic cooperation (Tang et al., 2016), and anterior cingulate cortex and prefrontal areas synchrony, between brains of dyad teams playing the prisoner's dilemma game (PDG) (Astolfi et al., 2011).

Previous studies on TCDM are mostly from the perspective of management, which fully investigated the influence of the attributes and composition of team members on TCDM, such as the influence of emotional intelligence, psychological safety of individual, status difference and the number of team members on TCDM (Carton and Cummings, 2012, 2013; Awino, 2013; Owens et al., 2015; Zhou et al., 2020). Still, a brain-based model that informs how collaborative-team interaction contributes and impacts the outcome of decision-making is needed.

Here, we studied the neural mechanisms of cooperative decision-making behavior, especially TCDM behavior in a two-person interaction, via functional near-infrared spectroscopy (fNIRS) based hyperscanning. Our experimental paradigm adopts an improved dilemma-game paradigm; two participants sitting side by side act as collaborators to play a computer-based PDG, which was composed of individual decision-making (ISDM) stages and TCDM stages. The effects of the different stages, sex and incentive levels were assessed.

Therefore, for the present study, we had three following hypothesis:

- H1: The context of decision patterns (individual and team decision stages) may providing some evidence for intra- and inter-brain neural substrate differences between ISDM and TCDM behavior.
- H2: Previous researches have confirmed the impact of social environmental cues (e.g., incentive levels) (Vlaev and Chater, 2006), which may also modulated the neural substrate differences between ISDM and TCDM behavior.
- H3: Meanwhile, male and female participants may exhibit different neural patterns during ISDM and TCDM processes.

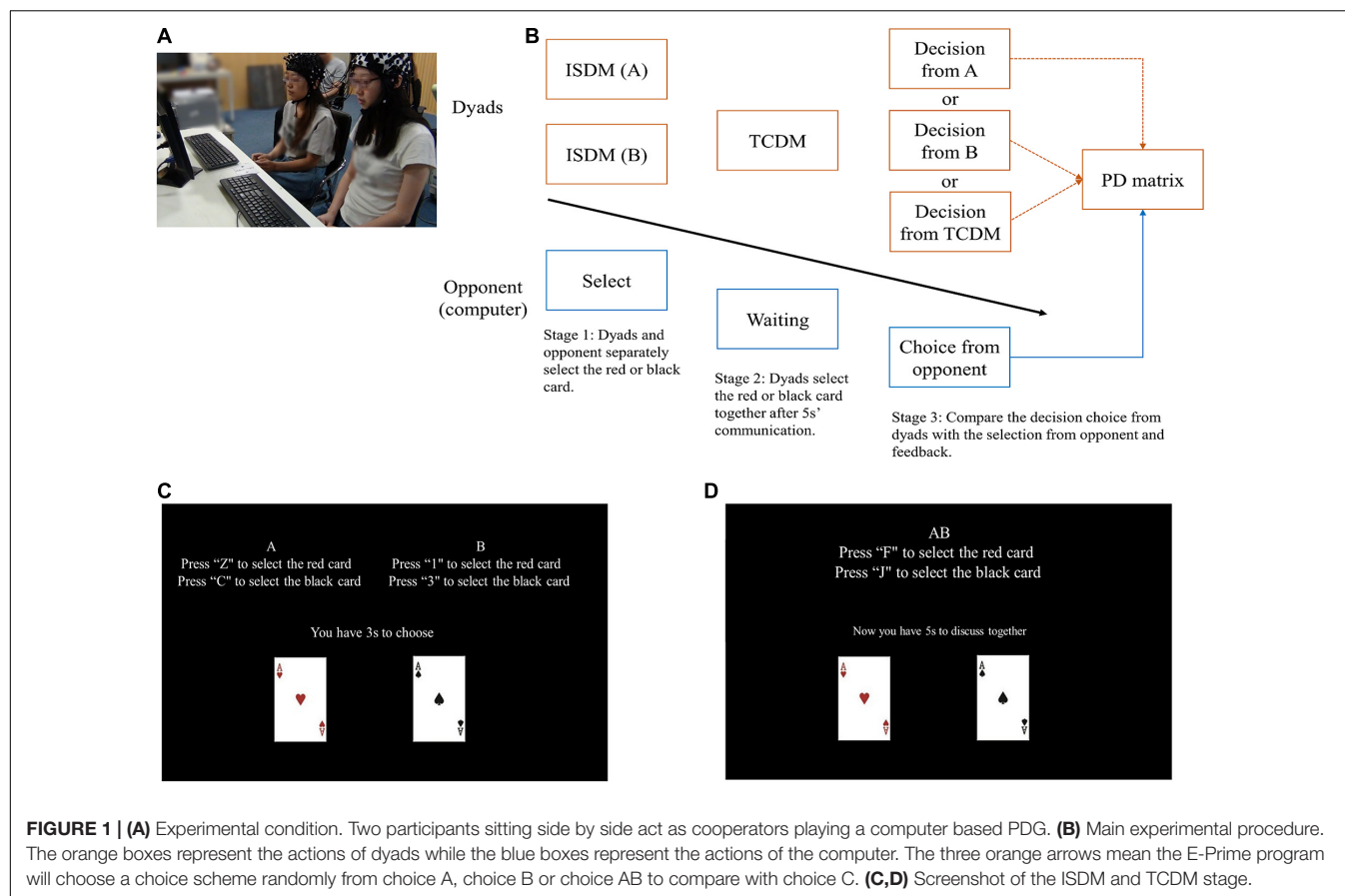
MATERIALS AND METHODS

Participants

We recruited sixty-two healthy, right-handed students from several universities (32 males and 30 females, mean age = 22.3 ± 2.4 years). All subjects were strangers and participated in same-sex pairs (total of thirty-one pairs). Informed written consent was collected from all participants. All experiments were approved by the Southeast University Institutional Review Board.

Experimental Procedure

Two participants (P1 and P2), sitting side by side, act as cooperators playing with a computer. Dyads were told that they need to play with another player in the next room, they do not know their opponent is a computer. We added a cooperative decision-making stage to the classical PDG (**Figure 1**). The prisoner's dilemma is a classic game paradigm that studies the conflict between individual interests and groups interests. In the classical two-person PDG, two participants were asked to cooperate or defect with each other, so as to win as many rewards as possible. The reward were classified in to four types, i.e., the reward outcome (R), temptation outcome (T), the sucker outcome (S) and the punishment outcome (P). In each round of the game, if two opponents choose to cooperate with each other, they both receive the reward outcome (R); If one opponent chooses to cooperate and the other chooses to defect, the opponent who defected receives the temptation outcome (T), while the other receives the sucker outcome (S); If two participants choose to defect with each other, they both receive the punishment outcome (P) (Rapoport, 1967). In the present study, Dyad members were asked to choose a red (choice A) or black card (choice B), separately (Participants were informed that the red and black card imply cooperation and defection, respectively). After that, the dyads have to decide together and form a choice scheme for TCDM (choice AB) without language communicate, before all the experiments, we told the dyads to communicate by using hand gestures which were discussed and established during the practical trials. Meanwhile, dyads were told that their opponent (the computer) also made a decision at the same time (choice C). After the two decision-making stages, the E-Prime program will choose a choice scheme randomly from choice A, choice B or choice AB to compare with choice C. It



should be noted that the decision of computer is controlled by the E-Prime program, which adopt the tit-for-tat strategy. i.e., always makes the same choice with the participants (the scheme executed in the judging stage in the present study) in the previous trial (Sheldon, 1999).

In order to investigate the effect of social environmental cues, we set up two types of reward outcomes (R). The R for mutual cooperation is 3 yuan (low incentive reward; LIR) or 7 yuan (high incentive reward; HIR), and were performed in a random order. Participants were told that their winnings in each round would be proportionately convert to remuneration at the end of the experiment. e.g., when the uncover result was cooperation (the choice scheme selected and the choice C were both red), P1 and P2 get 3 yuan (China's currency) together (1.5 yuan each), If the dyad got 280 yuan in total after all the experiments, we will give them 140 (The total reward/2) yuan as the remuneration (70 yuan each, equal to 10.85 USD). The entire experiment consisted of 40 rounds (20 rounds of randomized LIR/HIR tasks), each one lasting 40–50 s. All experimental procedures were implemented using E-prime 2.0 (Psychology Software Tools Inc., Pittsburgh, PA, United States). The reward outcome (R), temptation outcome (T), sucker outcome (S) and the punishment outcome (P) of the improved PDG are reported in **Table 1**.

Meanwhile, we prepared a text version of experimental guidance and a video of experimental instructions to help participants understand the experimental process. After the

TABLE 1 | The PD matrix in the experiment.

Dyads	P3	Cooperate (red)		Defect (black)	
Cooperate (red)		3/7	3/7	10	0
Defect (black)		0	10	0	0

practice experiment, in order to establish a baseline for the data analysis, the dyads were asked to rest for 60-s, during which they were required to relax mind and keep motionless as much as possible (Jiang et al., 2015). Formal experiment will start until all participants were familiar with the rules of the game.

By focusing on the two stage of decision actions, the improved PDG in the present study provided an opportunity to assess the behavioral and neural difference of ISDM and TCDM behavior in human-to-human interaction context. Moreover, the design of experiment setting of HIR and LIR, and the team composition of same sex allow us to evaluate the influence of social environmental cues and sex effect.

Apparatus

We used a 16 emitters -16 detectors fNIRS system (LABNIRS; Shimadzu Co., Japan) operated at 780, 805, and 830 nm wavelengths, which could simultaneously measure relative changes in concentrations of oxygenated hemoglobin (Oxy-Hb), deoxygenated hemoglobin (Deoxy-Hb), and total hemoglobin.

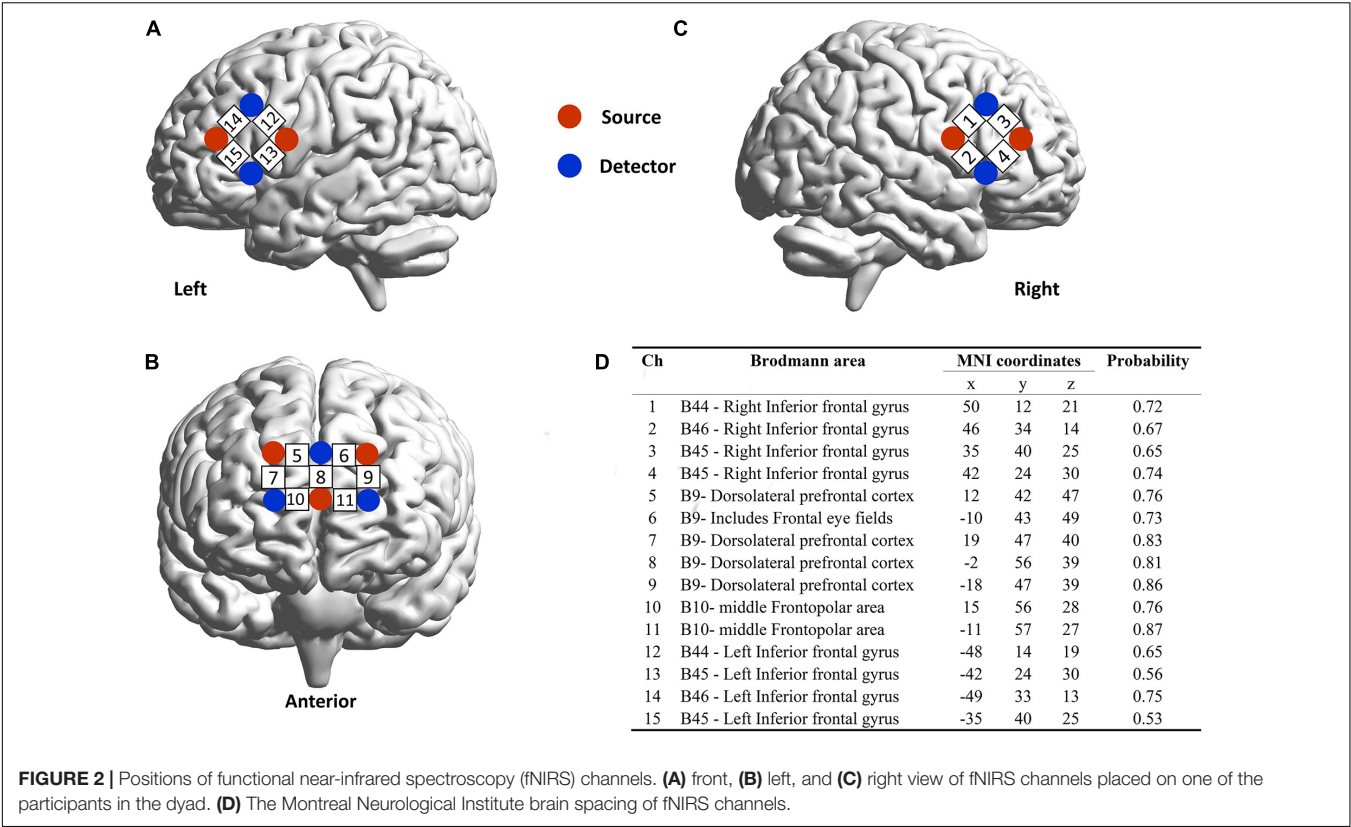


FIGURE 2 | Positions of functional near-infrared spectroscopy (fNIRS) channels. **(A)** front, **(B)** left, and **(C)** right view of fNIRS channels placed on one of the participants in the dyad. **(D)** The Montreal Neurological Institute brain spacing of fNIRS channels.

The relative brain activation was assessed via the conversion of light intensity signals using the modified Beer-Lambert law (Cope et al., 1988). For each participant, one “3 × 3” (forming 12 measurement channels) and two “2 × 2” (forming 8 measurement channels) measurement patches were placed on the prefrontal cortex (PFC) and bilateral IFG, respectively. The center point of each emitter-detector pair (located 3 cm apart) was defined as the measurement channel (we provide a video of the precise positions of the fNIRS channels in **Supplementary Materials**). The precise positions of the fNIRS channels was measured by a 3D electromagnetic tracking device (FASTRAK; Polhemus, United States) and registered on the Montreal neurological Institute (MNI) brain space using a virtual registration method (**Figure 2**) (Zhang et al., 2021). The sampling rate is 42 Hz.

Data Analysis

We used the data preprocessing program in LABNIRS system to remove longitudinal signal drift, motion artifacts, and physiological noise from the raw data, a 0.01–0.1 Hz band-pass filter was applied. Then, a linear baseline correction was performed on the filtered data using the mean value of the Oxy-Hb during the last 10 s of the pre-task period. The fNIRS data were further divided into three stages, ISDM, TCDM, and feedback. Here, we mainly focused on two decision stages. Actions by subject dyads were classified as cooperation or defection, based upon the card chosen. Note that we only focused on the mutual cooperation and defection decisions; in

the ISDM stage, the dyads may choose differently, e.g., one chose red while the other chose black. This situation was not considered in this study. Finally, the average group data were calculated from all these categories. The significant level was set at $p < 0.05$. False discovery rate correction was used to minimize false positive results. We primarily focused on the oxy-Hb data, since the oxygenated signal was more sensitive to changes in cerebral blood flow (Hoshi, 2003; Lindenberger et al., 2009). The wavelet coherence between the activations of each dyad was evaluated to examine the inter-brain coupling in each stage (Grinsted et al., 2004).

RESULTS

Behavioral Data

In order to examine the effects of sex, task-type and decision-type on the reaction times (RTs) and reaction choices, three-factor repeated measures ANOVA [sex (male vs. female) × decision type (cooperation vs. defection) × task type (low incentive task vs. high incentive task)] were conducted on the behavioral data (i.e., RT and reaction choice).

For the reaction choices, there was a significant interaction between decision-type and task-type during the ISDM stage [$F(1, 60) = 7.514, p = 0.003, \eta_p^2 = 0.326$]. Simple effect analysis revealed that participants formed more cooperative decisions than competitive decisions under the HIR task. There was an interaction between sex and decision-type [$F(1,$

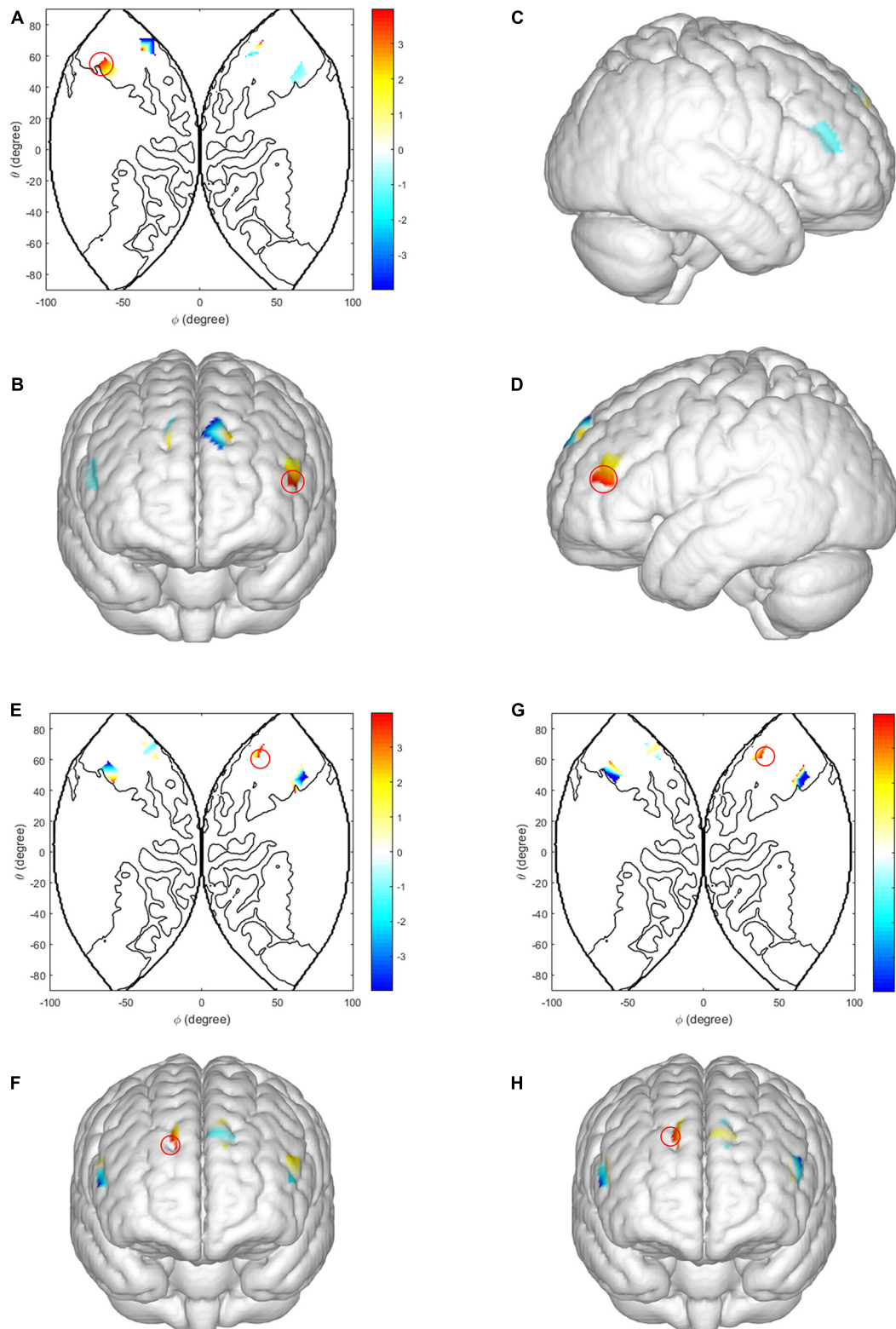


FIGURE 3 | Intra-brain activation in ISDM stage. **(A–D)** Higher activations in the left IFG (LIR condition vs. HIR condition). **(A)** The 2D model of activation map of the left IFG; **(B)** front, **(C)** right, and **(D)** left view of a 3D model of activation map of the left IFG. **(E,F)** Higher activations in the DLPFC and the middle frontopolar area of males (LIR condition vs. HIR condition, when male dyads made defection decision). **(G,H)** Higher activations in the DLPFC and the middle frontopolar area of males (cooperation decision vs. defection decision under HIR condition). Color bar indicated the paired-samples *T*-test value, the red circle represents the statistical significance areas, similarly hereinafter.

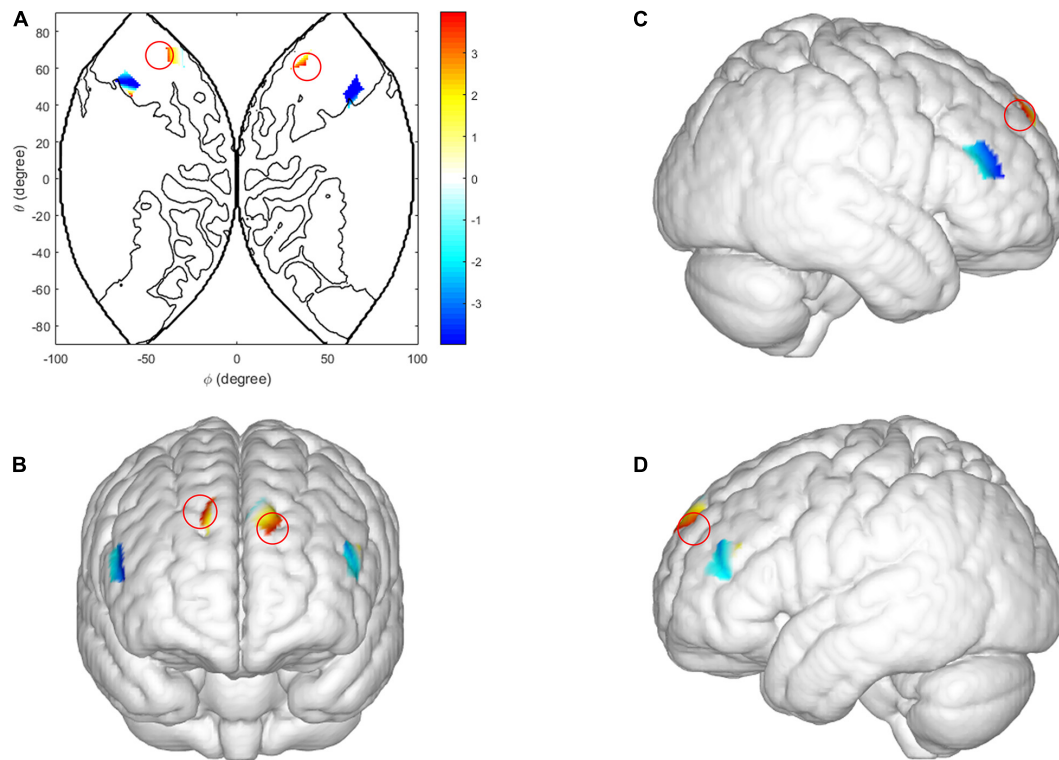


FIGURE 4 | Intra-brain activation in the TCDM stage. **(A–D)** Higher activations in the DLPFC and the middle frontopolar area (LIR condition vs. HIR condition). **(A)** The 2D model of activation map of the DLPFC and the middle frontopolar area; **(B)** front, **(C)** right, and **(D)** left view of a 3D model of activation map of the DLPFC and the middle frontopolar area.

60) = 4.326, $p = 0.012$, $\eta_p^2 = 0.245$]. Simple effect analysis revealed that male participants formed more defection decisions than cooperation decisions. There were no significant main effects or other interaction effects ($P > 0.05$). There was no significant difference in RTs.

In the TCDM stage, for the reaction choices, there was a significant main effect of the decision-type [$F(1, 29) = 9.675$, $p < 0.001$, $\eta_p^2 = 0.341$]. Participants formed more cooperation decision than competition decision. For the RTs, decision-type had a significant main effect [$F(1, 29) = 7.237$, $p = 0.004$, $\eta_p^2 = 0.306$]. The average RTs for cooperation decision is shorter than the defection decision. There were no significant main effects of sex, task type or other interaction effects ($P > 0.05$).

Near-Infrared Spectroscopy Data

Intra-Brain Activation Differences

We conducted three-factor, repeated measures ANOVA [sex (male vs. female) \times decision-type (cooperation vs. competition) \times task-type (low incentive task vs. high incentive task)] on the activation values from all dyads.

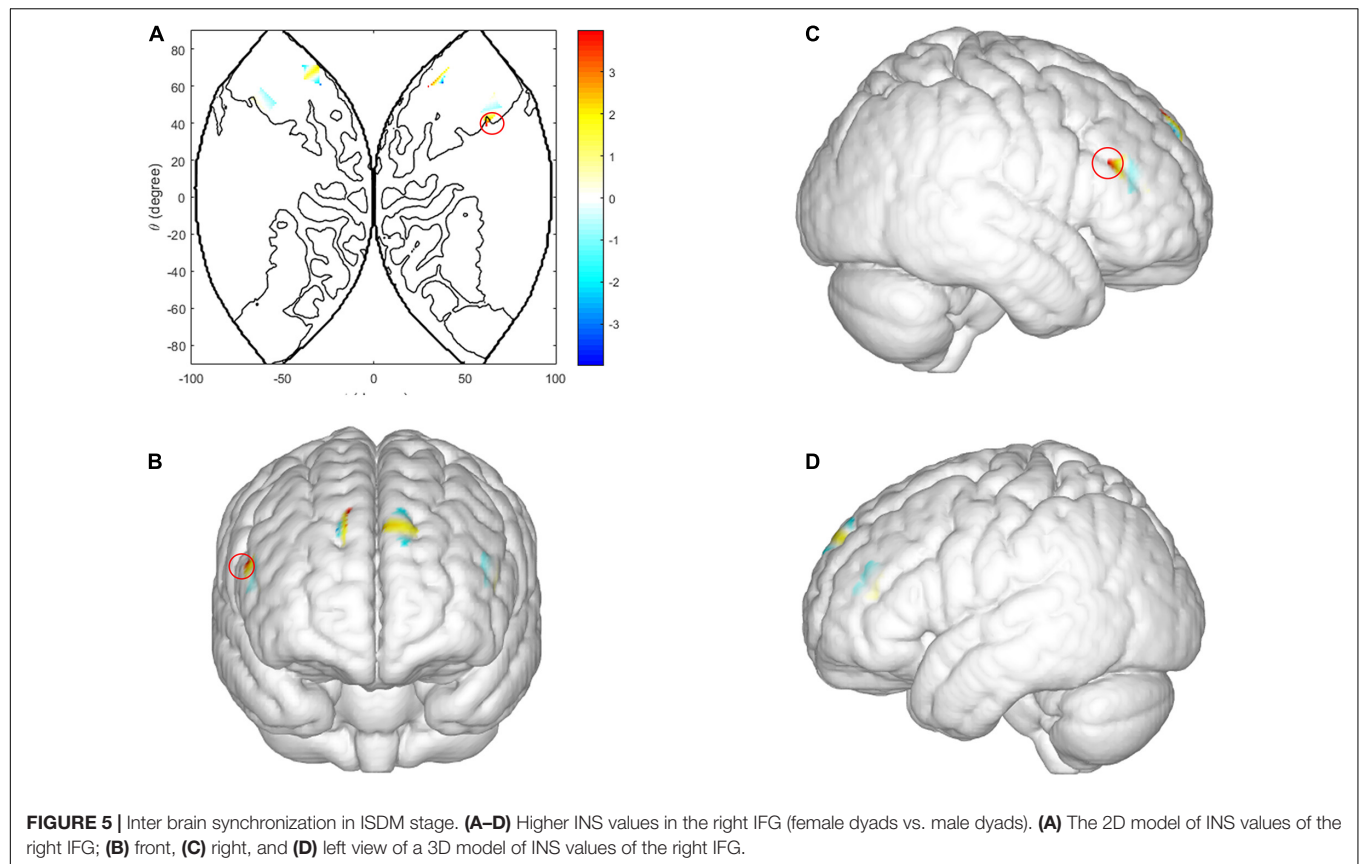
In the ISDM stage, we analyzed the intra-brain activation between two subjects in one dyad. We observed a main effect of task-type in the left IFG [CH 14: $F(1, 60) = 8.241$, $p = 0.010$, $\eta_p^2 = 0.314$; CH 15: $F(1, 60) = 7.282$, $p = 0.015$, $\eta_p^2 = 0.288$, FDR corrected]. The LIR condition evoked higher activations than the HIR condition. There were also significant interaction

effects between decision-type, task-type and sex in the DLPFC [CH 7: $F(1, 60) = 6.132$, $p = 0.01$, $\eta_p^2 = 0.419$, FDR corrected] and the middle frontopolar area [CH 10: $F(1, 60) = 6.709$, $p = 0.007$, $\eta_p^2 = 0.441$, FDR corrected]. Simple effect analysis revealed that under the HIR task, there was a significantly higher activation in cooperation decision than defection decision, in male dyads [DLPFC (CH 7): $p = 0.015$; middle frontopolar area (CH 10): $p = 0.009$]; when male dyads made defection decisions, there was a significantly higher activation in LIR condition than HIR condition [DLPFC (CH 7): $p = 0.017$; middle frontopolar area (CH 10): $p = 0.011$] (see Figure 3).

In the TCDM stage, there was a main effect of the task-type in the DLPFC [CH 5: $F(1, 29) = 7.889$, $p = 0.009$, $\eta_p^2 = 0.220$; CH 7: $F(1, 29) = 11.105$, $p = 0.002$, $\eta_p^2 = 0.284$; CH 9: $F(1, 29) = 8.088$, $p = 0.008$, $\eta_p^2 = 0.224$, FDR corrected] and the middle frontopolar area [CH 11: $F(1, 29) = 9.049$, $p = 0.006$, $\eta_p^2 = 0.244$, FDR corrected]. Activation of the LIR task was significantly higher than in the HIR task, in the DLPFC and the middle frontopolar area (see Figure 4).

Inter-Brain Synchrony

The length of defined interaction behaviors (decision-making) was a minimum of 20 s, and the length of one trial was about 50 s, so we identified a frequency band between 0.02 and 0.05 Hz, where the task occurred, and removed high- and low-frequency noise.



Three factor repeated measures ANOVA [sex (male vs. female) \times behavior-type (cooperation vs. competition) \times task-type (low incentive task vs. high incentive task)] was conducted on the coherence values of all scalp channels, from all dyads. In the ISDM stages there was a significant main effect of sex in the right IFG [CH1: $F(1, 29) = 7.876$, $p = 0.014$, $\eta_p^2 = 0.353$]. Female dyads yielded a significant coherence increase compared with male dyads (see **Figure 5**).

In the TCDM stages, there was a significant main effect of task-type in the right IFG [CH 2: $F(1, 29) = 9.985$, $p = 0.004$, $\eta_p^2 = 0.263$]. The coherence value under the HIR task is significantly higher than the LIR task (see **Figure 6**).

DISCUSSION

The experiment performed herein indicate that: (1) Cooperation is a common strategy adopted by the participants, indicating that they tend to cooperate with their opponents (computers in this experiment) in order to obtain the cooperative reward outcome. (2) The experimental setting of HIR promotes a cooperative outcome. Specifically, participants formed more cooperative decisions than competitive decisions under the HIR task. (3) Males were inclined to form more defection decisions than cooperation decisions. This is consistent with the results of previous research on cooperative behavior in social dilemmas, i.e., in team collaborative behavior, males were more willing

to defect their opponent, while females preferred cooperation (Balliet et al., 2011).

Regarding the neural substrates activated during cooperative decision-making, the present study found significant results in intra-brain activation. I.e., The LIR condition yielded higher intra-brain activation than the HIR condition (i.e., ISDM stage: in the left IFG, TCDM stage: In the DLPFC and middle frontopolar area). Regarding the activation of cortices between TCDM and ISDM, previous research has shown that the left IFG was activated in inhibitory processes, including the tendency to inhibit learning from undesirable information. For example, TMS to the left IFG has activity has been shown to release such inhibition, increasing the ability to learn from undesirable information (Acheson and Hagoort, 2013). We presume that in the ISDM stage, during the LIR task, the defection decision may obtain a higher reward and as a result, the individual inhibited their tendency for cooperation in order to satisfy the temptation outcome (T), which leads to higher activation in the left IFG. Meanwhile, Sanfey et al. (2003) found that in an economic game, the DLPFC is primarily responsible for suppressing the negative emotions activated by the insula, thus obtaining a rational economic outcome (Sanfey et al., 2003). The middle frontopolar area plays a central role in higher cognitive functions such as planning, problem solving, reasoning, and episodic memory retrieval (Braver and Bongiolatti, 2002). Due to the influence of the partner in the dyad, selfish behavior being inhibited and the tendency to cooperate increases, which caused the activation of the DLPFC and middle

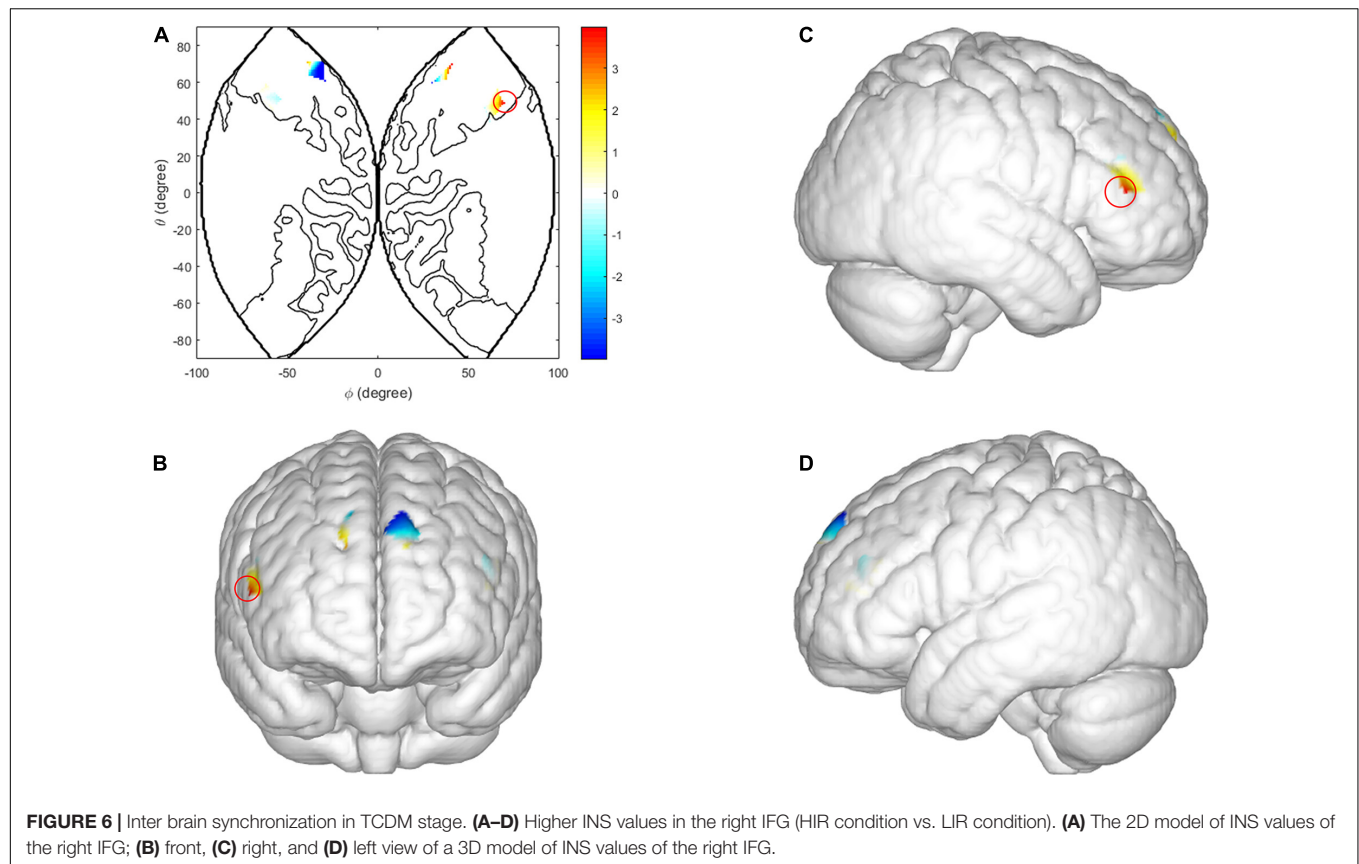


FIGURE 6 | Inter brain synchronization in TCDM stage. **(A–D)** Higher INS values in the right IFG (HIR condition vs. LIR condition). **(A)** The 2D model of INS values of the right IFG; **(B)** front, **(C)** right, and **(D)** left view of a 3D model of INS values of the right IFG.

frontopolar area during the TCDM stage, as opposed to the ISDM stage. Potentially explaining the activation differences caused by social environmental cues during the ISDM and TCDM stages. This activation diversity reveals the different intra-brain neural substrates between ISDM and TCDM.

With respect to the inter-brain neural substrate, the higher IBS during the TCDM (in the right IFG), was not located in the same cortex where we found in the intra-brain activations (in the DLPFC and middle frontopolar area). Previous studies have confirmed that the regions of intra-brain activation and IBS do not coincide (Baker et al., 2016). On the other hand, this discrepancy might be due to the neural substrates of TCDM, in the social interaction contexts. Meanwhile, the IFG, especially the superior inferior frontal, is a core cortex area of the mirror neurons (MNS), which is activated in imitation, language comprehension and interpretation of intentions (Iacoboni et al., 1999; Tettamanti et al., 2005; Iacoboni, 2008), and plays a key role in reward expectation (Sip et al., 2012). Overall, these behavioral results indicate that the HIR rewards involve common goals and fewer self-other differences, suggesting that it is relatively simple to understand the dyad members actions and intentions. Moreover, this difference in IBS values modulated by social environmental cues, only existed during the TCDM stage, which reveals the differences of inter-brain neural mechanisms between TCDM and ISDM.

The results of intra-brain activation and IBS also found significant sex differences, which was modulated by social

environmental cues. Overall, males evoked different intra-brain activations in the DLPFC and the middle frontopolar area while females yielded different IBS in the right IFG, during the ISDM stage. Regarding the sex effect in prisoner dilemmas, accumulated studies have indicated that males maintain a better balance between revenue expectation and trust behavior, which allows them to strategically deal the social dilemma situations (Balliet et al., 2011), which may lead to the higher activation of the cooperation decision during the HIR task. Other studies have found that male tend to be individualized, while females show more altruistic behavior because they are naturally more compassionate (Heilman and Chen, 2005). As mentioned earlier, defection decisions may earn higher rewards and this temptation is stronger in LIR conditions. Therefore, males expect to maximize their reward in the LIR task, which results in the higher activation when males made a defection decision (compared with the HIR task). Meanwhile, Declerck et al. (2013) suggested that the motivation for cooperation was modulated by cognitive control systems (located in the prefrontal cortex, i.e., the DLPFC) and social-brain networks (i.e., the TPJ, IFG, mPFC, and the amygdala) (Declerck et al., 2013). They assumed that pro-selfs are more likely to adopt economic rational strategies and rely on cognitive control systems to make decisions, while pro-social individuals are more likely to adopt social rational strategies and dependent on social-brain networks (Balliet et al., 2011). As such, it is not difficult to explain the sex effect between intra- and inter-brain frames found in this study. That said,

influenced by different levels of social environmental cues, males tended to form strategic decisions, while females were inclined to establish cooperation and interpersonal promotion, which leads to the significant results that decision-making behavior in males, showed activation differences in the cognitive control system (DLPFC and middle frontopolar area), while a female's decision-making behavior showed INS increased in the social brain networks (IFG). This was consistent with the hypothesis of Declerck (Declerck et al., 2013).

Finally, it should also be noted that, despite the increasing popularity of hyperscanning tools in social neuroscience research, we should always be aware of limitations and further questions for future research. First, the sample size was relatively small. In spite of this, the intra- and the inter-brain analysis revealed significant differences between individual and team decision-making behaviors. Second, the dyads in the present study were concentrated among the same sex, further study is thus needed to confirm the sex effect in mixed sex dyads. Third, the ecological validity of the present study still cannot completely generalize to a wider range of team decision-making behaviors, but even so, this study provide a compelling basis and interesting perspective for future neuroscience work of TCDM behaviors.

CONCLUSION

That said, the present study draws three conclusions. First, ISDM and TCDM behavior yielded different intra- brain activation and inter- brain synchronization patterns in the neural network of cognitive control system and MNS, this indicated the different neural substrate of ISDM and TCDM behavior. Second, social environmental cues modulated different intra brain activation and inter brain synchronization of ISDM and TCDM behavior. These results confirmed the modulation effect of social environmental cues, also conclude new impact model in the human-human interaction frame. Third, the INS results suggest that in ISDM situation, males and females may primarily depend on non-social and social cognitive ability to make decision separately, while in the social interaction situation of TCDM, females may use both social and non-social cognitive abilities.

REFERENCES

- Acheson, D. J., and Hagoort, P. (2013). Stimulating the Brain's language network: syntactic ambiguity resolution after tms to the inferior frontal gyrus and middle temporal gyrus. *J. Cogn. Neurosci.* 25, 1664–1677. doi: 10.1162/jocn_a_00430
- Astolfi, L., Toppi, J., Fallani, F. D. V., Vecchiato, G., Cincotti, F., Wilke, C. T., et al. (2011). Imaging the social brain by simultaneous hyperscanning during subject interaction. *IEEE Intell. Syst.* 26, 38–45. doi: 10.1109/mis.2011.61
- Awino, Z. B. (2013). Top Management team diversity, quality decisions and organizational performance in the service industry. *J. Manag. Strategy* 4, 113–123.
- Bahrami, B., Olsen, K., Latham, P. E., Roepstorff, A., Rees, G., and Frith, C. D. (2010). Optimally interacting minds. *Science* 329, 1081–1085. doi: 10.1126/science.1185718
- Baker, J. M., Liu, N., Cui, X., Vrticka, P., Saggar, M., Hosseini, S. M., et al. (2016). Sex differences in neural and behavioral signatures of cooperation revealed by fNIRS hyperscanning. *Sci. Rep.* 6:30512.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Southeast University Institutional Review Board. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

MZ and HJ conceived the project, performed the study, and analyzed the data. MZ designed the study. MZ, HJ, and GW wrote and reviewed the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The work was supported by the China Postdoctoral Science Foundation (2020M671175), the Foundation of Shanghai Normal university (310-AC7031-20-003028), and the Shanghai Science and Technology Commission (19QA1405800 and 19411968800).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnhum.2021.702959/full#supplementary-material>

- Balliet, D., Li, N. P., Macfarlan, S. J., and Van Vugt, M. (2011). Sex differences in cooperation: a meta-analytic review of social dilemmas. *Psychol. Bull.* 137, 881–909. doi: 10.1037/a0025354
- Braver, T. S., and Bongiolatti, S. R. (2002). The role of frontopolar cortex in subgoal processing during working memory. *Neuroimage* 15, 523–536. doi: 10.1006/nimg.2001.1019
- Carton, A. M., and Cummings, J. N. (2012). A theory of subgroups in work teams. *Acad. Manag. Rev.* 37, 441–470. doi: 10.5465/amr.2009.0322
- Carton, A. M., and Cummings, J. N. (2013). The impact of subgroup type and subgroup configurational properties on work team performance. *J. Appl. Psychol.* 98, 732–758. doi: 10.1037/a0033593
- Cope, M., Delpy, D. T., Reynolds, E. O. R., Wray, S., Wyatt, J., and van der Zee, P. (1988). "Methods of quantitating cerebral near infrared spectroscopy data," in *Oxygen Transport to Tissue X. Advances in Experimental Medicine and Biology*, Vol. 222, eds M. Mochizuki, C. R. Honig, T. Koyama, T. K. Goldstick, and D. F. Bruley (New York, NY: Springer).

- Cui, X., Bryant, D. M., and Reiss, A. L. (2012). NIRS-based hyperscanning reveals increased interpersonal coherence in superior frontal cortex during cooperation. *Neuroimage* 59, 2430–2437. doi: 10.1016/j.neuroimage.2011.09.003
- Dawes, R. M., and Messick, D. M. (2000). Social dilemmas. *Int. J. Psychol.* 35, 111–116.
- Decety, J., Jackson, P. L., Sommerville, J. A., Chaminade, T., and Meltzoff, A. N. (2004). The neural bases of cooperation and competition: an fMRI investigation. *Neuroimage* 23, 744–751. doi: 10.1016/j.neuroimage.2004.05.025
- Declerck, C. H., Boone, C., and Emonds, G. (2013). When do people cooperate? The neuroeconomics of prosocial decision making. *Brain Cogn.* 81, 95–117. doi: 10.1016/j.bandc.2012.09.009
- Dommer, L., Jäger, N., Scholkmann, F., Wolf, M., and Holper, L. (2012). Between-brain coherence during joint n-back task performance: a two-person functional near-infrared spectroscopy study. *Behav. Brain Res.* 234, 212–222. doi: 10.1016/j.bbr.2012.06.024
- Funane, T., Kiguchi, M., Atsumori, H., Sato, H., Kubota, K., and Koizumi, H. (2011). Synchronous activity of two people's prefrontal cortices during a cooperative task measured by simultaneous near-infrared spectroscopy. *J. Biomed. Opt.* 16:077011. doi: 10.1117/1.3602853
- Grinsted, A., Moore, J. C., and Jevrejeva, S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Process Geophys.* 11, 561–566. doi: 10.5194/npg-11-561-2004
- Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., and Keysers, C. (2012). Brain-to-brain coupling: a mechanism for creating and sharing a social world. *Trends Cogn. Sci.* 16, 114–121. doi: 10.1016/j.tics.2011.12.007
- Heilman, M. E., and Chen, J. J. (2005). Same behavior, different consequences: reactions to men's and women's altruistic citizenship behavior. *J. Appl. Psychol.* 90, 431–441. doi: 10.1037/0021-9010.90.3.431
- Hoshi, Y. (2003). Functional near-infrared optical imaging: utility and limitations in human brain mapping. *Psychophysiology* 40, 511–520. doi: 10.1111/1469-8986.00053
- Iacoboni, M. (2008). Imitation, empathy, and mirror neurons. *Annu. Rev. Psychol.* 60, 653–670. doi: 10.1146/annurev.psych.60.110707.163604
- Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazzotta, J. C., and Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science* 286, 2526–2528. doi: 10.1126/science.286.5449.2526
- Jiang, J., Chen, C., Dai, B., Shi, G., Ding, G., Liu, L., et al. (2015). Leader emergence through interpersonal neural synchronization. *Proc. Nat. Acad. U. S. A.* 112, 4274–4279. doi: 10.1073/pnas.1422930112
- Lindenberger, U., Li, S. C., Gruber, W., and Muller, V. (2009). Brains swinging in concert: cortical phase synchronization while playing guitar. *BMC Neurosci.* 10:22.
- Liu, T., and Pelowski, M. (2014). A new research trend in social neuroscience: towards an interactive-brain neuroscience. *Psych. J.* 3, 177–188. doi: 10.1002/pchj.56
- Owens, B. P., Wallace, A. S., and Waldman, D. A. (2015). Leader narcissism and follower outcomes: the counterbalancing effect of leader humility. *J. Appl. Psychol.* 100, 1203–1213. doi: 10.1037/a0038698
- Rapoport, A. (1967). Optimal policies for the Prisoner's dilemma. *Psychol. Rev.* 74, 136–148. doi: 10.1037/h0024282
- Rilling, J., Gutman, D., Zeh, T., Pagnoni, G., Berns, G., and Kilts, C. (2002). A neural basis for social cooperation. *Neuron* 35:395. doi: 10.1016/s0896-6273(02)00755-9
- Rilling, J. K., Goldsmith, D. R., Glenn, A. L., Jairam, M. R., Elfenbein, H. A., Dagenais, J. E., et al. (2008). The neural correlates of the affective response to unreciprocated cooperation. *Neuropsychologia* 46, 1256–1266. doi: 10.1016/j.neuropsychologia.2007.11.033
- Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., and Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science* 300, 1755–1758. doi: 10.1126/science.1082976
- Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., et al. (2013). Toward a second-person neuroscience. *Behav. Brain Sci.* 36, 393–414. doi: 10.1017/s0140525x12000660
- Sheldon, K. M. (1999). Learning the lessons of tit-for-tat: even competitors can get the message. *J. Pers. Soc. Psychol.* 77, 1245–1253. doi: 10.1037/0022-3514.77.6.1245
- Sip, K. E., Skewes, J. C., Marchant, J. L., McGregor, W. B., Roepstorff, A., and Frith, C. D. (2012). What if I Get Busted? deception, choice, and decision-making in social interaction. *Front. Neurosci.* 6:58.
- Tang, H., Mai, X., Wang, S., Zhu, C., Krueger, F., and Liu, C. (2016). Interpersonal brain synchronization in the right temporo-parietal junction during face-to-face economic exchange. *Soc. Cogn. Affect. Neurosci.* 11:23. doi: 10.1093/scan/nsv092
- Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., et al. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *J. Cogn. Neurosci.* 17, 273–281. doi: 10.1162/0898929053124965
- Vlaev, I., and Chater, N. (2006). Game relativity : how context influences strategic decision making. *J. Exp. Psychol.* 32, 131–149. doi: 10.1037/0278-7393.32.1.131
- Zhang, M., Jia, H., Zheng, M., and Liu, T. (2021). Group decision-making behavior in social dilemmas: inter-brain synchrony and the predictive role of personality traits. *Pers. Ind. Diff.* 168:110315. doi: 10.1016/j.paid.2020.110315
- Zhou, W., Zhu, Z., and Vredenburg, D. (2020). Emotional intelligence, psychological safety, and team decision making. *Team Perform. Manag. Int. J.* 26, 123–141. doi: 10.1108/tpm-10-2019-0105

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Zhang, Jia and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Shared and Unshared Feature Extraction in Major Depression During Music Listening Using Constrained Tensor Factorization

Xiulin Wang^{1,2}, Wenya Liu^{2,3}, Xiaoyu Wang², Zhen Mu⁴, Jing Xu⁵, Yi Chang⁵, Qing Zhang¹, Jianlin Wu^{1*} and Fengyu Cong^{2,3,6,7*}

¹ Department of Radiology, Affiliated Zhongshan Hospital of Dalian University, Dalian, China, ² School of Biomedical Engineering, Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, China, ³ Faculty of Information Technology, University of Jyväskylä, Jyväskylä, Finland, ⁴ Department of Psychology, College of Humanities and Social Sciences, Dalian Medical University, Dalian, China, ⁵ Department of Neurology and Psychiatry, First Affiliated Hospital, Dalian Medical University, Dalian, China, ⁶ School of Artificial Intelligence, Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, China, ⁷ Key Laboratory of Integrated Circuit and Biomedical Electronic System, Dalian University of Technology, Dalian, China

OPEN ACCESS

Edited by:

Zhishan Hu,
Beijing Normal University, China

Reviewed by:

Xiaoli Li,
Beijing Normal University, China
Meng-Yun Wang,
University of Bergen, Norway

*Correspondence:

Jianlin Wu
cjr.wujianlin@vip.163.com
Fengyu Cong
cong@dlut.edu.cn

Specialty section:

This article was submitted to
Brain Imaging and Stimulation,
a section of the journal
Frontiers in Human Neuroscience

Received: 21 October 2021

Accepted: 22 November 2021

Published: 15 December 2021

Citation:

Wang X, Liu W, Wang X, Mu Z, Xu J, Chang Y, Zhang Q, Wu J and Cong F (2021) Shared and Unshared Feature Extraction in Major Depression During Music Listening Using Constrained Tensor Factorization.
Front. Hum. Neurosci. 15:799288.
doi: 10.3389/fnhum.2021.799288

Ongoing electroencephalography (EEG) signals are recorded as a mixture of stimulus-elicited EEG, spontaneous EEG and noises, which poses a huge challenge to current data analyzing techniques, especially when different groups of participants are expected to have common or highly correlated brain activities and some individual dynamics. In this study, we proposed a data-driven shared and unshared feature extraction framework based on nonnegative and coupled tensor factorization, which aims to conduct group-level analysis for the EEG signals from major depression disorder (MDD) patients and healthy controls (HC) when freely listening to music. Constrained tensor factorization not only preserves the multilinear structure of the data, but also considers the common and individual components between the data. The proposed framework, combined with music information retrieval, correlation analysis, and hierarchical clustering, facilitated the simultaneous extraction of shared and unshared spatio-temporal-spectral feature patterns between/in MDD and HC groups. Finally, we obtained two shared feature patterns between MDD and HC groups, and obtained totally three individual feature patterns from HC and MDD groups. The results showed that the MDD and HC groups triggered similar brain dynamics when listening to music, but at the same time, MDD patients also brought some changes in brain oscillatory network characteristics along with music perception. These changes may provide some basis for the clinical diagnosis and the treatment of MDD patients.

Keywords: CANDECOMP/PARAFAC, constrained tensor factorization, EEG, major depressive disorder, naturalistic music stimuli

1. INTRODUCTION

Major depressive disorder (MDD) is a globally prevalent mental disorder with multifactorial causes (Belmaker and Agam, 2008; Gotlib and Joormann, 2010; Jia et al., 2010). Over the past decades, the neural mechanisms of MDD have been widely explored using non-invasive neuroimaging techniques, like functional magnetic resonance imaging (fMRI), electroencephalogram (EEG), and magnetoencephalography (Gotlib and Hamilton, 2008; Kaiser et al., 2015). Most previous studies of MDD are under the conditions of resting states or well-controlled stimuli. In recent years, naturalistic paradigms have been challenging conventional paradigms because they can approximate real-life experiences using naturalistic and continuous stimuli, like music listening and movie watching (Hasson et al., 2004; Sonkusare et al., 2019). Naturalistic paradigms have shown a clinical potential in mental disorders, such as MDD, autism-spectrum disorder, paranoia, and so on (Sonkusare et al., 2019). Music perception can induce emotional arousal for affective processing, and music therapy has shown the feasibility in MDD treatment (Michael et al., 2005; Maratos et al., 2008). However, few studies have correlated the music perceptive arousal with neural mechanisms in MDD, and the studies mainly explored the networks of brain connectivity at the source space (Liu et al., 2020, 2021; Zhu et al., 2021). Furthermore, the current findings are often inconsistent or even contradictory due to the different methodological approaches and involved participants, unified neural mechanisms of MDD (in music perception) can not be concluded (Zhi et al., 2018). Therefore, it is still urgent and important to develop novel experimental designs and advanced computational methodologies to better investigate the neural biomarkers of MDD. In our study, we aim to investigate the biomarkers in MDD during music listening using EEG data at the sensor level.

Due to the high temporal resolution, electroencephalography (EEG) signals contain rich spectral contents. During continuous cognition, the spatial reconfiguration will be dynamically sustained along time, and the spatial signatures are modulated by oscillations (Buzsaki, 2006; Yan et al., 2020; Sadaghiani et al., 2021). Apparently, the EEG signals can be represented by the high-order multi-way array, i.e., tensor, which can fully describe the inherent interaction relationships among multiple dimensions in the data (Kolda and Bader, 2009; Cichocki et al., 2015; Cong et al., 2015). Recently, some studies have investigated the electrophysiological signatures characterized by spatio-temporal-spectral modes of covariation from the tensor representation of EEG data via Canonical Polyadic (CP) decomposition (Mørup et al., 2007; Cong et al., 2012, 2015; Zhu et al., 2020). These studies are based on the assumption of spatial-, spectral-, temporal consistency, which means that each subject or group shares the same frequency-specific brain topography or networks with the same temporal dynamics (Cong et al., 2012, 2015). However, except for common features, individual features should also be considered for subject or group differences (Wang et al., 2020; Liu et al., 2021). Therefore, the incomplete spatial, temporal and spectral consistency should be assumed to better fit the data characteristics and practical applications. Meanwhile,

numerous versions of independent component analysis (ICA)-based methods and their group analysis variants have also been popularly adopted to analyse EEG signals (Cong et al., 2013; Labounek et al., 2018; Zhu et al., 2021). For example, Zhu et al. performed group-level spatial Fourier ICA to explore the frequency-specific brain networks of musical feature processing, and found the alpha lateral component engaged in music perception in MDD (Zhu et al., 2021). These two-way methods simply stack or concatenate the extra modes of EEG signals into two-way matrix for processing, but lose the potential internal relationships among modes and destroy the inherent multi-linear structure of the data (Cong et al., 2015). Considering the high-dimensional structure of the data and the incomplete consistency of different modes, we applied a constrained tensor factorization model by imposing nonnegative and coupled constraints in the present study, by which we can access to the shared and unshared features simultaneously. Liu et al. only considered the coupling structure in spectral and connectivity modes and explored the connectivity alteration in EEG signals during music perception in MDD using tensor decomposition-based methods (Liu et al., 2021). Therefore, different from the previous work (Liu et al., 2020, 2021; Zhu et al., 2021), we further consider the coupling characteristics in the temporal modes between the MDD and healthy controls (HC) data at the sensor-level, i.e., we assume that some of the spatio-temporal-spectral patterns are the same between the two group data while the rest are different.

In this study, for the EEG signals of MDD and HC groups during music listening, we investigated spatio-temporal-spectral modes of covariation using a coupled nonnegative tensor factorization framework, aiming to extract the shared and unshared features between/in the two groups. Specifically, we first recorded the EEG signals during freely listening to a piece of 512-s tango music. Using the time-frequency representation, we then constructed two fourth-order tensors of time, frequency, space and participant for the two groups. Considering the incomplete consistency in spatio-temporal-spectral modes, we applied the triple-coupled nonnegative tensor factorization model optimized by alternating direction method of multipliers (ADMM, Boyd et al., 2011) algorithm, which enables the simultaneous decomposition of shared components and unshared components among tensors. Meanwhile, we extracted five long-term musical features from the musical stimulus using musical information retrieval in order to build the connections with the extracted components from EEG signals. Next, correlation analysis was performed between temporal courses of musical features and the extracted temporal components from EEG signals, and we obtained the spatio-temporal-spectral brain dynamics of interest that were believed to be activated by music modulation. Following this, hierarchical clustering was conducted on the shared and unshared spatial components of the results from multiple runs. Finally, we obtained two clusters of feature patterns shared by MDD and HC groups, as well as one cluster from the HC group and two clusters from the MDD group, which may contribute to the biomarkers for the clinical diagnosis and treatment for MDD patients. The proposed framework based on the constrained tensor factorization is completely data-driven and provides a solution to extract the

TABLE 1 | Basic information of the participants in HC and MDD groups.

	HC group	MDD group	HC vs. MDD
	Mean \pm SD	Mean \pm SD	p-value
Age (years)	38.4 \pm 11.8	42.9 \pm 11.0	> 0.05*
Gender (F:M)	14:5	14:6	> 0.05**
Education (years)	13.6 \pm 3.8	12.8 \pm 3.4	> 0.05*
Duration (months)	-	12.8 \pm 8.5	-
HRSD	2.4 \pm 1.3	23.3 \pm 3.6	< 0.01*
HAMA	2.4 \pm 1.3	19.2 \pm 3.0	< 0.01*
MMSE	28.2 \pm 0.9	28.1 \pm 1.1	> 0.05*

*The p-value is calculated via t-test. **The p-value is calculated via chi-squared test. Duration is the duration of illness.

HC, healthy controls; MDD, major depression disorder patients; F, Female; M, Male; HRSD, Hamilton Rating Scale for Depression; HAMA, Hamilton Anxiety Rating Scale; MMSE, Mini-Mental State Examination.

shared and unshared spatio-temporal-spectral features of EEG signals from different groups.

2. MATERIALS AND METHODS

2.1. Data Description

2.1.1. Participants

In this study, we analyzed the data from 39 participants in total, including nineteen healthy controls (HC) and 20 major depression disorder (MDD) patients. No one was reported to have hearing loss and formal training in music. The mental health of each participant was evaluated and diagnosed by a clinical expert using Hamilton Rating Scale for Depression (HRSD), Hamilton Anxiety Rating Scale (HAMA) and Mini-Mental State Examination (MMSE). The relative values of these indices as well as age, gender, education, duration of illness for HC and MDD groups are listed in **Table 1**. All participants signed the informed consent forms approved by the ethics committee of First Affiliated Hospital of Dalian Medical University and Dalian University of Technology.

2.1.2. EEG data

A 512-second modern tango music “Adios Nonino” played by Astor Piazzolla was adopted as the naturalistic stimulus in this experiment. The participants were told to seat as still as possible with eyes open and listen to the tango music. The ongoing EEG data were recorded using the international 10–20 system-based Neuroscan Quik-cap device of 64 electrodes with the sampling frequency of 1,000 Hz. The recorded EEG data were preprocessed off-line using MATLAB software and EEGLAB toolbox (Delorme and Makeig, 2004), down-sampled to 256 Hz, and filtered by the high-pass and low-pass filters with 4 Hz and 30 Hz cut-off frequencies. The components indicating eye movements artifacts were rejected by independent component analysis (ICA). The data were also visually checked to remove the obvious artifacts brought by head movement, and then used for further analysis.

2.1.3. Musical Features

In this study, five long-term musical features (including two tonal and three rhythmic features) were extracted by a frame-by-frame

analysis method using MIR toolbox (Lartillot and Toivainen, 2007). The duration of each frame was 3 s and the overlap ratio between two frames was 66.7%, which was consistent with the window settings in the time-frequency representation of EEG data. Finally, in order to match the length of recorded EEG data, we selected the first 500 samples for each time course of these musical features with a 1 Hz sampling rate. For the tonal features, Mode denotes the strength of major or minor mode, and Key Clarity is the measure of the tonal clarity. For the rhythmic features, Fluctuation Centroid is defined as the geometric mean of the fluctuation spectrum, and it represents the global repartition of rhythm periodicities within the range of 0~10 Hz, indicating the average frequency of these periodicities. Fluctuation entropy is the Shannon entropy of the fluctuation spectrum, and it represents the global repartition of rhythm periodicities. Pulse Clarity is regarded as an estimate of clarity of the pulse. The details of musical features and extraction method can be found in the previous studies (Alluri et al., 2012; Cong et al., 2013).

2.2. Constrained Tensor Factorization

2.2.1. Tensor Construction

In order to comprehensively analyze the data from more aspects, we first converted the data from the time domain to the time-frequency domain. Specifically, we obtained the time-frequency representation of the preprocessed EEG data via performing the short-time Fourier transform (STFT) on the time series of each channel for each participant. The Hamming window was adopted as the window function, with the window length of 3 s and 66.7% overlap ratio between windows. The number of Fourier points in each window was 1,280, which was five times of the sampling rate. Finally, for the data of each channel, we obtained the spectrograms with the size of 130 (frequency bins) \times 500 (time samples). Therefore, for the HC group and MDD group, the EEG data were reconstructed to two fourth-order tensors with the dimensions of channel (64), frequency (130), time (500) and participants (19 or 20), i.e., $\mathcal{X}_{\text{HC}} \in \mathbb{R}_{+}^{64 \times 130 \times 500 \times 19}$ and $\mathcal{X}_{\text{MDD}} \in \mathbb{R}_{+}^{64 \times 130 \times 500 \times 20}$.

2.2.2. Tensor Factorization

Tensors, also known as multi-way arrays, are the higher-order generalizations of scalars, vectors and matrices. So far, the two most commonly used models for tensor factorization are the canonical polyadic [CP (Hitchcock, 1927), also known as CANDECOMP/PARAFAC (Carroll and Chang, 1970; Harshman, 1970)] model and the Tucker model (Tucker, 1966). The CP model, as a special case of the Tucker model, has better unique identifiability under mild conditions, and was adopted in this study. Thus, the factorizations of the data \mathcal{X}_{HC} and \mathcal{X}_{MDD} can be, respectively expressed, using a sum of fourth-order rank-one tensors or a set of factor matrices, as

$$\mathcal{X}_{\text{HC}} \approx \sum_{r=1}^{R_{\text{HC}}} \mathbf{a}_r^{(1)} \circ \mathbf{a}_r^{(2)} \circ \mathbf{a}_r^{(3)} \circ \mathbf{a}_r^{(4)} = [\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \mathbf{A}_4] \quad (1)$$

and

$$\mathcal{X}_{\text{MDD}} \approx \sum_{r=1}^{R_{\text{MDD}}} \mathbf{b}_r^{(1)} \circ \mathbf{b}_r^{(2)} \circ \mathbf{b}_r^{(3)} \circ \mathbf{b}_r^{(4)} = [\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3, \mathbf{B}_4] \quad (2)$$

where R_{HC} and R_{MDD} are the tensor rank or the number of components that will be extracted from \mathcal{X}_{HC} and \mathcal{X}_{MDD} . “ \circ ” denotes the vector outer product. $\mathbf{A}_i = [\mathbf{a}_1^{(i)} \cdots \mathbf{a}_{R_{\text{HC}}}^{(i)}] \in \mathbb{R}_+^{I_i \times R_{\text{HC}}}$ and $\mathbf{B}_i = [\mathbf{b}_1^{(i)} \cdots \mathbf{b}_{R_{\text{MDD}}}^{(i)}] \in \mathbb{R}_+^{I_i \times R_{\text{MDD}}}$, $i = 1, 2, 3, 4$, correspond to the factor matrices in the spatial, spectral, temporal and participant modes, respectively.

2.2.3. Non-negative Coupled Tensor Factorization

Constrained tensor factorization can accurately extract and explain the hidden components from the input data, by imposing particular penalties/regularizations (e.g., nonnegativity, sparsity, smoothness, coupling) on the corresponding factor matrices in the factorization process. Consider the connections and inherent characteristics of the data \mathcal{X}_{HC} and \mathcal{X}_{MDD} , we imposed the nonnegative constraint in all of modes and the coupling structure in spatial, spectral, temporal modes across the data during the optimization. For the factor matrices \mathbf{A}_i and \mathbf{B}_i , $i = 1 \cdots 4$, we assume each one consists of two parts as $\mathbf{A}_i = [\mathbf{A}_i^C \ \mathbf{A}_i^I]$ and $\mathbf{B}_i = [\mathbf{B}_i^C \ \mathbf{B}_i^I]$, where $\mathbf{A}_i^C = \mathbf{B}_i^C \in \mathbb{R}_+^{I_i \times L_i}$ denotes L_i components are shared by the data, while $\mathbf{A}_i^I \in \mathbb{R}_+^{I_i \times (R_{\text{HC}} - L_i)}$ and $\mathbf{B}_i^I \in \mathbb{R}_+^{I_i \times (R_{\text{MDD}} - L_i)}$ correspond to the individual components in each data. L_i is the number of shared components among data ($L_4 = 0$). Thus, the shared and unshared components can be simultaneously extracted via the formulated nonnegative coupled tensor factorization (NCTF) model, in which the objective function should be minimized as follows:

$$\begin{aligned} \text{minimize } \mathcal{L}(\mathbf{A}_i, \mathbf{B}_i) = & \|\mathcal{X}_{\text{HC}} - [\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \mathbf{A}_4]\|_F^2 + \|\mathcal{X}_{\text{MDD}} \\ & - [\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3, \mathbf{B}_4]\|_F^2 \end{aligned} \quad (3)$$

$$\text{subject to } \mathbf{A}_i \geq 0, \mathbf{B}_i \geq 0 \text{ for } i = 1 \cdots 4$$

where $\|\cdot\|_F$ denotes the Frobenius norm.

2.2.4. Algorithm Optimization

Obviously, the minimization in Equation (3) is not convex but in fact an NP-hard problem. Aiming for an easy-to-handle and robust approximation, we propose to use the ADMM method within the framework of block coordinate descent (BCD) to solve the above optimization problem, which has been proven to be very efficient in the regularized matrix and tensor factorizations (Huang et al., 2016; Schenker et al., 2020). Specifically, BCD framework can obtain a local solution of Equation (3) by converting it into a set of subproblems, in which the factor matrices $\mathbf{A}_1 \& \mathbf{B}_1, \mathbf{A}_2 \& \mathbf{B}_2, \mathbf{A}_3 \& \mathbf{B}_3$ and $\mathbf{A}_4 \& \mathbf{B}_4$ will be updated alternatively one by one in each iteration, then each subproblem can be solved using ADMM strategy. Taking the update of the primal variable pair \mathbf{A}_i and \mathbf{B}_i , $i = 1, \cdots 4$ as an example, the

problem in Equation (3) can be reformulated by introducing the auxiliary variables $\tilde{\mathbf{A}}_i$ and $\tilde{\mathbf{B}}_i$ as follows:

$$\begin{aligned} \text{minimize } \mathcal{L}(\mathbf{A}_i, \tilde{\mathbf{A}}_i, \mathbf{B}_i, \tilde{\mathbf{B}}_i) = & \|\mathcal{X}_{\text{HC}} - [\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \mathbf{A}_4]\|_F^2 \\ & + \|\mathcal{X}_{\text{MDD}} - [\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3, \mathbf{B}_4]\|_F^2 \end{aligned} \quad (4)$$

$$\text{subject to } \mathbf{A}_i = \tilde{\mathbf{A}}_i, \mathbf{B}_i = \tilde{\mathbf{B}}_i, \tilde{\mathbf{A}}_i \geq 0, \tilde{\mathbf{B}}_i \geq 0.$$

The augmented Lagrangian function of Equation (4) is given as:

$$\begin{aligned} \text{minimize } \mathcal{L}(\mathbf{A}_i, \tilde{\mathbf{A}}_i, \mathbf{B}_i, \tilde{\mathbf{B}}_i, \boldsymbol{\Lambda}_i, \boldsymbol{\Gamma}_i) = & \|\mathcal{X}_{\text{HC}} - [\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \mathbf{A}_4]\|_F^2 \\ & + \rho_i \|\mathbf{A}_i - \tilde{\mathbf{A}}_i + \boldsymbol{\Lambda}_i\|_F^2 + \|\mathcal{X}_{\text{MDD}} - [\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3, \mathbf{B}_4]\|_F^2 \\ & + \sigma_i \|\mathbf{B}_i - \tilde{\mathbf{B}}_i + \boldsymbol{\Gamma}_i\|_F^2 \end{aligned} \quad (5)$$

where $\boldsymbol{\Lambda}_i \in \mathbb{R}_+^{I_i \times R_{\text{HC}}}$ and $\boldsymbol{\Gamma}_i \in \mathbb{R}_+^{I_i \times R_{\text{MDD}}}$ are the Lagrangian multipliers or dual variables, ρ_i and σ_i are the penalty parameters. The solutions for Equation (5) can be calculated by successively minimizing \mathcal{L} with respect to $\mathbf{A}_i, \mathbf{B}_i, \tilde{\mathbf{A}}_i, \tilde{\mathbf{B}}_i, \boldsymbol{\Lambda}_i$ and $\boldsymbol{\Gamma}_i$ one at a time while fixing the others until convergence. The update rules of these variables can be seen in Equation (6), where $\mathbf{F}_A = \mathbf{A}_4 \odot \cdots \mathbf{A}_{i+1} \odot \cdots \mathbf{A}_{i-1} \odot \mathbf{A}_1$, and $\mathbf{F}_B = \mathbf{B}_4 \odot \cdots \mathbf{B}_{i+1} \odot \cdots \mathbf{B}_{i-1} \odot \mathbf{B}_1$, “ \odot ” is the Khatri-Rao product. $\mathbf{X}_{\text{HC},i} \in \mathbb{R}_+^{I_i \times \prod_{k \neq i}^4 I_k}$ and $\mathbf{X}_{\text{MDD},i} \in \mathbb{R}_+^{I_i \times \prod_{k \neq i}^4 I_k}$ mean the mode- i matricization of tensors \mathcal{X}_{HC} and \mathcal{X}_{MDD} . $(\cdot)^C$ represents the first L_i columns of the matrix and $(\cdot)^I$ is the remaining columns. Analogously, we can obtain the updating solutions of other variables. It should be noted that the derivation of the coupling parts $\mathbf{A}_i^C = \mathbf{B}_i^C$, $i = 1, 2, 3$ should be refer to the information from \mathcal{X}_{HC} and \mathcal{X}_{MDD} . The entire optimization process is termed as NCTF-ADMM algorithm and summarized in Algorithm 1. Moreover, in this study, two stopping criteria were adopted in the algorithm optimization. (i) $\|\text{RelErr}_{\text{new}} - \text{RelErr}_{\text{old}}\| < \text{tol}$, it means that the relative error (RelErr) change of data fittings between the adjacent iterations should be smaller than tol (here we set $\text{tol} = 10e - 8$). RelErr is defined as $\text{RelErr} = \frac{\|\mathcal{X}_{\text{HC}} - \tilde{\mathcal{X}}_{\text{HC}}\|_F}{\|\mathcal{X}_{\text{HC}}\|_F} + \frac{\|\mathcal{X}_{\text{MDD}} - \tilde{\mathcal{X}}_{\text{MDD}}\|_F}{\|\mathcal{X}_{\text{MDD}}\|_F}$, $\tilde{\mathcal{X}}_{\text{HC}}$ and $\tilde{\mathcal{X}}_{\text{MDD}}$ are the recovered tensors. Meanwhile, the Fit value is defined as $\text{Fit} = 1 - \frac{\text{RelErr}}{2}$. (ii) The maximum number of iterations is no more than 1,000.

2.3. Correlation Analysis

To discover the relationships between musical stimuli and the EEG data, five musical features were first extracted from the music stimuli. After performing NCTF-ADMM algorithm on the HC and MDD tensor data, a correlation analysis method was conducted between time courses of extracted temporal components and time courses of musical features. We adopted Pearson correlation to calculate the correlation efficient, and applied the Monte Carlo method and permutation test to determine the significant thresholds of the correlation and correct for multiple comparisons (Alluri et al., 2012; Cong et al.,

$$\begin{cases}
\mathbf{A}_i^C = \mathbf{B}_i^C = \\
\left[\mathbf{X}_{\text{HC},i} \mathbf{F}_A^C + \mathbf{X}_{\text{MDD},i} \mathbf{F}_B^C - \mathbf{A}_i^l (\mathbf{F}_A^l)^T \mathbf{F}_A^C - \mathbf{B}_i^l (\mathbf{F}_B^l)^T \mathbf{F}_B^C - \mathbf{\Lambda}_i^C - \mathbf{\Gamma}_i^C + \rho_i \tilde{\mathbf{A}}_i^C + \sigma_i \tilde{\mathbf{B}}_i^C \right] \\
\left[(\mathbf{F}_A^C)^T \mathbf{F}_A^C + (\mathbf{F}_B^C)^T \mathbf{F}_B^C + (\rho_i + \sigma_i) \mathbf{I} \right]^{-1} \\
\mathbf{A}_i^l = \left[\mathbf{X}_{\text{HC},i} \mathbf{F}_A^l - \mathbf{A}_i^C (\mathbf{F}_A^l)^T \mathbf{F}_A^C - \mathbf{\Lambda}_i^l + \rho_i \tilde{\mathbf{A}}_i^l \right] \left[(\mathbf{F}_A^l)^T \mathbf{F}_A^l + \rho_i \mathbf{I} \right]^{-1} \\
\mathbf{B}_i^l = \left[\mathbf{X}_{\text{MDD},i} \mathbf{F}_B^l - \mathbf{B}_i^C (\mathbf{F}_B^l)^T \mathbf{F}_B^C - \mathbf{\Gamma}_i^l + \sigma_i \tilde{\mathbf{B}}_i^l \right] \left[(\mathbf{F}_B^l)^T \mathbf{F}_B^l + \sigma_i \mathbf{I} \right]^{-1} \\
\tilde{\mathbf{A}}_i = [\mathbf{A}_i + \mathbf{\Lambda}_i]_+, \tilde{\mathbf{B}}_i = [\mathbf{B}_i + \mathbf{\Gamma}_i]_+, \mathbf{\Lambda}_i = \mathbf{\Lambda}_i + \mathbf{A}_i - \tilde{\mathbf{A}}_i, \mathbf{\Gamma}_i = \mathbf{\Gamma}_i + \mathbf{B}_i - \tilde{\mathbf{B}}_i
\end{cases} \quad (6)$$

Algorithm 1: NCTF-ADMM algorithm

Input: $\mathcal{X}_{\text{HC}}, \mathcal{X}_{\text{MDD}}, R_{\text{HC}}$ and R_{MDD}

1 Initialization:

2 $\mathbf{A}_i, \mathbf{B}_i, \tilde{\mathbf{A}}_i, \tilde{\mathbf{B}}_i, \mathbf{\Lambda}_i, \mathbf{\Gamma}_i$ and $L_i, i = 1 \cdots 4$

3 **while** stopping criterion is not satisfied **do**

4 **for** $i = 1, \cdots, 4$ **do**

5 According to Equation (6)

6 Update primal variables: \mathbf{A}_i and \mathbf{B}_i ;

7 Update auxiliary variables: $\tilde{\mathbf{A}}_i$ and $\tilde{\mathbf{B}}_i$;

8 Update dual variables: $\mathbf{\Lambda}_i$ and $\mathbf{\Gamma}_i$;

9 **end**

10 **end**

Output: $\mathbf{A}_i, \mathbf{B}_i, i = 1, \cdots, 4$

2013; Wang et al., 2020). For the time course of each musical feature, a threshold of correlation coefficient at a significant level of $p < 0.05$ was calculated with the time courses of extracted temporal components. Then those components whose temporal components satisfied significant correlation were considered to be related to musical stimuli, and will be of interest and further analyzed.

2.4. Shared and Unshared Feature Clustering

In order to guarantee the reliability of the results, we independently performed the constrained tensor factorization multiple times (in this study we set 50 times). After performing correlation analysis for the multiple results, we adopted clustering method to cluster the shared and unshared spatial components, respectively. Meanwhile, we merged the corresponding spectral component and counted the musical feature distributions that were involved in the same cluster. For stable clustering, we adopted hierarchical agglomerative clustering algorithm, in which complete linkage was used to calculate the furthest distance (here we used correlation) between pairs of clusters and the pairs of clusters with the nearest distance were merged.

3. RESULTS

The EEG data used in this study can be obtained from the corresponding authors according to reasonable requirements,

and the code to reproduce the simulation in Section 3.1 is available at <https://github.com/xiulinwang/FrontierHN-NCTF-ADMM>.

The following experiments are done with the following computer configurations; CPU: Intel® Xeon(R) E5-2680 v2 @ 2.80 Hz \times 40; Memory: 125.80 GiB; System: 64-bit ubuntu 16.04; Matlab R2014b.

3.1. Simulation Results

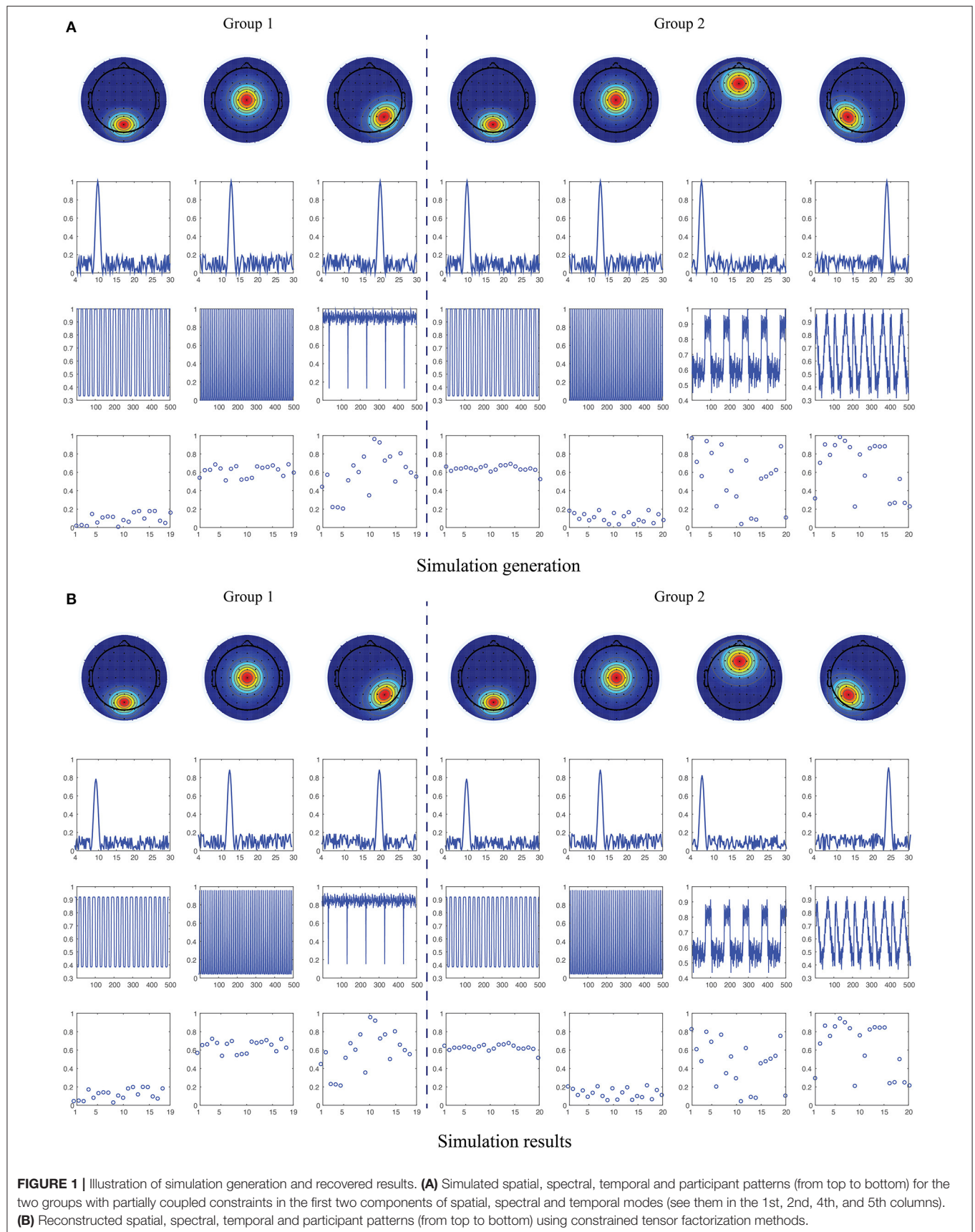
In this study, we first designed the simulation data to verify the performance of the proposed constrained tensor factorization method. We generated four kinds of predefined component factors indicating spatial, spectral, temporal and participant information, respectively, and then constructed two fourth-order tensors representing the simulated HC and MDD data via the outer product of corresponding vectors as follows:

$$\mathcal{X} = \sum_{r=1}^R \mathbf{u}_r^{(1)} \mathbf{u}_r^{(2)} \mathbf{u}_r^{(3)} \mathbf{u}_r^{(4)} \quad (7)$$

where $\mathbf{u}_r^{(1)} \in \mathbb{R}_+^{64 \times 1}$, $\mathbf{u}_r^{(2)} \in \mathbb{R}_+^{130 \times 1}$, $\mathbf{u}_r^{(3)} \in \mathbb{R}_+^{500 \times 1}$ and $\mathbf{u}_r^{(4)} \in \mathbb{R}_+^{19(20) \times 1}$ present topography, power spectrum, waveform and magnitude of participant, respectively, as shown in **Figure 1A**. $\mathcal{X} \in \mathbb{R}_+^{64 \times 130 \times 500 \times 19(20)}$ denotes the ground true EEG data, and the noised synthetic EEG data was generated as

$$\mathcal{Z} = \sigma_x \frac{\mathcal{X}}{\|\mathcal{X}\|} + \sigma_n \frac{\mathcal{N}}{\|\mathcal{N}\|} \quad (8)$$

where σ_x and σ_n denote the levels of signal and noise. \mathcal{N} is the noise tensor data uniformly distributed on the open interval (0, 1) and of the same size with \mathcal{X} . SNR refers to the signal-to-noise ratio defined as $\text{SNR} = 10 \log_{10}(\sigma_x / \sigma_n)$, and we set SNR to 20dB in this experiment. For the two synthetic tensors, the number of component is set to $R_{\text{HC}} = 3$ and $R_{\text{MDD}} = 4$. We assume there are two common components in the spatial, spectral and temporal modes between two tensors, i.e., $L_1 = L_2 = L_3 = 2$ and $L_4 = 0$, which are parallel to the assumptions in the following ongoing EEG data processing. The spatial patterns were generated based on the brain activations located in the occipito-parietal, center, right occipital, frontal and left occipital regions (the 1st row of **Figure 1A**), respectively, corresponding to the frequency fluctuations centered at 10, 13, 20, 7, and 24 Hz (the 2nd row of **Figure 1A**). The temporal patterns were constructed



using the time courses from the benchmark simulated complex fMRI dataset¹ (the 3rd row of **Figure 1A**). The magnitude of participants is uniformly distributed, but for the common components, we limited the corresponding magnitude to (0, 0.2) and (0.5, 0.7) in order to better discriminate the two groups (the 4th row of **Figure 1A**).

We applied NCTF-ADMM algorithm to the simulated HC and MDD data, and the extracted spatial, spectral, temporal and participant components can be seen in **Figure 1B**. We conducted the algorithm 50 times, and got the averaged tensor fit of 0.9343 with the averaged running time of 111.85 s. The averaged correlation between the two sets of true factor matrices and recovered factor matrices is close to 1.

3.2. EEG Results

In terms of the number of components for the tensor data, a simple explained variance-based principal component analysis (PCA) method was adopted (Liu et al., 2021), and the number of principal components with 99% accumulated explained variance were assigned to R_{HC} and R_{MDD} , then we set $R_{HC} = 26$, $R_{MDD} = 36$. Regarding the number of coupled components, we first ran the NCTF-ADMM without coupling constraints 10 times, and then we directly calculated the correlations in the spectral/spatial modes and performed correlation analysis in the temporal mode between the two groups of data, respectively. Finally, we selected the averaged number of highly correlated (0.87 and 0.90) and significantly correlated ($p < 0.05$) components as the number of shared components, i.e., $L_1 = L_2 = L_3 = 17$ and $L_4 = 0$.

We first carried out the proposed NCTF-ADMM algorithm 50 times on the two groups of ongoing EEG data, and then through correlation analysis and hierarchical clustering, we totally obtained 5 clusters of shared and unshared component patterns between/in HC and MDD groups. The averaged topographies, power spectrum, musical feature distribution and spatial correlation maps in the same cluster are plotted in **Figure 2**. Specifically, from the shared components extracted from HC and MDD data, we found two clusters of interested component patterns which were considered to be activated by the music modulation, and the probabilities of components in clusters #1 and #2 occurring in 50 runs reach 96% and 90%. Regarding cluster #1, the topography reveals that the right parietal region of the brain was activated with the low alpha oscillations modulated by the music feature “Mode”, while the cluster #2 represents the activation of parietal region of the brain with high alpha oscillations. The averaged correlations of spatial components in clusters #1 and #2 reached 0.9618 and 0.9757. For unshared components in HC group, we obtained one cluster in which the left occipital region was activated with the alpha oscillations and mostly modulated by the music features “Key Clarity” and “Fluctuation Centroid”. This cluster was unstable probably due to the low signal-to-noise ratio nature of EEG signals, and it was only clustered 32 times out of 50 times with the averaged spatial correlation of 0.8765. Moreover, the shared components in the MDD group were included into two clusters, cluster #4 reveals that the frontal region of the brain

was activated with theta oscillations and modulated by the music feature “Mode”, and cluster #5 reveals that the modulation of musical feature “Key Clarity” brought about the activation of alpha oscillations in the parietal-occipital region of the brain. The occurrence probabilities of spatial components from clusters #4 and #5 in 50 runs are 100% and 96% with the averaged spatial correlations of 0.9831 and 0.8408.

4. DISCUSSION

In this study, we investigated the shared and unshared brain activities of spatio-temporal-spectral modes from the HC and MDD data using EEG collections during freely listening to music. To this end, we proposed a complete solution combining constrained tensor factorization, musical information retrieval and spatial clustering, in which the brain activities of interest that were believed to be activated by music modulation were discovered. Through the time-frequency representation, we constructed two fourth-order tensors of time \times frequency \times space \times participant for HC and MDD groups, and the two tensors were decomposed at the same time with the extraction of common and individual components using nonnegative coupled tensor factorization, and its performance has been verified using simulation data in section 3.1. Meanwhile, five long-term musical features, including two tonal and three rhythmic features, were extracted using the MIR toolbox. Then we adopted the correlation analysis to identify the components of interest whose temporal components were significantly correlated with any of the five music features, and performed clustering analysis on the outcomes of correlation analysis of the repeated runs in order to obtain reliable and convincing results.

For the simulated data, as shown in **Figure 1**, we can see that the simulated factor matrices representing space, frequency, time and participant loadings were successfully recovered using nonnegative and coupled tensor factorization with high tensor fittings. The participant loadings of the two coupled components are significantly different in the magnitude distribution. Compared with the conventional tensor factorization (Cichocki et al., 2015; Cong et al., 2015; Sidiropoulos et al., 2017), the constrained tensor factorization applied in this study simultaneously considered the shared and unshared information between/in the two groups of data via imposing the coupling structure (Zhou et al., 2016; Wang et al., 2019, 2020). Moreover, it can achieve unique solutions and interpretable components, while circumventing the independence constraint compared to its matrix counterparts (Calhoun et al., 2009; Hunyadi et al., 2017; Labounek et al., 2018; Zhu et al., 2021). As we all know, EEG signals are recorded as the multi-way tensors of time, space, subject, or group modes, thus multi-way analysis methods are attractive and promising tools for processing such tensor data, while the two-way analysis will lose the multilinear structure and hidden internal relationships in the original data (Cong et al., 2015).

The researchers have reported that a considerable amount of neural dynamic changes distributed in multiple subcortical and cortical regions (such as auditory, tactile, visual) were found in

¹http://mlsp.umbc.edu/simulated_complex_fmri_data.html

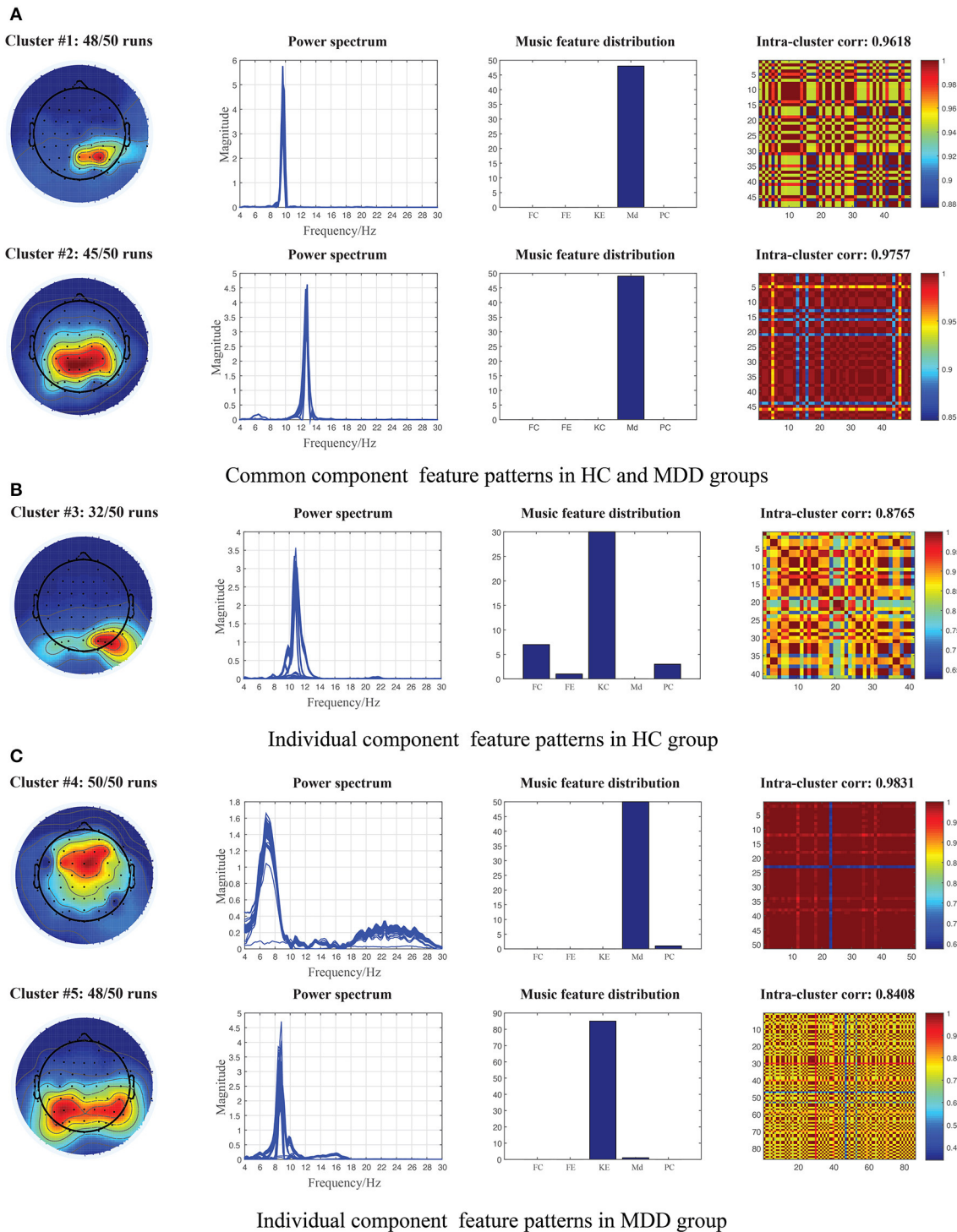


FIGURE 2 | Illustration of the extracted component patterns (from left to right column: mean topography, overall power spectrum, music feature distribution and intra-cluster correlation maps) from HC and MDD EEG data via 50 runs of constrained tensor factorization method, and the parallel temporal component was significantly correlated with at least one of the musical features. **(A)** Common component patterns clustered from the shared components of 50 runs between HC and MDD data. **(B)** Individual component patterns clustered from the individual components of 50 runs in HC data. **(C)** Individual component patterns clustered from the individual components of 50 runs in MDD data. Md, Mode; KE, Key Clarity; FC, Fluctuation Centroid; FE, Fluctuation entropy; PC, Pulse Clarity.

EEG/fMRI recordings during music perception (Andrade and Bhattacharya, 2003; Khalfa et al., 2005; King et al., 2019). These regions including hippocampus, hypothalamus, orbitofrontal and ventral medial prefrontal cortex are typically involved during emotion evocation, processing and hedonic regulation in voices (Menon and Levitin, 2005). Therefore, the potential of music to modulate activities in brain networks is worth investigating in MDD. Using the constrained tensor factorization-based framework, we were able to provide the feasibility to identify the brain dynamics involved in the processing of acoustical music from the ongoing EEG signals, and finally obtained two shared and three unshared feature patterns between or in HC and MDD data as shown in **Figure 2**. Studies have found that music can evoke a variety of emotions from simple arousal responses, basis emotions to more complex emotions such as subjective feeling, emotional expression or physiological changes in listeners (Witvliet and Vrana, 2007; Juslin et al., 2015). First, from the shared components extracted from both HC and MDD data, we obtained two clusters of interest patterns including a series of ~ 10 Hz right parietal components and ~ 13 Hz centro-parietal components, which were believed to be elicited by the tonal music feature of “Mode”. As we all know, MDD is a kind of mental disorders characterized by affective and cognitive dysfunctions, and existing studies have shown that brain networks of MDD patients have abnormal network topology structure (Gotlib and Joormann, 2010; Jia et al., 2010; Mulders et al., 2015). Moreover, the study also reported individuals with MDD were associated with impaired recognition of emotion in music as well as in facial and vocal stimuli (Naranjo et al., 2011). Therefore, in addition to some basic emotional processing and regulation that are indistinguishable from the MDD patients and the HCs, some uncontrolled responses with music of MDD patients may be more negative due to their cognitive dysfunctions. We also observed the right occipital components of ~ 10 Hz oscillations mostly elicited by the tonal feature of “Key clarity” and the rhythmic feature “Fluctuation Centroid” which were clustered from the individual components of HC group, but such brain dynamics in MDD patients were not sensitive to music perception and were suppressed. The dopaminergic system is activated during music processing, however, some dopamine responses to music in MDD patients may be weakened, which makes some brain neural dynamics that should be appeared not captured (Menon and Levitin, 2005; Blum et al., 2010). Our findings replicate some of the results of our previous studies in which similar alpha brain oscillations located in the centro-parietal or occipital regions were found from the EEG signals of 14 healthy participants when freely listening to music (Cong et al., 2013; Wang et al., 2020). Lin et al. also found the electrodes of parietal lobes across alpha band contributed a lot in the emotion recognition during music listening (Lin et al., 2010). Second, for the individual components extracted from MDD patients, we obtained two clusters of interest: theta oscillations in the frontal region and alpha oscillations in the bilateral parieto-occipital region, which were considered more overactive than the HC groups. From the perspective of functional connectivity in the source level, Liu et al. revealed three oscillatory hyperconnectivity networks including right hemisphere of alpha and beta bands, left auditory

region of delta band and prefrontal region of delta band in MDD (Liu et al., 2021). The frontal region was involved in planning complex cognitive behaviour, decision making and working memory (Liu et al., 2021). In other words, the frontal regions played an important role in depression development and have received widespread attention (Rajkowska et al., 1999). Previous studies reported that the fronto-limbic neural networks were implicated in MDD, particularly in relation to the subgenual anterior cingulate cortex (ACC) which was considered to regulate amygdala activity in order to prevent excessive emotional reactivity and stress responses (Drevets et al., 1997; Phillips et al., 2003). The fMRI studies also uncovered the modification of functions in frontal and temporal regions (Wang et al., 2012). The alpha temporo-occipital component located in the left angular gyrus was engaged in music perception from most MDD patients (Zhu et al., 2021). Significant alternations of brain dynamics in the left frontal lobe, (left) parieto-occipital lobe in theta and alpha bands were observed from the functional networks of MDD patients when using resting-state EEG signals (Sun et al., 2019; Zhang et al., 2020). The abnormal regions near parieto-occipital sulcus in MDD may be associated with the inability to detach from the visual dorsal stream, robust biases in attention or inhibitory control of irrelevant sensory (Gotlib and Joormann, 2010; Sacchet et al., 2016). Meanwhile, the sources in parieto-occipital regions were considered to contribute to the working memory load in the alpha band (Tuladhar et al., 2007). Our results are indeed consistent with some of the research findings in MDD, but in the end, it is difficult to compare directly due to the different methodological approaches and selected participants.

In conclusion, we provide a comprehensive framework for the shared and unshared feature extraction from the EEG recordings of MDD and HC groups during music listening, our findings are well supported and in line with the results of some previous studies to some extent, and contribute to providing some novel biomarkers for the clinical diagnosis and treatment of MDD patients. Meanwhile, the proposed methods based on nonnegative coupled tensor factorization may provide a new perspective for the analysis of other EEG recordings or the data with other psychiatric disorders. However, there are still some limitations in this study. First, we directly assume the temporal, spatial and spectral consistency among participants in MDD or HC group; that is, we only pay attention to group similarities and differences between the MDD and HC groups, and ignore the participant differences in each individual group, which will be a key issue that needs to be considered in our future work. Second, the analysis in this study was performed at the sensor level, which will be extended to the source level in the following work. Third, we have proposed a way to select the number of common components by correlation analysis. However, its selection is still subjective to some extent, which remains an open issue and invites more discussion.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by First Affiliated Hospital of Dalian Medical University and Dalian University of Technology. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

XLW conceived and carried out the idea of the study, and drafted the manuscript under the guidance of JW and FC throughout the preparation process. WL took part in the introduction and discussion parts and provided critical revisions and feedback with the rest of authors. XYW preprocessed the

data and revised the manuscript. ZM designed the experiment and EEG collected the data. JX and YC provided the EEG data used in this study, and QZ reviewed the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work is supported by National Natural Science Foundation of China (grant no. 91748105), National Foundation in China (no. JCKY2019110B009 and 2020-JCJQ-JJ-252), the Fundamental Research Funds for the Central Universities (DUT2019 and DUT20LAB303) in Dalian University of Technology in China, Dalian Science and Technology Innovation Fund Project (2021JJ12SN38), and the scholarship from China scholarship Council (no. 201706060263).

REFERENCES

- Alluri, V., Toiviainen, P., Jääskeläinen, I. P., Glerean, E., Sams, M., and Brattico, E. (2012). Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. *Neuroimage* 59, 3677–3689. doi: 10.1016/j.neuroimage.2011.11.019
- Andrade, P. E., and Bhattacharya, J. (2003). Brain tuned to music. *J. R. Soc. Med.* 96, 284–287. doi: 10.1177/014107680309600607
- Belmaker, R., and Agam, G. (2008). Major depressive disorder. *N. Engl. J. Med.* 358, 55–68. doi: 10.1056/NEJMra073096
- Blum, K., Chen, T. J., Chen, A. L., Madigan, M., Downs, B. W., Waite, R. L., et al. (2010). Do dopaminergic gene polymorphisms affect mesolimbic reward activation of music listening response? therapeutic impact on reward deficiency syndrome (rds). *Med. Hypoth.* 74, 513–520. doi: 10.1016/j.mehy.2009.10.008
- Boyd, S., Parikh, N., and Chu, E., Peleato, B., and Eckstein, J. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.* 3, 1–122. doi: 10.1561/22000000016
- Buzsaki, G. (2006). *Rhythms of the Brain*. New York, NY: Oxford University Press.
- Calhoun, V. D., Liu, J., and Adahi, T. (2009). A review of group ica for fmri data and ica for joint inference of imaging, genetic, and erp data. *Neuroimage* 45, S163–S172. doi: 10.1016/j.neuroimage.2008.10.057
- Carroll, J. D., and Chang, J.-J. (1970). Analysis of individual differences in multidimensional scaling via an n-way generalization of "eckart-young" decomposition. *Psychometrika* 35, 283–319. doi: 10.1007/BF02310791
- Cichocki, A., Mandic, D., De Lathauwer, L., Zhou, G., Zhao, Q., Caiafa, C., et al. (2015). Tensor decompositions for signal processing applications: from two-way to multiway component analysis. *IEEE Signal Process Mag.* 32, 145–163. doi: 10.1109/MSP.2013.2297439
- Cong, F., Alluri, V., Nandi, A. K., Toiviainen, P., Fa, R., Abu-Jamous, B., et al. (2013). Linking brain responses to naturalistic music through analysis of ongoing eeg and stimulus features. *IEEE Trans. Multimedia* 15, 1060–1069. doi: 10.1109/TMM.2013.2253452
- Cong, F., Lin, Q.-H., Kuang, L.-D., Gong, X.-F., Astikainen, P., and Ristaniemi, T. (2015). Tensor decomposition of eeg signals: a brief review. *J. Neurosci. Methods* 248, 59–69. doi: 10.1016/j.jneumeth.2015.03.018
- Cong, F., Phan, A. H., Zhao, Q., Nandi, A. K., Alluri, V., Toiviainen, P., et al. (2012). "Analysis of ongoing eeg elicited by natural music stimuli using nonnegative tensor factorization," in *2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)* (Bucharest: IEEE), 494–498.
- Delorme, A., and Makeig, S. (2004). Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009
- Drevets, W. C., Price, J. L., Simpson, J. R., Todd, R. D., Reich, T., Vannier, M., et al. (1997). Subgenual prefrontal cortex abnormalities in mood disorders. *Nature* 386, 824–827. doi: 10.1038/386824a0
- Gotlib, I. H., and Hamilton, J. P. (2008). Neuroimaging and depression: current status and unresolved issues. *Curr. Dir. Psychol. Sci.* 17, 159–163. doi: 10.1111/j.1467-8721.2008.00567.x
- Gotlib, I. H., and Joormann, J. (2010). Cognition and depression: current status and future directions. *Annu. Rev. Clin. Psychol.* 6, 285–312. doi: 10.1146/annurev.clinpsy.121208.131305
- Harshman, R. A. (1970). Foundations of the parafac procedure: models and conditions for an "explanatory" multimodal factor analysis. *UCLA Work. Pap. Phon.* 16, 1–84.
- Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., and Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science* 303, 1634–1640. doi: 10.1126/science.1089506
- Hitchcock, F. L. (1927). The expression of a tensor or a polyadic as a sum of products. *J. Math. Phys.* 6, 164–189. doi: 10.1002/sapm192761164
- Huang, K., Sidiropoulos, N. D., and Liavas, A. P. (2016). A flexible and efficient algorithmic framework for constrained matrix and tensor factorization. *IEEE Trans. Signal Proc.* 64, 5052–5065. doi: 10.1109/TSP.2016.2576427
- Hunyadi, B., Dupont, P., Van Paesschen, W., and Van Huffel, S. (2017). Tensor decompositions and data fusion in epileptic electroencephalography and functional magnetic resonance imaging data. *Wiley Interdiscipl. Rev. Data Min. Knowl. Discov.* 7, e1197. doi: 10.1002/widm.1197
- Jia, Z., Huang, X., Wu, Q., Zhang, T., Lui, S., Zhang, J., et al. (2010). High-field magnetic resonance imaging of suicidality in patients with major depressive disorder. *Am. J. Psychiatry* 167, 1381–1390. doi: 10.1176/appi.ajp.2010.09101513
- Juslin, P. N., Barradas, G., and Eerola, T. (2015). From sound to significance: exploring the mechanisms underlying emotional reactions to music. *Am. J. Psychol.* 128, 281–304. doi: 10.5406/amerjpsyc.128.3.0281
- Kaiser, R. H., Andrews-Hanna, J. R., Wager, T. D., and Pizzagalli, D. A. (2015). Large-scale network dysfunction in major depressive disorder: a meta-analysis of resting-state functional connectivity. *JAMA Psychiatry* 72, 603–611. doi: 10.1001/jamapsychiatry.2015.0071
- Khalfa, S., Schon, D., Anton, J.-L., and Liégeois-Chauvel, C. (2005). Brain regions involved in the recognition of happiness and sadness in music. *Neuroreport* 16, 1981–1984. doi: 10.1097/00001756-200512190-00002
- King, J., Jones, K., Goldberg, E., Rollins, M., MacNamee, K., Moffit, C., et al. (2019). Increased functional connectivity after listening to favored music in adults with alzheimer dementia. *J. Prevent. Alzheimers Dis.* 6, 56–62. doi: 10.14283/jpad.2018.19
- Kolda, T. G., and Bader, B. W. (2009). Tensor decompositions and applications. *SIAM Rev.* 51, 455–500. doi: 10.1137/07070111X
- Labounek, R., Bridwell, D. A., Mareček, R., Lamoš, M., Mikl, M., Slaviček, T., et al. (2018). Stable scalp eeg spatio-spectral patterns across paradigms estimated by group ica. *Brain Topogr.* 31, 76–89. doi: 10.1007/s10548-017-0585-8

- Lartillot, O., and Toivainen, P. (2007). "A matlab toolbox for musical feature extraction from audio," in *International Conference on Digital Audio Effects*, Vol. 237 (Bordeaux), 244.
- Lin, Y.-P., Wang, C.-H., Jung, T.-P., Wu, T.-L., Jeng, S.-K., Duann, J.-R., et al. (2010). Eeg-based emotion recognition in music listening. *IEEE Trans. Biomed. Eng.* 57, 1798–1806. doi: 10.1109/TBME.2010.2048568
- Liu, W., Wang, X., Xu, J., Chang, Y., Hämäläinen, T., and Cong, F. (2021). Identifying oscillatory hyperconnectivity and hypoconnectivity networks in major depression using coupled tensor decomposition. *IEEE Trans. Neural Syst. Rehabil. Eng.* 29, 1895–1904. doi: 10.1109/TNSRE.2021.3111564
- Liu, W., Zhang, C., Wang, X., Xu, J., Chang, Y., Ristaniemi, T., et al. (2020). Functional connectivity of major depression disorder using ongoing eeg during music perception. *Clin. Neurophysiol.* 131, 2413–2422. doi: 10.1016/j.clinph.2020.06.031
- Maratos, A., Gold, C., Wang, X., and Crawford, M. (2008). Music therapy for depression. *Cochrane Database Syst. Rev.* 2017:CD004517. doi: 10.1002/14651858.CD004517.pub2
- Menon, V., and Levitin, D. J. (2005). The rewards of music listening: response and physiological connectivity of the mesolimbic system. *Neuroimage* 28, 175–184. doi: 10.1016/j.neuroimage.2005.05.053
- Michael, A. J., Krishnaswamy, S., and Mohamed, J. (2005). An open label study of the use of eeg biofeedback using beta training to reduce anxiety for patients with cardiac events. *Neuropsychiatr. Dis. Treat.* 1, 357.
- Morup, M., Hansen, L. K., and Arnfred, S. M. (2007). ErpwaveLab: a toolbox for multi-channel analysis of time-frequency transformed event related potentials. *J. Neurosci. Methods* 161, 361–368. doi: 10.1016/j.jneumeth.2006.11.008
- Mulders, P. C., van Eijndhoven, P. F., Schene, A. H., Beckmann, C. F., and Tendolcar, I. (2015). Resting-state functional connectivity in major depressive disorder: a review. *Neurosci. Biobehav. Rev.* 56, 330–344. doi: 10.1016/j.neubiorev.2015.07.014
- Naranjo, C., Kornreich, C., Campanella, S., Noël, X., Vandriette, Y., Gillain, B., et al. (2011). Major depression is associated with impaired processing of emotion in music as well as in facial and vocal stimuli. *J. Affect. Disord.* 128, 243–251. doi: 10.1016/j.jad.2010.06.039
- Phillips, M. L., Drevets, W. C., Rauch, S. L., and Lane, R. (2003). Neurobiology of emotion perception ii: Implications for major psychiatric disorders. *Biol. Psychiatry* 54, 515–528. doi: 10.1016/S0006-3223(03)00171-9
- Rajkowska, G., Miguel-Hidalgo, J. J., Wei, J., Dille, G., Pittman, S. D., Meltzer, H. Y., et al. (1999). Morphometric evidence for neuronal and glial prefrontal cell pathology in major depression. *Biol. Psychiatry* 45, 1085–1098. doi: 10.1016/S0006-3223(99)00041-4
- Sacchet, M. D., Ho, T. C., Connolly, C. G., Tymofiyeva, O., Lewinn, K. Z., Han, L. K., et al. (2016). Large-scale hypoconnectivity between resting-state functional networks in unmedicated adolescent major depressive disorder. *Neuropsychopharmacology* 41, 2951–2960. doi: 10.1038/npp.2016.76
- Sadaghiani, S., Brookes, M. J., and Baillet, S. (2021). Connectomics of human electrophysiology. *PsyArXiv*. doi: 10.31234/osf.io/dr7zh
- Schenker, C., Cohen, J. E., and Acar, E. (2020). A flexible optimization framework for regularized matrix-tensor factorizations with linear couplings. *IEEE J. Select. Top. Signal Proc.* 15, 506–521. doi: 10.1109/JSTSP.2020.3045848
- Sidiropoulos, N. D., De Lathauwer, L., Fu, X., Huang, K., Papalexakis, E. E., and Faloutsos, C. (2017). Tensor decomposition for signal processing and machine learning. *IEEE Trans. Signal Proc.* 65, 3551–3582. doi: 10.1109/TSP.2017.2690524
- Sonkusare, S., Breakspear, M., and Guo, C. (2019). Naturalistic stimuli in neuroscience: critically acclaimed. *Trends Cogn. Sci.* 23, 699–714. doi: 10.1016/j.tics.2019.05.004
- Sun, S., Li, X., Zhu, J., Wang, Y., La, R., Zhang, X., et al. (2019). Graph theory analysis of functional connectivity in major depression disorder with high-density resting state eeg data. *IEEE Trans. Neural Syst. Rehabil. Eng.* 27, 429–439. doi: 10.1109/TNSRE.2019.2894423
- Tucker, L. R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika* 31, 279–311. doi: 10.1007/BF02289464
- Tuladhar, A. M., Huurne, N., t., Schoffelen, J.-M., Maris, E., Oostenveld, R., et al. (2007). Parieto-occipital sources account for the increase in alpha activity with working memory load. *Hum. Brain Mapp.* 28, 785–792. doi: 10.1002/hbm.20306
- Wang, L., Hermens, D. F., Hickie, I. B., and Lagopoulos, J. (2012). A systematic review of resting-state functional-mri studies in major depression. *J. Affect. Disord.* 142, 6–12. doi: 10.1016/j.jad.2012.04.013
- Wang, X., Liu, W., Toivainen, P., Ristaniemi, T., and Cong, F. (2020). Group analysis of ongoing eeg data based on fast double-coupled nonnegative tensor decomposition. *J. Neurosci. Methods* 330:108502. doi: 10.1016/j.jneumeth.2019.108502
- Wang, X., Ristaniemi, T., and Cong, F. (2019). "Fast implementation of double-coupled nonnegative canonical polyadic decomposition," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (Brighton: IEEE), 8588–8592.
- Witvliet, C. V., and Vrana, S. R. (2007). Play it again sam: Repeated exposure to emotionally evocative music polarises liking and smiling responses, and influences other affective reports, facial emg, and heart rate. *Cogn. Emot.* 21, 3–25. doi: 10.1080/02699930601000672
- Yan, Y., Qian, T., Xu, X., Han, H., Ling, Z., Zhou, W., et al. (2020). Human cortical networking by probabilistic and frequency-specific coupling. *Neuroimage* 207:116363. doi: 10.1016/j.neuroimage.2019.116363
- Zhang, B., Yan, G., Yang, Z., Su, Y., Wang, J., and Lei, T. (2020). Brain functional networks based on resting-state eeg data for major depressive disorder analysis and classification. *IEEE Trans. Neural Syst. Rehabil. Eng.* 29, 215–229. doi: 10.1109/TNSRE.2020.3043426
- Zhi, D., Calhoun, V. D., Lv, L., Ma, X., Ke, Q., Fu, Z., et al. (2018). Aberrant dynamic functional network connectivity and graph properties in major depressive disorder. *Front. Psychiatry* 9:339. doi: 10.3389/fpsy.2018.00339
- Zhou, G., Zhao, Q., Zhang, Y., Adali, T., Xie, S., and Cichocki, A. (2016). Linked component analysis from matrices to high-order tensors: Applications to biomedical data. *Proc. IEEE* 104, 310–331. doi: 10.1109/JPROC.2015.2474704
- Zhu, Y., Liu, J., Ye, C., Mathiak, K., Astikainen, P., Ristaniemi, T., et al. (2020). Discovering dynamic task-modulated functional networks with specific spectral modes using meg. *Neuroimage* 218:116924. doi: 10.1016/j.neuroimage.2020.116924
- Zhu, Y., Wang, X., Mathiak, K., Toivainen, P., Ristaniemi, T., Xu, J., et al. (2021). Altered eeg oscillatory brain networks during music-listening in major depression. *Int. J. Neural Syst.* 31, 2150001. doi: 10.1142/S0129065721500015

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Wang, Liu, Wang, Mu, Xu, Chang, Zhang, Wu and Cong. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Behavioral Experience-Sampling Methods in Neuroimaging Studies With Movie and Narrative Stimuli

Iiro P. Jääskeläinen^{1,2*}, Jyrki Ahveninen³, Vasily Klucharev², Anna N. Shestakova² and Jonathan Levy^{1,4}

¹ Brain and Mind Laboratory, Department of Neuroscience and Biomedical Engineering, Aalto University School of Science, Espoo, Finland, ² International Laboratory of Social Neurobiology, Institute of Cognitive Neuroscience, National Research University Higher School of Economics, Moscow, Russia, ³ Massachusetts General Hospital, Massachusetts Institute of Technology and Harvard Medical School, Athinoula A. Martinos Center for Biomedical Imaging, Charlestown, MA, United States, ⁴ Baruch Ivcher School of Psychology, Interdisciplinary Center Herzliya, Reichman University, Herzliya, Israel

OPEN ACCESS

Edited by:

Klaus Kessler,
Aston University, United Kingdom

Reviewed by:

Robert A. Seymour,
University College London,
United Kingdom
Samantha Elizabeth Anne
Gregory,
University of Salford, United Kingdom

*Correspondence:

Iiro P. Jääskeläinen
iiro.jaaskelainen@aalto.fi

Specialty section:

This article was submitted to
Cognitive Neuroscience,
a section of the journal
Frontiers in Human Neuroscience

Received: 12 November 2021

Accepted: 06 January 2022

Published: 27 January 2022

Citation:

Jääskeläinen IP, Ahveninen J, Klucharev V, Shestakova AN and Levy J (2022) Behavioral Experience-Sampling Methods in Neuroimaging Studies With Movie and Narrative Stimuli. *Front. Hum. Neurosci.* 16:813684. doi: 10.3389/fnhum.2022.813684

Movies and narratives are increasingly utilized as stimuli in functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG), and electroencephalography (EEG) studies. Emotional reactions of subjects, what they pay attention to, what they memorize, and their cognitive interpretations are all examples of inner experiences that can differ between subjects during watching of movies and listening to narratives inside the scanner. Here, we review literature indicating that behavioral measures of inner experiences play an integral role in this new research paradigm *via* guiding neuroimaging analysis. We review behavioral methods that have been developed to sample inner experiences during watching of movies and listening to narratives. We also review approaches that allow for joint analyses of the behaviorally sampled inner experiences and neuroimaging data. We suggest that building neurophenomenological frameworks holds potential for solving the interrelationships between inner experiences and their neural underpinnings. Finally, we tentatively suggest that recent developments in machine learning approaches may pave way for inferring different classes of inner experiences directly from the neuroimaging data, thus potentially complementing the behavioral self-reports.

Keywords: naturalistic stimuli, memory, attention, emotion, social cognition, language, movies, narratives

INTRODUCTION

Movies and narratives are increasingly used as stimuli in functional magnetic resonance imaging (fMRI), electroencephalography (EEG), and magnetoencephalography (MEG) studies (Bartels and Zeki, 2004; Hasson et al., 2004; Dmochowski et al., 2012; Lankinen et al., 2014) (for a recent review, see Jaaskelainen et al., 2021). Such media-based stimuli are especially advantageous in that they allow for elicitation of a range of ecologically valid emotional and cognitive states in experimental subjects that would be difficult to achieve using traditional neuroimaging paradigms. This opens up possibilities to study the underlying neural mechanisms, which is one of the key questions in cognitive neuroscience. Elicitation of genuine strong emotions provides one good example where movies and narratives are more robust than other types of stimuli (Westermann et al., 1996). It is, however, highly important to take into account inter-subject variability in how the media-based stimuli are experienced. For example, a movie clip that is positive-emotional for one subject

can be negatively experienced by another. Similarly, cognitive states and memories evoked by a narrative can significantly differ between subjects. Inter-subject variability in the elicited emotional and cognitive states can be utilized in the analysis of their neural basis. For example, if mental imagery is elicited only in some subjects during listening to an audiobook and if only those subjects exhibit activity in visual cortical areas, one can infer that visual cortical areas are important for mental imagery. It is worth mentioning that throughout the history of experimental psychology, there have been two competing schools of thought arguing on whether or not inner experiences can be consciously and reliably accessed by experimental subjects. For a recent account of this debate, and for arguments in favor of the validity of self-reporting subjective experiences see Hurlburt et al. (2017). A host of methods have been developed to assess the variability in behavioral experiences while subjects are watching movies or listening to narratives, and utilize the variability to guide neuroimaging data analyses (see **Figure 1**). Here, we review some of these important approaches.

INDIVIDUAL VARIABILITY IN MEMORY ENCODING GUIDING NEUROIMAGING DATA ANALYSIS

Post-experiment questionnaires can be used to quantify which aspects of movies or narratives have been memorized during neuroimaging. These provide a good example of a way to quantify the variability in how movies or narratives are experienced. The information provided by such questionnaires can then be used to guide neuroimaging data analyses. In one study, higher inter-subject correlation (ISC) of hippocampal hemodynamic activity, determined by calculating voxel-wise correlations between all pairs of subjects (Hasson et al., 2004; Kauppi et al., 2010; Pajula et al., 2012), went hand-in-hand with memorizability of movie contents 3 weeks later (Hasson et al., 2008). In another study, ISC of EEG activity correlated with subsequent memorization of movie clip content (Cohen and Parra, 2016). In a third example of this approach, subjects watched emotionally aversive vs. neutral movie clips during positron emission tomography. Memorization of events in the clips 3 weeks after the scan correlated with activity levels in the amygdala and orbitofrontal areas (Cahill et al., 1996). Naturally, assessing what is memorized of a stimulus afterward is straightforward to accomplish with post-neuroimaging questionnaires or recoups. Assessing how the subjects emotionally experienced media-based stimuli during neuroimaging is already a bit trickier. We review these approaches in the following.

ASSESSING EMOTIONAL EXPERIENCES ELICITED BY MOVIES AND NARRATIVES AFTER NEUROIMAGING

Movies and narratives represent a powerful tool for neuroimaging studies of emotions *via* their capability to

induce robust emotions. Since emotional experiences can be highly variable across individual subjects, we argue that it is essential to ask the subjects about their experiences to inform the fMRI data analysis. The simplest approach is to ask the subjects afterward (Jones and Fox, 1992; Cahill et al., 1996; Aftanas et al., 1998; Jaaskelainen et al., 2008; Nummenmaa et al., 2012, 2014; Raz et al., 2012). In one study, subjects were asked to rate their experienced emotional valence and arousal continuously in two separate re-showings of the movie clips after the neuroimaging session. The ratings were then used to inform the neuroimaging data analyses. ISC of brain activity in default-mode and dorsal attentional networks were observed to co-vary with experienced valence and arousal, respectively (Nummenmaa et al., 2012). Naturally, obtaining continuous ratings for more than two classes of emotional experiences would be highly taxing for subjects, and obtaining even just one continuous rating is strenuous in case the movie or narrative is a long one. To circumvent this problem, we advise to obtain ratings of experienced categorical emotions as non-continuous variables (Jones and Fox, 1992).

In addition to emotional experiences *per se* one can obtain self-reports of specific aspects of emotional processing such as emotion regulation. Success in emotion regulation during watching of anger-eliciting movie clips has been observed to correlate with greater involvement of VMPFC in emotion regulation network (Jacob et al., 2018). Frontal EEG activity has been associated with better emotion regulation during induction of emotions (Nitschke et al., 2004; Dennis and Solomon, 2010). Degree of self-reported suspense has also been successfully used in fMRI data analyses (Metz-Lutz et al., 2010; Hsu et al., 2014). As a word of caution, the chances for obtaining reliable ratings are lower for more ambiguous and interpretable emotion categories (e.g., social emotions). We advise careful attention to be paid to how subjects are instructed, and running pilot experiments when asking to self-report categories of higher ambiguity.

In case of experienced humorousness, sadness, emotional valence, and emotional arousal, post-neuroimaging recall closely matches what is experienced during first-time watching of movie clips during neuroimaging. This is true at least when subjects are re-viewing the movie clips as memory cues for recall of what they had experienced during the first viewing of the clips in the scanner (Hutcherson et al., 2005; Raz et al., 2012; Jaaskelainen et al., 2016). Further studies are needed to determine for which emotional and other experiences the post-neuroimaging self-reports are accurate enough [e.g., using our published method (Jaaskelainen et al., 2016)]. In cases where post-scan emotion-ratings do not seem reliable, there are ways to sample emotional experiences during neuroimaging. These are briefly discussed below.

ASSESSING EMOTIONAL EXPERIENCES ELICITED BY MOVIES AND NARRATIVES DURING NEUROIMAGING

In addition to the post-scanning self-reports of emotional experiences, paradigms have been developed where

Self Reports	Indirect models	Biological measures
During neuroimaging <ul style="list-style-type: none"> + Potentially capture inner experiences most reliably - Self-reporting may disturb task performance - Self-report task may alter the quality of experience - Limited number of experiences can be sampled After neuroimaging <ul style="list-style-type: none"> + Task performance is not affected by self-reports + Quality of experience is not affected by self-reports + Re-showing the stimulus serves as a recall cue - Inner experiences may be inaccurately recalled - Limited number of experiences can be sampled 	Associations (LSA) <ul style="list-style-type: none"> + Circumvents direct inquiry - Unclear specificity Phenomenology-based <ul style="list-style-type: none"> + Emulates naturalistic experience - Empirical approaches are yet to be determined 	Eye-movements <ul style="list-style-type: none"> + Can be collected during neuroimaging + Neither task performance nor experience affected - Very limited in disclosing inner experiences Physiological measures (ECG, GSR, fEMG) <ul style="list-style-type: none"> + Can be collected during neuroimaging + Neither task performance nor experience affected - Limited to specific inner experiences Neuroimaging <ul style="list-style-type: none"> + Neither task performance nor experience affected - So far limited to specific inner experiences

FIGURE 1 | Summary illustration of the different types of approaches that are available for sampling of inner experiences. The potential strengths and limitations are indicated with + and – signs.

subjects continuously rate their experienced emotions while viewing the emotional movie clips in the scanner (Hutcherson et al., 2005). This approach potentially captures the emotional experiences more accurately than recall-based self-reporting, as the experiences might be different when viewing the emotional clips for the first vs. second time. Lending support for this concern, at least in case of some inner experiences there seem to be mismatches between retrospective reports and moment-to-moment experiences obtained in the scanner (Hurlburt et al., 2015). On the other hand, emotional responses might be altered by the ongoing rating task due to differential attentional demands.

There are findings showing that brain activity patterns elicited during passive viewing of emotional movie clips differ from brain activity patterns elicited while subjects are rating their emotions during viewing of the clips in the fMRI scanner (Hutcherson et al., 2005). Given this, we recommend quantification of emotional experiences *via* asking the subjects after the neuroimaging session how they had experienced the media-based stimuli while in the scanner. We advise to use the stimuli as recall cues when obtaining the ratings and validation of the post-scan ratings with separate behavioral experiments (Jaaskelainen et al., 2016). In addition to emotions, there are cognitive states, such as mental imagery or mentalization, that are elicited by media-based stimuli. In the following, we review behavioral methods with which such inner experiences can be quantified.

SAMPLING INNER EXPERIENCES OTHER THAN EMOTIONS

In addition to experienced emotions, there are methods that have been developed to assess other inner experiences during neuroimaging such as mental imagery elicited by movies and narratives. These are called experience-sampling methods. One recent experience-sampling approach involved re-playing a narrative in segments to subjects after the neuroimaging

session. The subjects were then tasked to produce a list of words in 20–30 s after each segment best describing what had been on their minds while they had heard the segment in the fMRI scanner. Word-list similarities were then estimated with latent semantic analysis (LSA). This was followed by representational similarity analysis (RSA; Kriegeskorte et al., 2008) of shared similarities between word-listings and voxel-wise fMRI activity as assessed by ISC metrics (Saalasti et al., 2019). Naturally, such word lists could also be probed for, e.g., occurrence of emotional words to estimate emotional reactions. Multi-dimensional scaling provides a manual alternative for automated algorithms such as LSA. In this approach, volunteers rate pair-wise similarities between word lists using, e.g., visual-analog scale (Bisanz et al., 1978).

There are also studies that have used experience sampling with other types of task paradigms. One example of such is a study where experimental subjects were presented intermittently with beeps during recording of resting-state fMRI. The beeps prompted the subjects to describe their thoughts and feelings just before each beep *via* answering specific questions (Hurlburt et al., 2015). In another study, it was observed that differences in inner experiences as sampled with post-scanning self-reports predicted differences in resting-state functional connectivity (Gorgolewski et al., 2014). In this approach, a questionnaire was utilized that prompted for agreeing or disagreeing with given statements. Examples of the statements include presence of social cognition, thinking about present vs. past, and thinking in words vs. images. Subjects rich with imagery showed differential visual-area connectivity (Gorgolewski et al., 2014).

Notably, it seems that there can be differences between online vs. post-scanning experience sampling (Hurlburt et al., 2015). The challenge of sampling subjective experience without affecting the experience itself is well known in studies on consciousness (Block, 2005; Dehaene et al., 2006). How can we know whether one has a conscious experience without resorting to cognitive functions such as attention, memory, or inner speech? Naturally, this is an intrinsic limitation in cognitive science: how to sample experience

without altering its neural underpinnings. We discuss this in the following.

SOME CAVEATS OF SELF-REPORTS AND POTENTIAL WAYS TO CIRCUMVENT THEM

Self-report methods have their potential caveats. First of all, it is poorly understood to what extent humans can be self-aware of what is on their minds, due to limits of consciousness. While we may not be fully self-aware, it can be argued that with proper experience-sampling methods a good deal of the inner workings of our minds can be uncovered (Hurlburt et al., 2017). Importantly, self-reports can be used in human research, in contrast to animal models. Thus, further development of experience-sampling methods has the potential of broadening the scope of findings that can be obtained exclusively in human studies. In order to achieve reliable and valid ratings, subjects need to be explained very clearly what they are to rate. For example, inter-rater reliability was reduced when subjects rated goal-directed movements in movie clips compared with presence of faces in them (Lahnakoski et al., 2012).

Another limitation of self-reports is that they take time. For example, obtaining continuous valence and arousal measures for emotional movie clips shown to subjects during fMRI, the clips had to be played twice to the experimental subjects after the fMRI scanning session, one time for assessment of valence, and another time for assessment of arousal (Nummenmaa et al., 2012). In practice, the amount of time that a single subject can be expected to participate in an experiment is limited to a few hours and thus the number of variables that can be sampled is very limited.

The third limitation of self-reports is that they easily alter the experience of movies and narratives if they are obtained during neuroimaging, and on the other hand they can be less reliable if they are obtained separately after scanning (Hurlburt et al., 2015). In practice, however, it seems re-presenting the stimulus after scanning offers a powerful memory cue for retrospective self-reports of how the stimulus was experienced. This has been empirically shown in case of ratings of the humorousness and certain experienced emotions elicited by movie clips (Raz et al., 2012; Jaaskelainen et al., 2016), however, the assumption needs to be validated for other types of self-reported experiences.

There are other types of measures that can be obtained to either replace or complement verbal self-reports. It is possible to ask subjects to paint body maps of feelings that relate to their emotional experience (Nummenmaa et al., 2013). This can be one way to circumvent potential confounds that arise when asking subjects to verbally categorize their emotions. Besides self-reports, emotions can be estimated from physiological data (i.e., galvanic skin response and heart-rate measures; Wang et al., 2018) and facial muscle activity (Jonghwa and Ande, 2008; Jerritta et al., 2014), respectively. Notably, these measures can be recorded simultaneously with fMRI.

Recording of eye-movements provides yet another complementary behavioral measure. In one study, it was

observed that variations in eye-movements failed to predict variations in brain activity beyond the early visual areas (Lahnakoski et al., 2014). Other studies have noted that visual representations are relatively independent of eye-movements in hierarchically higher visual areas (Lu et al., 2016; Nishimoto et al., 2017). However, there are findings demonstrating a coupling between eye-movements and hippocampal activity (Hannula and Ranganath, 2009). This suggests that eye-movements have the potential of offering indirect information about the inner workings of the human mind. In the context of eye-movement recordings, variations in pupil size can be taken as an additional measure of arousal (Wang et al., 2018). However, this is easily confounded by constantly changing luminance levels in the case of watching movies, which need to be taken into account. Pupil size has been also linked to processing load, yet on this we failed to obtain significant results in one of our studies (Smirnov et al., 2014). Thus, more studies are needed to determine in which ways eye-movement recordings can be useful in enhancing our understanding of how subjects experience media-based stimuli.

NEURO-PHENOMENOLOGY

One potential way to circumvent these caveats and to optimize the synergy between experience sampling and neuroimaging was proposed over two decades ago by Francisco Varela (Varela, 1996; Rodriguez et al., 1999), who coined the term “neuro-phenomenology.” Phenomenology is the philosophical study of the structures of experience and consciousness from the first-person perspective (Callagher and Zahavi, 2020). Varela was inspired by pioneering works of the phenomenologists Merleau-Ponty (1978) and Husserl (1999). Following-up previous attempts to bridge the gap between neural activity of the mind and the subjective experience (e.g., “Hard Problem of Consciousness”) (Chalmers, 1995), Varela reframed the gap by methodologically integrating the two realms, while merging the quantitative statistical power of neuroimaging parameters with the experiential first-person descriptions. However, thus far there has not been any systematic empirical implementation of the neurophenomenological approach (Berkovich-Ohana et al., 2020).

Movies and narratives provide an excellent opportunity to this end: while early neuroimaging tended to rely on dichotomous and over-simplified approaches, the paradigm shift to use media-based stimuli in neuroscience has emphasized the need for a neurophenomenological approach that would accommodate the complex and multifaceted features of natural phenomena. At the theoretical level, phenomenological representations and concepts can accommodate and clarify neural mechanisms and processes, as instantiated in the recent Graded Empathy Framework (Levy and Bader, 2020). This neurophenomenological model of empathy leans on phenomenological analyses revealing parametric leveling of the empathic experience, and these straightforwardly align with neural mechanisms during experiments with movie and narrative stimuli. Accordingly, movies and narratives often evoke multiple

levels of neural processes, and phenomenological analyses can describe these processes from a naturalistic and experiential point of view. In that way, neuro-phenomenology introduces a new dimension to the field of neuroscience – this dimension extends self-reported measures. Thus, neuro-phenomenological approaches may further advance neuroimaging research with media-based stimuli.

MOVING BEYOND SELF-REPORTS?

Another intriguing, yet completely different possibility is emerging for estimating experienced emotions based on distributed brain activity patterns. Specifically, it seems that it is possible to classify, in many instances well above chance level, basic emotions such as anger, fear, sadness, or happiness from fMRI data using multi-voxel pattern based machine-learning algorithms (Kragel and LaBar, 2015; Saarimäki et al., 2016). Recently, this has been demonstrated also in case of social emotions such as pride and love (Saarimäki et al., 2018). In addition, brain activity accompanying sadness differs between two types of sadness-inducing movies, one involving also sympathy and another involving also hate (Raz et al., 2012). Emotion classification can be accomplished also from EEG data at accuracies as high as 80% (Shuang et al., 2016; Yano and Suyama, 2016; Özderem and Polat, 2017). EEG also suffices for estimating emotional valence – i.e., negative vs. positive emotional state (Costa et al., 2006; Zhao et al., 2018). Notably, in a recent study, emotional brain activity states were tracked over time during resting state fMRI. The results suggested transitions between emotional states once every 5–15 s during the resting state (Kragel et al., 2021).

Taken together these results suggest that it is possible in principle to assess emotional states that a given subject experienced based on brain activity that is recorded during viewing of movies and listening to narratives. Intriguingly, while at this point rather tentative leads than readily usable tools, these types of approaches might provide ways to circumvent one of the major challenges in cognitive science by allowing one to sample experiences as they take place without altering its neural underpinnings *via* intrusive prompts for self-reports and without the potential loss of validity *via* asking the subjects afterward. These approaches should be further developed and validated in future studies.

REFERENCES

- Aftanas, L. I., Lotova, N. V., Koshkarov, V. I., and Popov, S. A. (1998). Non-linear dynamical coupling between different brain areas during evoked emotions: an EEG investigation. *Biol. Psychol.* 48, 121–138. doi: 10.1016/s0301-0511(98)00015-5
- Bartels, A., and Zeki, S. (2004). Functional brain mapping during free viewing of natural scenes. *Hum. Brain Mapp.* 21, 75–85. doi: 10.1002/hbm.10153
- Berkovich-Ohana, A., Dor-Ziderman, Y., Trautwein, F.-M., Schweitzer, Y., Nave, O., Fulder, S., et al. (2020). The hitchhiker's guide to neurophenomenology -

CONCLUSION

Having detailed annotation of movies and narratives (e.g., presence of people, social interactions, movements) and the recorded brain activity provides only partially the keys for understanding of the neural mechanisms underlying the emotional and cognitive states elicited by the media-based stimuli. To get the complete picture, the inner experiences of experimental subjects need to be additionally estimated. Historically, there have been two opposite schools of thought, one arguing for the usability of introspection (i.e., self-reports) as a method, and the other arguing that introspection is not useful. Several approaches have been developed that can be utilized to tap on different aspects of inner experiences, such as memorization of events, emotional experiences, thoughts, and mental imagery. It seems that post-scanning self-reports are valid for some classes of experiences, such as experiences of humorfulness and certain emotions. However, it also seems that for some other inner experiences online-sampling is more accurate. In neuroimaging studies, utilizing movies and narratives as stimuli and applying introspective self-reports can capture essential information that helps guide the neuroimaging data analysis. We foresee that further development of such methods will open new important avenues in cognitive neuroscience. Finally, machine learning approaches that allow for classification and tracking of emotional (and possibly other mental) states based on neuroimaging data may have potential for yielding complementary information to that provided by self-reports in the future.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

FUNDING

IJ, VK, and AS were supported by the International Laboratory of Social Neurobiology ICN HSE RF Government Grant Ag. No. 075-15-2019-1930. JA was supported by the National Institutes of Health (Grant Nos. R01DC016915, R01DC016765, and R01DC017991). JL was supported by the Academy of Finland (Grant No. 328674).

the case of studying self boundaries with meditators. *Front. Psychol.* 11:1680. doi: 10.3389/fpsyg.2020.01680

- Bisanz, G. L., LaPorte, R. E., Vesonder, G. T., and Voss, J. F. (1978). On the representation of prose: new dimensions. *J. Verbal. Learn. Verbal. Behav.* 17, 337–357. doi: 10.1016/s0022-5371(78)90219-0
- Block, N. (2005). Two neural correlates of consciousness. *Trends Cogn. Sci.* 9, 46–52.
- Cahill, L., Haier, R. J., Fallon, J., Alkire, M. T., Tang, C., Keator, D., et al. (1996). Amygdala activity at encoding correlated with long-term, free recall of emotional information. *Proc. Natl. Acad. Sci. U.S.A.* 93, 8016–8021. doi: 10.1073/pnas.93.15.8016

- Callagher, S., and Zahavi, D. (2020). *The Phenomenological Mind*, 3rd Edn. Oxfordshire: Taylor and Francis.
- Chalmers, D. J. (1995). Facing up to the problem of consciousness. *J. Conscious. Stud.* 2, 200–219.
- Cohen, S. S., and Parra, L. C. (2016). Memorable audiovisual narratives synchronize sensory and supramodal neural responses. *eNeuro* 3, ENEURO.203–ENEURO.216. doi: 10.1523/ENEURO.0203-16.2016
- Costa, T., Rognoni, E., and Galati, D. (2006). EEG phase synchronization during emotional response to positive and negative film stimuli. *Neurosci. Lett.* 406, 159–164. doi: 10.1016/j.neulet.2006.06.039
- Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., and Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends Cogn. Sci.* 10, 204–211. doi: 10.1016/j.tics.2006.03.007
- Dennis, T. A., and Solomon, B. (2010). Frontal EEG and emotion regulation: electrocortical activity in response to emotional film clips is associated with reduced mood induction and attention interference effects. *Biol. Psychol.* 85, 456–464. doi: 10.1016/j.biopsycho.2010.09.008
- Dmochowski, J. P., Sajda, P., Dias, J., and Parra, L. C. (2012). Correlated components of ongoing EEG point to emotionally laden attention - a possible marker of engagement? *Front. Hum. Neurosci.* 6:112. doi: 10.3389/fnhum.2012.00112
- Gorgolewski, K. J., Lurie, D., Urchs, S., Kipping, J. A., Craddock, R. C., Milham, M. P., et al. (2014). A correspondence between individual differences in the brain's intrinsic functional architecture and the content and form of self-generated thoughts. *PLoS One* 9:e97176. doi: 10.1371/journal.pone.0097176
- Hannula, D., and Ranganath, C. (2009). The eyes have it: hippocampal activity predicts expression of memory in eye movements. *Neuron* 63, 592–599. doi: 10.1016/j.neuron.2009.08.025
- Hasson, U., Furman, O., Clark, D., Dudai, Y., and Davachi, L. (2008). Enhanced intersubject correlations during movie viewing correlate with successful episodic encoding. *Neuron* 57, 452–462. doi: 10.1016/j.neuron.2007.12.009
- Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., and Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science* 303, 1634–1640. doi: 10.1126/science.1089506
- Hsu, C. T., Conrad, M., and Jacobs, A. M. (2014). Fiction feelings in Harry Potter: haemodynamic response in the mid-cingulate cortex correlates with immersive reading experience. *Neuroreport* 25, 1356–1361. doi: 10.1097/WNR.0000000000000272
- Hurlburt, R. T., Alderson-Day, B., Fernyhough, C., and Kuhn, S. (2015). What goes on in the resting-state? A qualitative glimpse into resting-state experience in the scanner. *Front. Psychol.* 6:1535. doi: 10.3389/fpsyg.2015.01535
- Hurlburt, R. T., Alderson-Day, B., Fernyhough, C., and Kuhn, S. (2017). Can inner experience be apprehended in high fidelity? Examining brain activation and experience from multiple perspectives. *Front. Psychol.* 8:43. doi: 10.3389/fpsyg.2017.00043
- Husserl, E. (1999). *The Idea of Phenomenology*. Dordrecht: Springer.
- Hutcherson, C. A., Goldin, P. R., Ochsner, K. N., Gabrieli, J. D., Feldman Barrett, L., and Gross, J. J. (2005). Attention and emotion: does rating emotion alter neural responses to amusing and sad films. *Neuroimage* 27, 656–668. doi: 10.1016/j.neuroimage.2005.04.028
- Jaaskelainen, I. P., Koskentalo, K., Balk, M. H., Autti, T., Kauramäki, J., Pomren, C., et al. (2008). Inter-subject synchronization of prefrontal cortex hemodynamic activity during natural viewing. *Open Neuroimag. J.* 2, 14–19. doi: 10.2174/1874440000802010014
- Jaaskelainen, I. P., Pajula, J., Tohka, J., Lee, H.-J., Kuo, W.-J., and Lin, F.-H. (2016). Brain hemodynamic activity during viewing and re-viewing of comedy movies explained by experienced humor. *Sci. Rep.* 6:27741. doi: 10.1038/srep27741
- Jaaskelainen, I. P., Sams, M., Glerean, E., and Ahveninen, J. (2021). Movies and narratives as naturalistic stimuli in neuroimaging. *Neuroimage* 224:117445. doi: 10.1016/j.neuroimage.2020.117445
- Jacob, Y., Gilam, G., Lin, T., Raz, G., and Hendler, T. (2018). Anger modulates influence hierarchies within and between emotional reactivity and regulation networks. *Front. Behav. Neurosci.* 12:60. doi: 10.3389/fnbeh.2018.00060
- Jerritta, S., Murugappan, M., Wan, K., and Yaacob, S. (2014). Emotion recognition from facial EMG signals using higher order statistics and principal component analysis. *J. Chin. Inst. Chem. Eng.* 37, 385–394. doi: 10.1080/02533839.2013.799946
- Jones, N. A., and Fox, N. A. (1992). Electroencephalogram asymmetry during emotionally evocative films and its relation to positive and negative affectivity. *Brain Cogn.* 20, 280–299. doi: 10.1016/0278-2626(92)90021-d
- Jonghwa, K., and Ande, E. (2008). Emotion recognition based on physiological changes in music listening. *IEEE Trans. Pattern. Anal. Mach. Intell.* 30, 2067–2083. doi: 10.1109/TPAMI.2008.26
- Kauppi, J. P., Jaaskelainen, I. P., Sams, M., and Tohka, J. (2010). Inter-subject correlation of brain hemodynamic responses during watching a movie: localization in space and frequency. *Front. Neuroinform.* 4:5. doi: 10.3389/fninf.2010.00005
- Kragel, P. A., Hariri, A. R., and LaBar, K. S. (2021). The temporal dynamics of spontaneous emotional brain states and their implications for mental health. *J. Cogn. Neurosci.* doi: 10.1162/jocn_a_01787 [Epub Online ahead of print].
- Kragel, P. A., and LaBar, K. S. (2015). Multivariate neural biomarkers of emotional states are categorically distinct. *Soc. Cogn. Affect. Neurosci.* 10, 1437–1448. doi: 10.1093/scan/nsv032
- Kriegeskorte, N., Mur, M., and Bandettini, P. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* 2:4. doi: 10.3389/neuro.06.004.2008
- Lahnakoski, J. M., Glerean, E., Jaaskelainen, I. P., Hyönä, J., Hari, R., Sams, M., et al. (2014). Synchronous brain activity across individuals underlies shared psychological perspectives. *Neuroimage* 100, 316–324. doi: 10.1016/j.neuroimage.2014.06.022
- Lahnakoski, J. M., Glerean, E., Salmi, J., Jaaskelainen, I. P., Sams, M., Hari, R., et al. (2012). Naturalistic fMRI mapping reveals superior temporal sulcus as the hub for distributed brain network for social perception. *Front. Hum. Neurosci.* 6:233. doi: 10.3389/fnhum.2012.00233
- Lankinen, K., Saari, J., Hari, R., and Koskinen, M. (2014). Intersubject consistency of cortical MEG signals during movie viewing. *Neuroimage* 92, 217–224. doi: 10.1016/j.neuroimage.2014.02.004
- Levy, J., and Bader, O. (2020). Graded empathy: a neuro-phenomenological hypothesis. *Front. Psychiatry* 11:554848. doi: 10.3389/fpsyg.2020.554848
- Lu, K. H., Hung, S. C., Wen, H., Marussich, L., and Liu, Z. (2016). Influences of high-level features, gaze, and scene transitions on the reliability of BOLD responses to natural movie stimuli. *PLoS One* 11:e0161797. doi: 10.1371/journal.pone.0161797
- Merleau-Ponty, M. (1978). *Phenomenology of Perception*. London: Routledge.
- Metz-Lutz, M. N., Bressan, Y., Heider, N., and Otzenberger, H. (2010). What physiological changes and cerebral traces tell us about adhesion to fiction during theater-watching? *Front. Hum. Neurosci.* 19:59. doi: 10.3389/fnhum.2010.00059
- Nishimoto, S., Huth, A. G., Bilenko, N. Y., and Gallant, J. L. (2017). Eye movement-invariant representations in the human visual system. *J. Vis.* 17:11. doi: 10.1167/17.1.11
- Nitschke, J. B., Heller, W., Etienne, M. A., and Miller, G. A. (2004). Prefrontal cortex activity differentiates processes affecting memory in depression. *Biol. Psychol.* 67, 125–143. doi: 10.1016/j.biopsycho.2004.03.004
- Nummenmaa, L., Glerean, E., Hari, R., and Hietanen, J. K. (2013). Bodily maps of emotions. *Proc. Natl. Acad. Sci. U.S.A.* 111, 646–651.
- Nummenmaa, L., Glerean, E., Viinikainen, M., Jaaskelainen, I. P., Hari, R., and Sams, M. (2012). Emotions promote social interaction by synchronizing brain activity across individuals. *Proc. Natl. Acad. Sci. U.S.A.* 109, 9599–9604.
- Nummenmaa, L., Saarimäki, H., Glerean, E., Gotsopoulos, A., Jaaskelainen, I. P., Hari, R., et al. (2014). Emotional speech synchronizes brains across listeners and engages large-scale dynamic brain networks. *Neuroimage* 102, 498–509.
- Özderm, M. S., and Polat, H. (2017). Emotion recognition based on EEG features in movie clips with channel selection. *Brain Inform.* 4, 241–252.
- Pajula, J., Kauppi, J. P., and Tohka, J. (2012). Inter-subject correlation in fMRI: method validation against stimulus-model based analysis. *PLoS One* 8:e41196. doi: 10.1371/journal.pone.0041196
- Raz, G., Winetraub, Y., Jacob, Y., Kinreich, S., Maron-Katz, A., Shaham, G., et al. (2012). Portraying emotions at their unfolding: a multilayered approach for probing dynamics of neural networks. *Neuroimage* 60, 1448–1461.
- Rodriguez, E., George, N., Lachaux, J. P., Martinerie, J., Renault, B., and Varela, F. J. (1999). Perception's shadow: long-distance synchronization of human brain activity. *Nature* 397, 430–433.

- Saastti, S., Alho, J., Bar, M., Glerean, E., Honkela, T., Kauppila, M., et al. (2019). Inferior parietal lobule and early visual areas support elicitation of individualized meanings during narrative listening. *Brain Behav.* 9:e01288.
- Saarimäki, H., Ejtehadian, L. F., Glerean, E., Jaaskelainen, I. P., Vuilleumier, P., Sams, M., et al. (2018). Distributed affective space represents multiple emotion categories across the brain. *Soc. Affect. Cogn. Neurosci.* 13, 471–482.
- Saarimäki, H., Gotsopoulos, A., Jaaskelainen, I. P., Lampinen, J., Vuilleumier, P., Hari, R., et al. (2016). Discrete neural signatures of basic emotions. *Cereb. Cortex* 26, 2563–2573.
- Shuang, L., Jingling, T., Xu, M., Yang, J., Qi, H., and Ming, D. (2016). Improve the generalization of emotional classifiers across time by using training samples from different days. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2016, 841–844.
- Smirnov, D., Glerean, E., Lahnakoski, J. M., Salmi, J., Jaaskelainen, I. P., Sams, M., et al. (2014). Fronto-parietal network supports context-dependent speech comprehension. *Neuropsychologia* 63, 293–303.
- Varela, F. J. (1996). Neurophenomenology: a methodological remedy for the hard problem. *J. Conscious. Stud.* 3, 330–349.
- Wang, C.-A., Baird, T., Huang, J., Coutinto, J. D., Brien, D. C., and Munoz, D. P. (2018). Arousal effects on pupil size, heart rate, and skin conductance in an emotional face task. *Front. Neurol.* 9:1029. doi: 10.3389/fneur.2018.01029
- Westermann, R., Spies, K., Stahl, G., and Hesse, F. W. (1996). Relative effectiveness and validity of mood induction procedures: a meta-analysis. *Eur. J. Soc. Psychol.* 26, 557–580.
- Yano, K., and Suyama, T. (2016). A novel fixed low-rank constrained EEG spatial filter estimation with application to movie-induced emotion recognition. *Comput. Intell. Neurosci.* 2016:6734720.
- Zhao, G., Zhang, Y., Ge, Y., Zheng, Y., Sun, X., and Zhang, K. (2018). Asymmetric hemisphere activation in tenderness: evidence from EEG signals. *Sci. Rep.* 8:8029.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Jääskeläinen, Ahveninen, Klucharev, Shestakova and Levy. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Influence of Auditory Cues on the Neuronal Response to Naturalistic Visual Stimuli in a Virtual Reality Setting

George Al Boustani^{1†}, Lennart Jakob Konstantin Weiß^{1†}, Hongwei Li^{2,3}, Svea Marie Meyer¹, Lukas Hiendlmeier¹, Philipp Rinklin¹, Bjoern Menze^{2,3}, Werner Hemmert⁴ and Bernhard Wolfrum^{1*}

¹ Neuroelectronics – Munich Institute of Biomedical Engineering, Department of Electrical and Computer Engineering, Technical University of Munich, Munich, Germany, ² Department of Quantitative Biomedicine, University of Zurich, Zurich, Switzerland, ³ Department of Informatics, Technical University of Munich, Munich, Germany, ⁴ Bio-Inspired Information Processing – Munich Institute of Biomedical Engineering, Department of Electrical and Computer Engineering, Technical University of Munich, Munich, Germany

OPEN ACCESS

Edited by:

Selina C. Wriessneger,
Graz University of Technology, Austria

Reviewed by:

Marta Matamala-Gomez,
University of Milano-Bicocca, Italy
Theerawit Wilaiprasitporn,
Vidyasirimedhi Institute of Science
and Technology, Thailand
Surej Mouli,
Aston University, United Kingdom

*Correspondence:

Bernhard Wolfrum
bernhard.wolfrum@tum.de

[†] These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Brain-Computer Interfaces,
a section of the journal
Frontiers in Human Neuroscience

Received: 04 November 2021

Accepted: 02 May 2022

Published: 02 June 2022

Citation:

Al Boustani G, Weiß LJK, Li H,
Meyer SM, Hiendlmeier L, Rinklin P,
Menze B, Hemmert W and Wolfrum B
(2022) Influence of Auditory Cues on
the Neuronal Response to Naturalistic
Visual Stimuli in a Virtual Reality
Setting.
Front. Hum. Neurosci. 16:809293.
doi: 10.3389/fnhum.2022.809293

Virtual reality environments offer great opportunities to study the performance of brain-computer interfaces (BCIs) in real-world contexts. As real-world stimuli are typically multimodal, their neuronal integration elicits complex response patterns. To investigate the effect of additional auditory cues on the processing of visual information, we used virtual reality to mimic safety-related events in an industrial environment while we concomitantly recorded electroencephalography (EEG) signals. We simulated a box traveling on a conveyor belt system where two types of stimuli – an exploding and a burning box – interrupt regular operation. The recordings from 16 subjects were divided into two subsets, a visual-only and an audio-visual experiment. In the visual-only experiment, the response patterns for both stimuli elicited a similar pattern – a visual evoked potential (VEP) followed by an event-related potential (ERP) over the occipital-parietal lobe. Moreover, we found the perceived severity of the event to be reflected in the signal amplitude. Interestingly, the additional auditory cues had a twofold effect on the previous findings: The P1 component was significantly suppressed in the case of the exploding box stimulus, whereas the N2c showed an enhancement for the burning box stimulus. This result highlights the impact of multisensory integration on the performance of realistic BCI applications. Indeed, we observed alterations in the offline classification accuracy for a detection task based on a mixed feature extraction (variance, power spectral density, and discrete wavelet transform) and a support vector machine classifier. In the case of the explosion, the accuracy slightly decreased by –1.64% p. in an audio-visual experiment compared to the visual-only. Contrarily, the classification accuracy for the burning box increased by 5.58% p. when additional auditory cues were present. Hence, we conclude, that especially in challenging detection tasks, it is favorable to consider the potential of multisensory integration when BCIs are supposed to operate under (multimodal) real-world conditions.

Keywords: brain computer interface, event-related potential (ERP), combinational audio-visual stimulus, visual evoked potential (VEP), virtual reality, support vector machine (SVM)

INTRODUCTION

Neuroscientists aim to understand the human brain by deciphering neuronal signals due to different tasks and stimuli (Adrian and Yamagiwa, 1935; Gross, 1999; Finger, 2001; Strotzer, 2009). Although there are other techniques, most research up to date is based on non-invasive electroencephalography (EEG) recordings, where the electrical activity across the scalp is monitored using distributed electrode arrays (Adrian and Yamagiwa, 1935; Homan et al., 1987; Cincotti et al., 2008; Nicolas-Alonso and Gomez-Gil, 2012). In the past, extensive research focused on unraveling basic neuronal patterns in response to different isolated conditions (Adrian and Yamagiwa, 1935; Penfield and Evans, 1935; Davis et al., 1939; Davis, 1939; Hill, 1958). Thus, an extensive collection of experimental paradigms that evoke specific responses – e.g., event-related potentials (ERPs), steady-state visually evoked potentials (SSVEPs), and motor imaginary related activity, among others – has been established (Ritter et al., 1979; Lines et al., 1984; Alho et al., 1994; Creel, 1995; Comerchero and Polich, 1999; Stige et al., 2007; Sur and Sinha, 2009). Nowadays, applied neuroscientists and engineers use these stimuli–response relations to design brain-computer interfaces (BCIs) that can automatically read out and analyze signals for a specific task. For instance, the P300-speller, a brain-controlled wheelchair, and a brain-controlled prosthetic arm are common BCI applications in the medical context (Rebsamen et al., 2010; Belitski et al., 2011; Nicolas-Alonso and Gomez-Gil, 2012; Abdulkader et al., 2015; Bright et al., 2016). Furthermore, recent technological improvements enable EEG recordings not only under “clean” laboratory conditions but also in natural environments via portable EEG devices. Hence, there is considerable interest in translating BCI applications into more complex real-world settings (Zander and Kothe, 2011). However, in such scenarios, the performance of BCIs and their discriminatory power are drastically affected by interfering signals and physiological artifacts (Fatourehchi et al., 2007; Zander et al., 2010; Minguillon et al., 2017). Here, a combined read-out of multiple cues and/or measurement modalities – a so-called hybrid BCI (hBCI) – addresses this issue by providing an enlarged dataset for classification (Allison et al., 2010; Pfurtscheller et al., 2010; Leeb et al., 2011; Amiri et al., 2013; Yin et al., 2015; Hong and Khan, 2017). For instance, ERPs were combined with motor or mental tasks to design multiple-cue hBCIs (Hong and Khan, 2017). Additionally, parallel recordings from EEG and electrooculography (EOG) or functional near-infrared spectroscopy (fNIRS) were reported to improve performance (Amiri et al., 2013; Hong and Khan, 2017). Consequently, hBCIs offer great potential in various fields, e.g., in diagnostics, rehabilitation, machine control, entertainment, and safety (Allison et al., 2010; Blankertz et al., 2010; Brumberg et al., 2010; Nicolas-Alonso and Gomez-Gil, 2012; Hong and Khan, 2017). Another promising area of application is in the context of industry 4.0, where the aim is to operate factories most efficiently by fusing data streams and monitoring all relevant processes digitally (Douibi et al., 2021).

However, the affiliated classification tasks will be very challenging in most real-world cases depending on the paradigm

and the interfering background signals. Although novel machine learning approaches help to find common patterns, they rely on massive amounts of input data. Here, virtual reality technology (VR) can help to gather consistent training data by simulating natural environments (Holper et al., 2010; Kober and Neuper, 2012; Lotte et al., 2012; Tauscher et al., 2019; Vourvopoulos et al., 2019; Marucci et al., 2021). It has been shown that VR enhances the feeling of presence and provides a real-world experience that keeps the subject more engaged (Kober and Neuper, 2012; Marucci et al., 2021). So far, most EEG-VR studies focused on 3D visual cues, disregarding the effect of simultaneous visual and acoustic stimuli in realistic situations. Previous studies on multimodal audio-visual cues, (Marucci et al., 2021) revealed that the simultaneous neuronal processing of vision and sound is strongly dependent on the exact experiment, determined by the nature, strength, and synchronicity of the stimulus.

This work aims to reveal the effect of additional auditory cues on visually-evoked ERPs within a complex naturalistic scene. To this end, we created an industrial VR environment and designed two visual stimuli that are different in the degree of event severity and stimulus strength. In our experiment, the subject's vision is a conveyor belt-based industrial warehouse, where packages are carried along a unilateral path during regular operation. However, as we target safety applications, in some instances, the regular operation is interrupted by either an exploding or an igniting/burning box.

Since both naturalistic stimuli are visually complex, we first investigate the neuronal response to such visual stimuli and study the effect of perceived severity. Then, we compare our previous findings (visual-only) to a set of experiments, where additional auditory cues match the subject's vision (audio-visual). Lastly, we apply three basic feature extraction methods – variance-, power-spectral-density- and discrete-wavelet-transform-based – to evaluate the effect of additional auditory cues on the classification performance by using a support vector machine (SVM) classifier. Throughout the study, we focused on hardware (24-channel portable EEG) and processing methods suitable for real-world applications.

MATERIALS AND METHODS

Participants

Eighteen subjects (7 females, 11 males) with a mean age of 26 ± 3.4 years participated in this study. Nine subjects were recorded in a visual-only experiment, and nine participated in an audio-visual experiment. To avoid interferences and adaptation, each participant took part only in one of the two experiments. All subjects had normal or corrected to normal vision, normal hearing, no history of neurological diseases, and no previous experience with BCIs or/and EEG recordings. Subjects that exhibited a skin-to-electrode impedance above 10 kOhm across the parietal-occipital lobe electrodes were not considered for further analysis. The study was approved by the Ethics Commission of the Technical University of Munich.

Experimental Setup

The experiments were conducted in a quiet room with a mean sound pressure level (SPL) of 32.1 ± 2.1 dBA (measured with a precision sound analyzer Nor140, Norsonic-Tippkemper GmbH). All subjects were seated comfortably in an idle state in front of a keyboard, see **Figure 1A**. The visual scene and stimuli were designed with Blender v2.81 (The Blender Foundation) and Unity 2018 (Unity Software Inc.) and displayed via an HTC Cosmos virtual reality headset (90 FPS). In the case of an audio-visual experiment, the subjects were facing an active loudspeaker (8020C, GENELEC) placed at a distance of 1 m in front of the subject, as shown in **Figure 1A**.

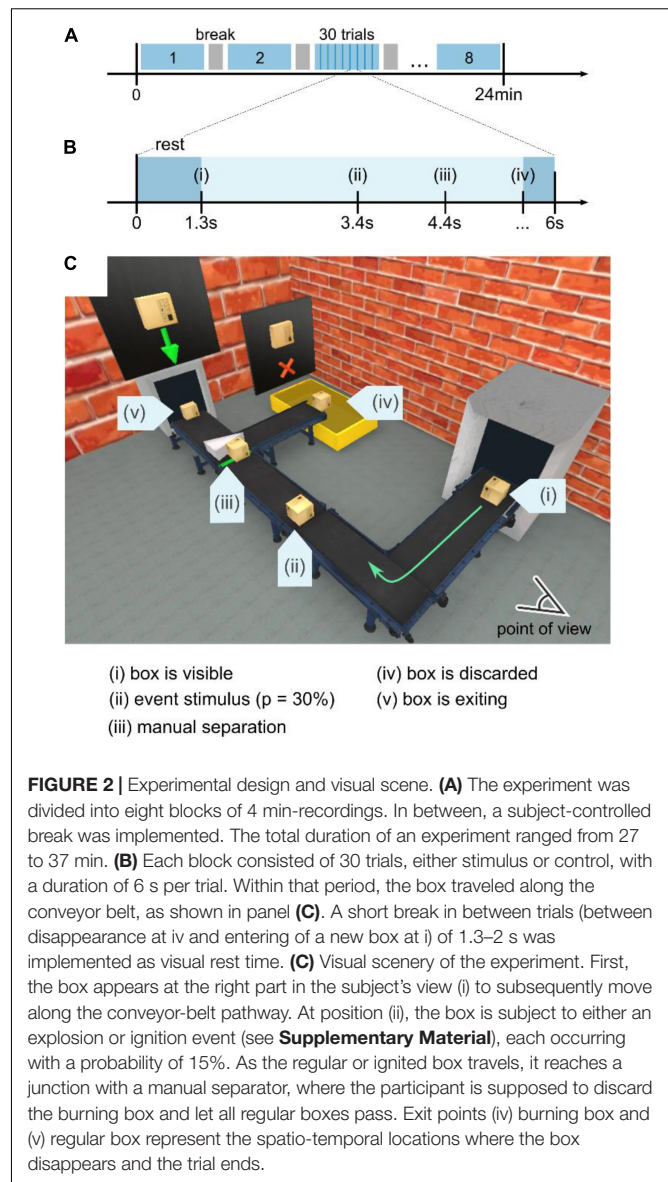
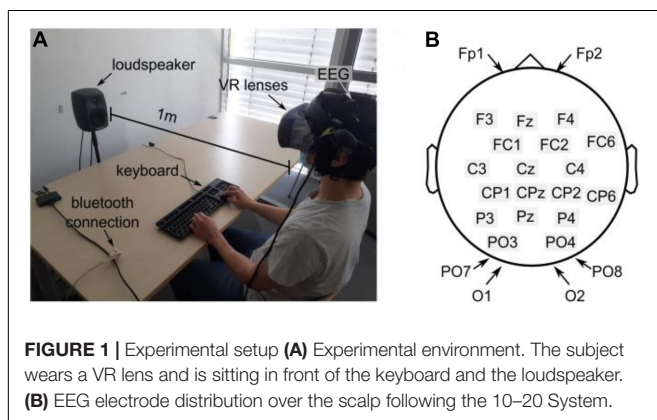
All experiments were recorded using a portable 24-channel EEG system (SMARTING, mbraintrain, Serbia) with a sampling frequency of 250 Hz. The EEG was equipped with passive Ag/AgCl electrodes from EASYCAP (Herrsching, Germany), and a chloride-based electrogel was used (Abralylt HiCl, EASYCAP) to achieve impedance below 10 k Ω . The system featured a reference electrode (common mode sense, CMS) at FCz and a driven right leg electrode (DRL) at Fpz. All electrode locations follow the 10–20 system (see **Figure 1B**) and mainly covered occipital and parietal areas. The electrodes at Fp1 and Fp2 were considered to account for artifacts from eye movements.

Markers that indicate the onset of an (audio-) visual event were streamed from Unity using the lab-streaming layer for Unity asset (LSL4UNITY). Furthermore, all streams were recorded and synchronized using the SMARTING built-in streamer v3.3 for the lab-streaming layer. The data was further processed and analyzed via Matlab (Matlab and Statistics Toolbox Release 2020b, The MathWorks, Inc) combined with the toolboxes EEGLab (Delorme and Makeig, 2004) and fieldtrip (Oostenveld et al., 2011).

Experimental Procedure and Stimulus Design

The study was divided into a visual-only and an audio-visual experiment containing additional auditory cues that matched the visual scene. In both experiments, the stimuli were simulated at the same positions in space and time during the trial. Moreover, the sequence of trials was the same for all subjects.

Each experiment (see **Figure 2A**) consisted of 8 blocks with a break of variable duration in between. Each block contained 30



trials with a fixed duration of 6 s per trial, as shown in **Figure 2B**. In general, three different conditions for the box's pathway were implemented – either the box exploded (a), the box ignited and kept on burning (b), or the box traveled unperturbed along the pathway (c). Regardless of the trial condition, the box initially appeared in the center of the conveyor belt in the right part of the subject's field of view (see (i), **Figure 2C**). Then, the box kept traveling along the conveyor belt for 2 s until it reached point (ii) in **Figure 2C**, where the safety-relevant events occurred with a probability of 33% (equal probability for either a burning or an exploding box) following the oddball paradigm. This probability ultimately leads to 40 stimulus trials for an exploding and 40 stimulus trials for a burning box.

The participants were told to stay seated with a visual point of view, as shown in **Figure 2C**. When a box appeared at point (i), the participant was instructed to track the box along the conveyor

belt visually. Moreover, a short break in between trials (between disappearance at iv and entering of a new box at i) of 1.3–2 s was implemented as visual rest time.

The deviating stimuli were designed to mimic real-world scenarios, consisting of different visual characteristics (e.g., a light flash, change in size and shape). For instance, the explosion (see the video in the **Supplementary Material**) combined a sudden rapid change in light intensity, a swiftly propagating spherical light wave, and a disappearing flying box that occupies the entire field of view. Contrarily, in case of ignition (see the video in the **Supplementary Material**), the box emitted flames of fire from the center of the box. Compared to the explosion, the ignition only partially affected the scenery and started with a slower change in light intensity. While the box was traveling, the fire intensity increased until a steady state was reached.

For the burning and the control condition, the boxes were traveling past position (ii) in **Figure 2C** to reach the manual separator at location (iii) after 1 s. There, the subject had to manually discard the burning box toward the waste container at location (iv) by pressing the right arrow key on the keyboard. A regular box was directed to the exit (v) by pressing the up arrow key. Depending on the discarding speed, the trial duration was ~6 s. Then, the subsequent trial started 1.3–1.5 s after the box had exited the scene at locations (iv) or (v).

In an audio-visual experiment, sounds matching the visual impressions were selected from an open-source library (freesound.org, see **Supplementary Material**). The sound source was attached to the traveling box in the virtual scene. However, reverberations usually stemming from walls were disabled in order to keep the acoustic scene simple. Before each experiment, the loudspeaker was adjusted to match a maximum sound level of 67 ± 0.5 dBA for the explosion and 55 ± 0.3 dBA for the burning box sound, respectively. Both sounds featured fast increasing and slowly decaying characteristics (see **Supplementary Material**). In the case of the burning box, the auditory cue was displayed at a constant level of 50 dBA SPL as long as the box traveled. Additionally, background noise was added to mimic a conveyor belt sound (42 ± 0.1 dBA).

Signal Processing

Eight out of the nine subjects per condition were considered while one of each group was excluded for hardware issues. The following signal processing pipeline is depicted in **Figure 3**. First, bad channels due to non-working electrodes were excluded. Thus, all non-working electrode were removed consistently for all participants. Then, notch filters with 50 and 100 Hz cutoff frequencies were applied to remove line noise and its second harmonic. Similar to other work, (Rozenkrants and Polich, 2008; Wang C. et al., 2012; Putze et al., 2014; Tidoni et al., 2014; Chang, 2018; Guo et al., 2019) the signal was subsequently bandpass-filtered using a low-pass FIR filter with a cutoff frequency of 40 Hz and a high-pass FIR filter with a cutoff frequency of 0.5 Hz. Consequently, all frequencies outside the narrow frequency band, such as slow drifts and high-frequency artifacts, were attenuated (Nicolas-Alonso and Gomez-Gil, 2012; de Cheveigné and Nelken, 2019). A re-referencing step was omitted due to the low number of channels and their

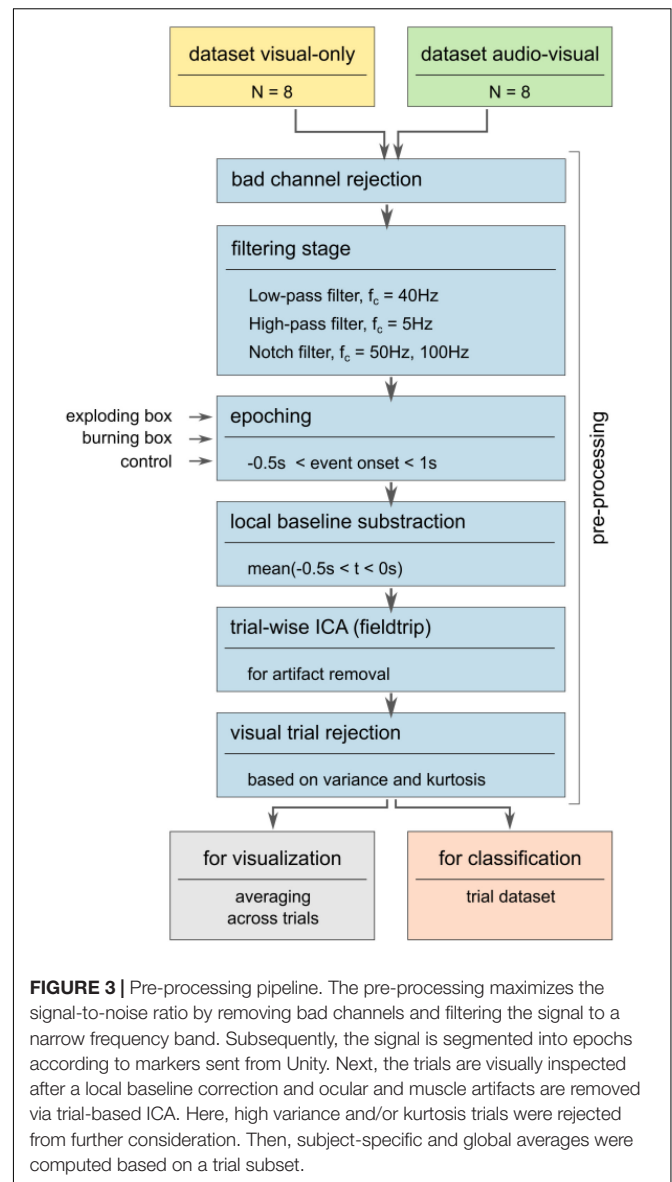


FIGURE 3 | Pre-processing pipeline. The pre-processing maximizes the signal-to-noise ratio by removing bad channels and filtering the signal to a narrow frequency band. Subsequently, the signal is segmented into epochs according to markers sent from Unity. Next, the trials are visually inspected after a local baseline correction and ocular and muscle artifacts are removed via trial-based ICA. Here, high variance and/or kurtosis trials were rejected from further consideration. Then, subject-specific and global averages were computed based on a trial subset.

heterogeneous distribution across the scalp (see **Supplementary Figure 1**). After the filter stage, the recordings were segmented into epochs according to the respective markers sent from Unity at the onset of the stimulus (position (ii) in **Figures 2B,C**). This segmentation resulted in a structural dataset containing all epochs ranging from $t_{(ii)} - 0.5s \leq t \leq t_{(ii)} + 1s$ for all three conditions, explosion (a), burning box (b), and control (c). A local baseline subtraction based on the mean signal before the onset accounted for offset differences. Then, an independent component analysis (ICA) was applied using the logistic infomax approach provided by the fieldtrip toolbox to decompose the signal (Donchin, 1966; Oostenveld et al., 2011; Chang, 2018). Subsequently, the independent components that stem from artifacts such as eye blinking and eye movement, electrode-pops, and muscle movements were visually rejected (Xue et al., 2006; Zhang et al., 2017). Here, the rejected independent

component frequency spectrum and the mixing topographical matrix was inspected to decide which component was identified as an artifact. Lastly, a visual trial rejection removed trials that significantly deviated from the ensemble in terms of variance and/or kurtosis (Oostenveld et al., 2011). In general, the signal processing pipeline was established to maximize the signal-to-noise ratio and, at the same time, to avoid large signal distortions by amplification or attenuation.

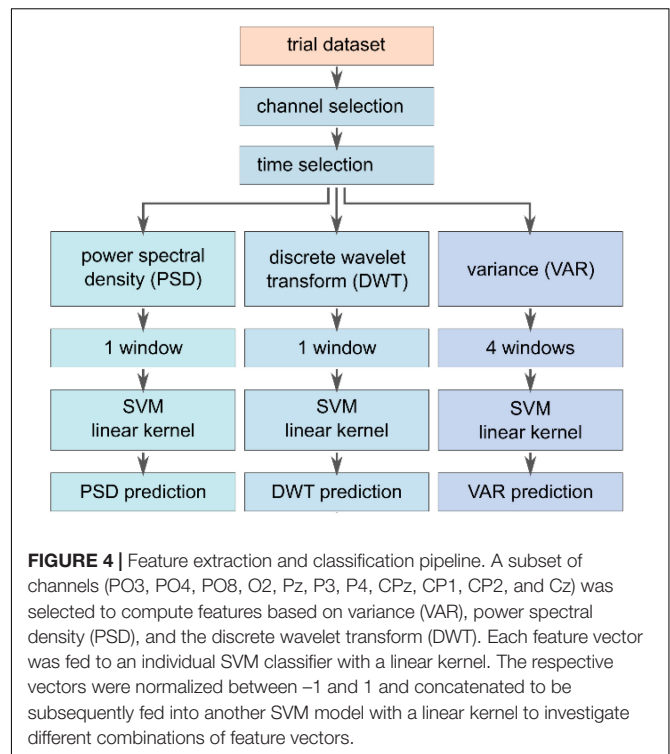
In order to be consistent, subject-specific averages were computed based on 38 out of 40 stimulus trials per subject. Similarly, 38 control trials per subject were randomly selected out of 200 trials. Finally, the global responses shown in the results section were calculated as mean and standard deviation based on the subject-specific characteristics. Hence, the global average indicates the mean neuronal response of the population, whereas the standard deviation visualizes the variability between subjects. Finally, the average control condition was computed based on a random selection of 304 out of 1500 possible trials.

Finally, a statistical analysis on the difference between visual-only vs. audio-visual experiments was performed using a Welch's *t*-test with a 5% significance level. The evaluation is based on the maximum (P1, P3b) and minimum (N2c) for each subject's average (channel O2) and there latencies. The *t*-test assumes that both ensembles are sampled from a normally distributed dataset with unequal variance.

Feature Extraction and Offline Classification

In order to assess the influence of additional auditory cues on the classification performance, different feature extraction methods, see **Figure 4**, – based on the variance (VAR), the power at a specific frequency band (PSD), and specific time-frequency characteristics acquired by a discrete wavelet transform (DWT) – are compared using a SVM classifier. The task of the SVM classifier was to detect the safety-relevant event – explosion (a) or ignition (b) – compared to the control condition (c), where the box was regularly traveling the pathway.

The feature extraction methods were evaluated based on the same dataset that was used for averaging. The feature vectors were computed based on channels covering the parietal and occipital lobe, namely PO3, PO4, PO8, O2, Pz, P3, P4, CPz, CP1, CP2, and Cz. Unfortunately, the channels O1 and PO7 had to be excluded due to inconsistency across subjects. The three methods were applied to the previously selected epochs for averaging with yet a smaller timeframe ranging between $t_{(ii)} \leq t \leq t_{(ii)} + 660$ ms. Each feature extraction method resulted in a dataset of feature vectors, as described in the following. The VAR method computes the variance in four different windows that have been chosen to capture the specific characteristics of the response signal, leading to a 44-element (4 values per channel, 11 channels) feature vector per trial. The first window evaluates the entire epoch from $0 \text{ ms} \leq t \leq 660 \text{ ms}$, whereas the other windows split the entire interval into three successive segments of 220 ms without any overlap. Thereby, the VAR method is supposed to extract information of the entire signal and the variance of early and late potential fluctuations. The PSD feature vector of the trial



was computed using the Welch-method from Matlab. Since we expect stimulus-related frequency information between 1 and 30 Hz,³⁸ all other frequencies outside this window were removed, leading to a vector of length 275 (25 frequency components per channel). The third feature extraction approach, DWT relied on a Matlab discrete wavelet transform decomposition method (Bostanov, 2004; Amin et al., 2015; Cheong et al., 2015; Yahya et al., 2019). In particular, a 3-level decomposition (mother wavelet db8, window size 660 ms) was used to separate the signal in an approximate coefficient vector that extracts low-frequency information and a detail coefficient vector including the high-frequency components. The DWT vector had a length of 341 (31 approximate features per trial). The considered features were normalized and concatenated into a single vector to investigate different feature vector combinations amongst the three approaches. Here, e.g., in the case of the combined VAR-PSD-DWT feature, the vector had a length of 660 elements and ranges between -1 and 1 . Subsequently, the feature vectors were individually fed to a support vector machine classifier with a linear kernel to investigate the different extraction methods (Oskoei et al., 2009; Putze et al., 2014; Li et al., 2018). Here, *k*-fold cross-validation ($k = 10$, 80% training data, 20% testing data) was applied to subject-independent input data stemming from a random selection across the entire dataset. To calculate subject-specific results, an individual SVM classifier for each subject was trained on the combined VAR-PSD-DWT data. Here, similar trial selection and *k*-fold cross-validation approaches were used as mentioned earlier.

Finally, a statistical analysis on the difference between visual-only vs. audio-visual *k*-folds classification results was performed

using a Welch's *t*-test with a 5% significance level. The evaluation is based on the accuracy performance for all folds. The *t*-test assumes that both ensembles are sampled from a normally distributed dataset with unequal variance.

RESULTS

Combined Visual Stimuli

The explosion and the ignition event are implemented as a combination of visual effects, see videos in **Supplementary Material**. Thus, we first want to study the neuronal response to such a combinational visual input. For instance, the explosion was mimicked by an upwards flying box and a bright white spherical wave starting at the box and rapidly propagating through space until the entire field of view is filled. Then, the white flash faded out, the box fell downwards until it disappeared at the floor, and the scene stayed blurry until all smoke had vanished. In total, the entire explosion event lasted ~ 2 s. Hence, we expect the explosion event to be a spatio-temporal mix of different effects leading to an early visually evoked potential (VEP) induced by the flash at the onset and an event-related potential (ERP) in response to the change of the visual scenery. The global responses to the visual-only exploding and burning box are depicted in **Figures 5A,B**, respectively.

As visualized in **Figure 5A** for channel O2, we found deviations at different time instances in the global average response for an explosion compared to the control condition. First, there was a positive rise in amplitude (P1) at O2 in **Figure 5A**, which started at stimulus onset and peaked with 11.5 ± 9.9 a.u. at ~ 125 ms. Then, a negative dip followed, beginning at ~ 200 ms and peaking at ~ 310 ms to -15 ± 6.9 a.u. Subsequently, a smaller positive rise was observed until a plateau of 4.3 ± 2.8 a.u. was reached at ~ 430 ms, which decayed slowly afterward. This finding was robust across trials, as the trial colormaps for a single subject show in **Supplementary Figure 2**. The high standard deviations in the global response, especially for the first peak P1, were caused by the subjects' large variability in terms of latency and amplitude, as depicted in **Supplementary Figure 3**. The first rise in amplitude for O2 was also present at the entire parietal-occipital lobe, but with higher amplitude over the primary visual cortex, see topoplots in **Figure 5A** and the average response for all channels of a single subject in **Supplementary Figure 4**.

In contrast to the explosion, the burning box (see videos in **Supplementary Material**) is designed as a progressive rather than a sudden event. Furthermore, it is modeled as less severe since the flames gradually evolve originating at the traveling box. The burning box stimulus was terminated when the box disappeared in the waste container after discarding. The global response to a burning box is visualized in **Figure 5B**. Here, we find a pattern similar to the explosion – a small P1 between 50 and 100 ms, then a N2c at ~ 280 ms, followed by a P3b at ~ 520 ms.

Additional Acoustic Stimuli

As real-world events naturally lead to a combination of visual and auditory cues, we further investigated the influence of additional

sounds that match the visual experience in the experiment. To this end, background noise (42 dBA SPL) related to the running conveyor belt was implemented. Furthermore, the explosion and the burning box events were synchronized with suitable audio signals (sounds see **Supplementary Material**). Here, we complied with the hierarchical approach and implemented different loudness levels for the explosion and the burning box event. The explosion audio signal had a peak level of ~ 65 dBA and faded slowly toward the conveyor belt noise floor, correlating with the visual impression. The burning box audio stimulus consisted of a transient signal (lighting a match) that reached a steady state of 50 dBA (fire) until the subject discarded the box. Apart from the additional sound, the experiment was the same as previously described. The global responses to the audio-visual exploding and burning box are depicted in **Figures 5C,D**, respectively.

In case of an explosion, five characteristic fluctuations at O2 are visible: a positive peak with ~ 4 a.u. between 70 and 140 ms (P1), two small-amplitude peaks around 220 ms, followed by a prominent negative peak with -13.0 ± 7.1 a.u. at 320 ms (N2c), and a subsequent positive peak with 7.4 ± 5.2 a.u. at ~ 530 ms (P3b). The global response to a burning box with additional auditory cues is shown in **Figure 5D**. Here, three peaks, P1 with 2.6 ± 2.1 a.u. at ~ 80 ms, N2c with -4.4 ± 3.9 a.u. at ~ 330 ms and P3b with 4.9 ± 2.0 a.u. at 550 ms are visible, similar to the fluctuations in **Figure 5B**.

Offline Classification

Since experiments based on virtual reality nowadays offer a great tool to study the applicability of BCIs, we lastly investigate the detectability of events based on visual-only and audio-visual input. This is particularly interesting, as real-world training data is not always easily accessible – especially if the event is rare and/or severe. Moreover, the implementation of multiple modalities in VR settings can be challenging as well. Thus, we aim to evaluate if the classifier that uses bimodal training data is outperforming the classifier based on unimodal input only. To this end, we tested different feature extraction methods – variance-based (VAR), power-spectral density-based (PSD), and discrete-wavelet-transform based – and performed an offline classification using support vector machines on a subject-independent dataset. Here, all subjects' data was merged to randomly select training and cross-validation trials afterward. The VAR method calculates the variance of four different windows containing the response in the P1-, the N2c- and the P3b-part, and the entire epoch as shown in **Figure 5**. The PSD method analyzes the power within the frequency band of 1–30 Hz. In the DWT method, we used a Daubechies mother wavelet to decompose the signal. Additionally, all three methods were combined by concatenation into a single feature vector (DVP) and assessed. The performance of the methods was evaluated with three indicators: (i) the average accuracy across folds indicating the overall model performance, (ii) the average specificity indicating the model performance toward detecting the control condition, and (iii) the average sensitivity that represents the model performance toward detecting the stimulus. The offline detection results are shown in **Table 1**.

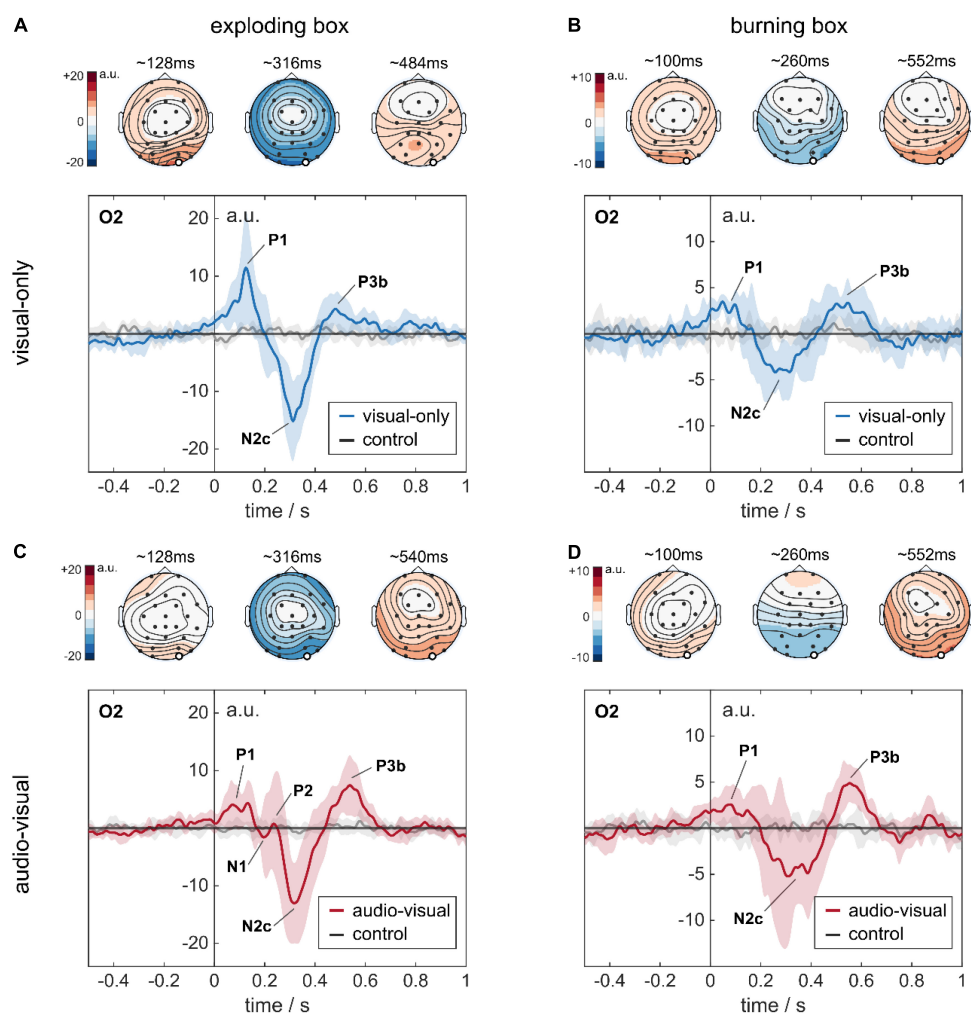


FIGURE 5 | Stimulus-response to complex audio-visual stimuli. All four temporal plots represent the EEG response at the occipital channel O2. The temporal signals are shown as the global average across eight subjects. The standard deviation indicates the variability between the subject-specific average responses. The topoplots represent the global average amplitude distribution across the scalp at three different time points (120, 320, and 540 ms), indicating P1, N2c, and P3b. **(A)** Response to an explosion in a visual-only experiment. **(B)** Response to a burning box in a visual-only experiment. **(C)** Response to an explosion in an audio-visual experiment. **(D)** Response to a burning box in an audio-visual experiment. Note the different y-scale for the exploding and the burning box in the temporal plots.

TABLE 1 | Classification results for the subject-independent dataset.

		Exploding box		Burning box	
		Visual-only	Audio-visual	Visual-only	Audio-visual
Variance method (VAR)	accuracy/%	86.18	87.09	74.01	76.89
	specificity/%	86.5	89.08	79.85	79.27
	sensitivity/%	86.09	85.40	68.39	76.3
Power-spectral density method (PSD)	accuracy/%	82.16	85.58	67.56	76.42
	specificity/%	83.90	87.58	72.6	80.47
	sensitivity/%	80.83	83.02	61.91	72.6
Discrete wavelet transform method (DWT)	accuracy/%	91.16	90.82	78.45	78.73
	specificity/%	90.07	92.36	78.65	79.55
	sensitivity/%	91.70	89.7	79.46	78.66
Feature Fusion Method (DVP)	accuracy/%	94.56	92.92	80.78	86.36
	specificity/%	96.10	94.25	84.22	89.47
	sensitivity/%	92.84	91.71	78.32	84.25

A support vector machine with a linear kernel was used to detect either the exploding or the burning box with respect to the control condition. The results are provided as mean across 10 folds. The Bold values represent the highest achieved accuracy.

DISCUSSION

In the following, we will first discuss the neuronal activity in response to the combinational visual stimuli of an explosion and burning box (see Section “Combined Visual Stimuli”). Afterward, the changes in neuronal activity for experiments with additional auditory cues are presented in Section “Additional Acoustic Stimuli”. Focusing on an industrial BCI application, we lastly compare in Section “Offline Classification” the detectability of an explosion or ignition event based on different feature extraction methods using a support vector machine classifier.

Combined Visual Stimuli

For the explosion box stimulus, we assign this first response P1 to a VEP stemming from a sudden change in light intensity (Connolly and Gruzelić, 1982; Lines et al., 1984; Creel, 1995; Kazai and Yagi, 2003; Sharma et al., 2015; Guo et al., 2019). Further, we associate the negative peak at ~ 310 ms with the N2c component of an ERP-response, as it is distributed across the occipital/posterior region (see **Supplementary Figure 4A**). The N2c component is generally related to visual attention and the processing of stimulus characteristics, which aligns with our expectations of an early primary reaction (P1) and a later activity that reflects the interpretation of the visual scene (N2c and further peaks) (Ritter et al., 1979, 1982; Folstein and Petten, 2008). Lastly, we identify the positive response at ~ 430 ms to be a late P300 signal being evoked by the oddball paradigm. Here, the processing in the visual cortex leads to a delayed response, called P3b, which is usually observed after an N2c component (Comerchero and Polich, 1999; Stige et al., 2007). Consistent with other published work, (Katayama and Polich, 1998; Comerchero and Polich, 1999; Stige et al., 2007; Folstein and Petten, 2008) we observed the P3b component to be higher in the posterior region than in the anterior region of the brain, see topoplots in **Figure 5A** and **Supplementary Figure 4A** as well.

In the burning box stimulus, the absolute amplitudes are notably reduced to a range of approx. ± 5 a.u., reflecting the lower degree of severity and/or lower attention accumulation compared to the explosion. Interestingly, the P1 amplitude for the burning box was in the same range as its N2c-P3b complex, which is in stark contrast to the explosion stimulus, where the P1 was significantly higher than the P3b. This difference might be firstly explained by the gradual increase of fire, secondly by its bounded extent, and thirdly by the red-orange color scheme of the fire animation compared to a full-screen white flash for the explosion.

The high standard deviations for both global responses can be explained by significant differences in amplitude and – even more critical – latencies across individual subjects (see **Supplementary Figures 3A,B**). Here, the response variation might also be affected by adaptation and/or the subjects’ engagement and focus throughout the experiment. In summary, we observed a similar neuronal activity – a combination of an early visually evoked potential (P1) and a delayed event-related potential (N2c-P3b complex) – in response to our virtual explosion and burning stimuli. Here, the degree of severity is reflected in the signal amplitudes, leading to a generally reduced response for the burning box compared to the explosion. Both events, however, showed clearly differentiable global average responses

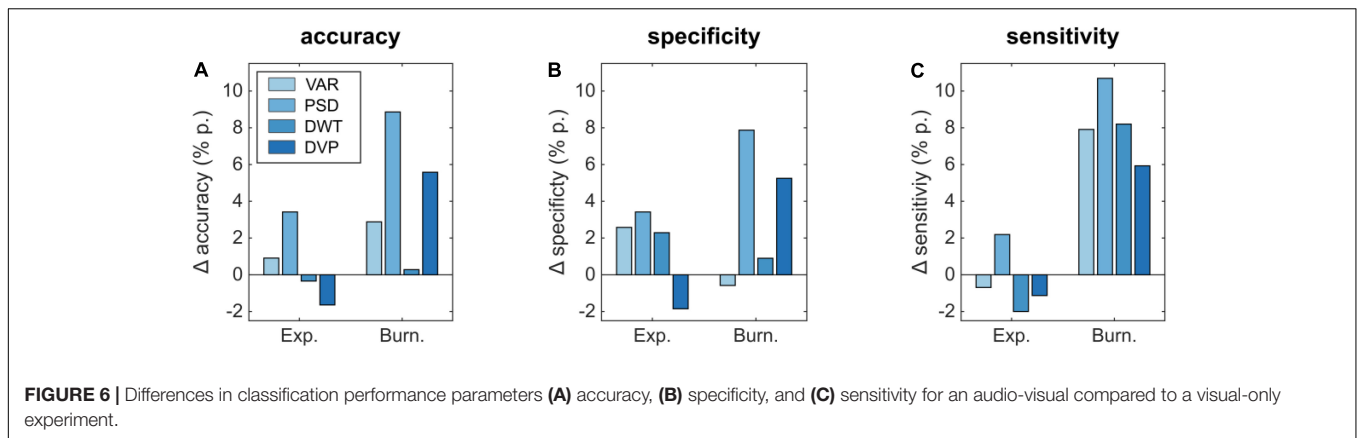
compared to the control condition where the box simply travels along the pathway.

Additional Acoustic Stimuli

In case of an explosion, we find the characteristics of N2c and P3b to be stable, yet their latencies and amplitudes differ (see **Supplementary Figure 5A**). Surprisingly, the VEP P1 is reduced by a factor of ~ 3 , whereas the N2c component is similar in amplitude. The P3b component is delayed by ~ 70 ms and increased by a factor of ~ 2 . Consequently, the additional sound had two effects, the primary visual cue is drastically suppressed, and the ERP components are robust (N2c) or enhanced and delayed (P3b) compared to the visual-only findings. Whereas the suppression of the VEP response P1 are surprising, the ERP enhancement seems plausible, as the additional sound provided congruent supplementary information to the subjects’ visual impression. Furthermore, the enhanced N2c signals could also be attributed to increased attention during the experiment since participants (that took both experiments in an initial pilot study) reported the audio-visual experiment to be more engaging in the burning box stimuli. Lastly, two new fluctuations around 220 ms appeared in the global averages, see **Figure 5C**. Therefore, in line with our hypothesis, the additional small-amplitude peaks could be interpreted as the N1 and P2 components of a strongly enhanced ERP and were not caused by the additional auditory cues. Generally, the N1 and P2 fluctuations of an ERP can be assigned to sensation-seeking behavior, thus reflecting a stronger focus of the participants (Sur and Sinha, 2009). A closer look at the individual responses (**Supplementary Figures 5A, 6A**) reveals the presence of N1 and P2 in 6 out of 8 subjects that participated in an audio-visual experiment. Surprisingly, the additional P2 is in most cases in the same amplitude range as the visually evoked P1 (see **Supplementary Figure 6A**), which is not visible in the global averages due to latency differences across subjects. However, we found N1-P2 components also in the visual-only experiment for some subjects (see **Supplementary Figure 3A**), yet with smaller amplitude compared to an audio-visual experiment. Thus, we conclude that the additional peaks most probably stem from the ERP, which might be altered in amplitude by attention, focus, severity, and congruent input.

The global response to a burning box with additional auditory cues compared to the visual-only experiment, we find the P1 component also to be suppressed by a factor of ~ 1.3 . However, the N2c and the P3b components are again enhanced by a factor of ~ 1.25 and 1.5 , respectively. Additionally, we also observed a delayed ERP response. This result is in line with the previous findings for the explosion. Moreover, large standard deviations around 300 ms indicate the presence of additional small-amplitude peaks as well, which is supported by inspecting the individual responses in **Supplementary Figures 5B, 6B**.

Statistical analysis based on Welch’s *t*-tests revealed a significant amplitude difference in the mean responses for the P1 ($p = 0.0393$) in the case of the explosion stimulus, and further for the N2c ($p = 0.0412$) in case of a burning box at channel O2. However, the P3b component for both conditions did not yield statistical significance as we calculated $p = 0.0764$, $p = 0.0704$ for the exploding box and burning box, respectively. Moreover, all other differences in amplitude and latencies did



not provide statistical significance, which can be also explained by the small dataset and large deviations across subjects (see **Supplementary Table 1**).

In summary, we noticed two different effects on the neuronal responses if additional matching auditory cues were present (see **Supplementary Figure 7**). Firstly, different levels of severity – explosion versus burning box – were again visible as differences in amplitude. Consistently for both stimuli we found the VEP or primary reaction in the visual cortex to be diminished, whereas the ERP components N2c and P3b were enhanced by the sound. Moreover, two other fluctuations, N1-P2, occurred around 220 ms, which we assign to ERP components prior to the large-amplitude peaks N2c and P3b.

Based on our observations, we conclude that additional auditory cues lead to a suppression of the VEP by inhibitory pathways. This was surprising, as we did not expect the sound to induce changes in the early processing stages of primary visual information. However, recent studies shed light on the complex interplay of the neuronal processing of multisensory input (Driver and Spence, 2000; Calvert, 2004; Marchant and Driver, 2013). Indeed, it has been demonstrated that there is “crosstalk” between modality-specific pathways in the associative cortex (Calvert, 2001; Hidaka and Ide, 2015) as well as the primary sensory cortices (Talsma et al., 2007; Senkowski et al., 2011) leading to an early audio-visual integration (Driver and Noesselt, 2008; Wang Y. et al., 2008; Iurilli et al., 2012; Ide and Hidaka, 2013; Hidaka and Ide, 2015). In line with our data, other groups demonstrated e.g., a decreased fNIRS response in the visual cortex (Wiggins and Hartley, 2015) as well as a suppressed visual perception (Hidaka and Ide, 2015) when sound is presented in a spatially and temporally consistent manner. However, we did not only observe the suppression of the primary reaction in the visual cortex (VEP) but also an enhancement and a delay of the following ERP response for additional sound. This could be caused due to differences in the population for both experimental conditions. However, we experienced the phenomenon on single subjects in pilot studies to be robust. In fact, various effects – both, facilitatory and inhibitory – have been reported for multimodal audio-visual input (Stein et al., 1996; Shams et al., 2005; Hidaka et al., 2009; Meredith et al., 2009; Romei et al., 2009; Gleiss and Kayser, 2013).

For instance, it was shown that a multimodal (e.g., visual, acoustic, and tactile) compared to unimodal (visual) stimulation can lead to a drastic enhancement of the P300 signal. (Wang W. Y. et al., 2012; Marucci et al., 2021). Interestingly, an additional delay of the ERP, as visible in our data, was not explored. One could attribute the ERP delay to originate from weak inhibition effects that eventually lead to longer responses (Wang W. Y. et al., 2012). Yet, we found the ERP responses to be more prominent and robust in the audio-visual experiment. Thus, we conclude that a multimodal stimulus leads to an increased certainty about visual perception. Especially in the case of the burning box, where the unimodal visual perception is less clear, the additional (informative) sound supports the understanding and discrimination of the scene (Stein et al., 1996; Talsma et al., 2007; Senkowski et al., 2011; Gleiss and Kayser, 2013).

Offline Classification

Last, we investigated the effect of multimodal stimuli on their classifiability by using offline classification. In this way, we are able to test different extraction methods in a time efficient manner and apply our findings to online classification schemes.

As expected, detecting an explosion is less challenging than detecting a burning box; see absolute values of all criteria in **Table 1**, both in a visual-only and in an audio-visual experiment. Here, we observe significant amplitude differences between the explosion and the burning box responses. The best single-method detection performance for both a visual-only and audio-visual experiment was achieved with the DWT approach (e.g., 91.16 % for an explosion and 78.45 % for the burning box in a visual-only experiment). In contrast, PSD and VAR-based detection performances were substantially lower. This can also be partially explained by correlations between the mother wavelet of the DWT and the neuronal response (Samar et al., 1999). Furthermore, the concatenation of all three feature vectors (DVP) led to an improvement in both conditions (visual-only and audio-visual) for both stimuli compared to DWT. Again, this was partially expected since a larger feature vector can provide more information to the classifier. In the subject specific classification, we achieved an average detection accuracy of 96.06 and 79.96 % for the explosion and burning box, respectively.

The effect of additional auditory cues on the detectability based on different features is shown in **Figure 6**. Here, the accuracy for the explosion (**Figure 6A**) improves by 0.91 and 3.42% p. for VAR and PSD, whereas the DWT and DVP-based performance decreased by -0.34 and -1.64% p., respectively. In case of the burning box, additional auditory cues lead in all cases to an improvement, most prominent for the PSD (8.86% p.) and the combined DVP (5.58% p.). Similarly, the specificity and sensitivity for the burning box are also increased in all but one method, if additional auditory cues are present. In case of the explosion, there is not always an improvement. Mainly if the extraction method relies on the strong P1-contribution in the visual-only experiment (VAR, DWT, and DVP), the performance is slightly decreased in case of additional sound. Similar to ensemble values, we observed a slightly decreased subject-specific classification accuracy (based on DVP) of 94.53% for the audio-visual explosion compared to the visual-only. Again, the burning box led to opposite results. Here, the accuracy increased to 85.18%.

Statistical analysis based on Welch's *t*-tests revealed a significant classification accuracy difference in case of the burning box stimulus for the VAR, PSD, and DVP feature extraction methods ($p = 0.0044$, $p < 0.001$, $p = 0.0050$). Moreover, the exploding box stimulus classification accuracy did not yield statistical significance, which can be explained by the small increase or decrease in performance and the overlapping standard deviation between folds.

The results for the burning box highlight that multimodal input can lead to more robust and enhanced ERP patterns that guarantee an enhanced classification performance. In fact, future real-world detection tasks will resemble most likely the burning box-type of situation, where isolated sensory inputs are less severe, hence, attention-grabbing. Here, the consistent multisensory experience leads to a stronger attentional shift and an increased certainty about the (complex) situation. Consequently, we expect that BCIs trained on multimodal input will show an enhanced classification performance in real-world settings compared to BCIs that consider only unimodal input.

CONCLUSION

Within this work, we studied neuronal responses to two complex stimuli – an exploding box and a burning box – with different perceived severities. The response consisted of a strong early VEP component and a smaller delayed ERP complex in the explosion. The burning box evoked a similar pattern consisting of a minor VEP component and the following ERP complex, but significantly smaller amplitudes than the explosion. Thus, the effect of different severity levels was reflected in the signal amplitudes. Surprisingly, the effect of additional auditory input was not consistent for all response components. Most prominently, for the exploding box, the initial VEP was significantly suppressed in the audio-visual experiment. Moreover, we observed additional small-amplitude peaks around 220 ms after stimulus onset, which we attribute to the early small-scale ERP fluctuations N1 and P2. Hence, we conclude that

congruent multimodal sensory input leads to greater attention and/or a more confident evaluation of the input data, resulting in a robust ERP signal.

In summary, experiments in a virtual environment offer great potential to test the potential of BCIs in different applications. However, stimuli that mimic real-world situations elicit complex neuronal patterns that highly depend on the exact stimulus and environment. As shown in this work, step-by-step VR-EEG studies provide means to bridge the gap from experiments under “clean” lab conditions toward specifically tailored BCI systems. Here, we demonstrated that inhibition and facilitation effects alter the signal for a combined audio-visual input. Based on a SVM classifier, we showed an improvement in the detectability of a bimodal audio-visual stimulus compared to a unimodal visual input. As real-world experiences are multimodal by nature, the early integration of multisensory input has a significant impact on the design of future VR BCI studies.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Commission of the Technical University of Munich. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

GA, LW, PR, and BW designed the study. GA, LW, and SM carried out the experiments. WH and LH helped with the experimental setup. LW and GA wrote the manuscript with support from PR, WH, and BW. All authors provided critical feedback.

FUNDING

The authors greatly appreciate funding from the IuK Förderungsprogramm of the Bavarian State (grant number IUK542/002).

ACKNOWLEDGMENTS

The authors thank J. Mangelberger and S. Hertl for their help and feedback during the study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnhum.2022.809293/full#supplementary-material>

REFERENCES

- Abdulkader, S. N., Atia, A., and Mostafa, M.-S. M. (2015). Brain computer interfacing: applications and challenges. *Egypt. Inform. J.* 16, 213–230. doi: 10.1016/j.eij.2015.06.002
- Adrian, E. D., and Yamagiwa, K. (1935). The origin of the berger rhythm. *Brain* 58, 323–351. doi: 10.1093/brain/58.3.323
- Alho, K., Woods, D. L., and Algazi, A. (1994). Processing of auditory stimuli during auditory and visual attention as revealed by event-related potentials. *Psychophysiology* 31, 469–479. doi: 10.1111/j.1469-8986.1994.tb01050.x
- Allison, B. Z., Brunner, C., Kaiser, V., Müller-Putz, G. R., Neuper, C., and Pfurtscheller, G. (2010). Toward a hybrid brain–computer interface based on imagined movement and visual attention. *J. Neural Eng.* 7:026007. doi: 10.1088/1741-2560/7/2/026007
- Amin, H. U., Malik, A. S., Ahmad, R. F., Badruddin, N., Kamel, N., Hussain, M., et al. (2015). Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques. *Australas. Phys. Eng. Sci. Med.* 38, 139–149. doi: 10.1007/s13246-015-0333-x
- Amiri, S., Fazel-Rezaei, R., and Asadpour, V. (2013). A review of hybrid brain-computer interface systems. *Adv. Hum.-Comp. Int.* 2013:187024. doi: 10.1155/2013/187024
- Belitski, A., Farquhar, J., and Desain, P. (2011). P300 audio-visual speller. *J. Neural Eng.* 8, 025022. doi: 10.1088/1741-2560/8/2/025022
- Blankertz, B., Tangermann, M., Vidaurre, C., Fazli, S., Sannelli, C., Haufe, S., et al. (2010). The Berlin Brain–Computer Interface: non-Medical Uses of BCI Technology. *Front. Neurosci.* 4:198. doi: 10.3389/fnins.2010.00198
- Bostanov, V. (2004). BCI competition 2003-data sets Ib and IIb: feature extraction from event-related brain potentials with the continuous wavelet transform and the t-value scalogram. *IEEE Transac. Biomed. Eng.* 51, 1057–1061. doi: 10.1109/TBME.2004.826702
- Bright, D., Nair, A., Salvekar, D., and Bhisikar, S. (2016). “EEG-based brain controlled prosthetic arm,” in *2016 Conference on Advances in Signal Processing (CASP)*, (Piscataway: IEEE), 479–483. doi: 10.1109/CASP.2016.7746219
- Brumberg, J., Nieto-Castanon, A., Kennedy, P., and Guenther, F. (2010). Brain-Computer Interfaces for Speech Communication. *Speech Commun.* 52, 367–379. doi: 10.1016/j.specom.2010.01.001
- Calvert, G., Spence, C., and Stein, B. E. (2004). *The Handbook of Multisensory Processes*. Cambridge, MA: MIT Press.
- Calvert, G. A. (2001). Crossmodal Processing in the Human Brain: insights from Functional Neuroimaging Studies. *Cereb. Cortex* 11, 1110–1123. doi: 10.1093/cercor/11.12.1110
- Chang, H. I. (2018). *Computational EEG Analysis: Methods and Applications. Biological and Medical Physics, Biomedical Engineering*. Singapore: Springer, doi: 10.1007/978-981-13-0908-3
- Cheong, L. C., Sudirman, R., and Hussin, S. (2015). Feature extraction of EEG signal using wavelet transform for autism classification. *ARN J. Eng. Appl. Sci.* 10, 8533–8540.
- Cincotti, F., Mattia, D., Aloise, F., Bufalari, S., Schalk, G., Oriolo, G., et al. (2008). Non-invasive brain–computer interface system: towards its application as assistive technology. *Brain Res. Bull.* 75, 796–803. doi: 10.1016/j.brainresbull.2008.01.007
- Comerchero, M. D., and Polich, J. (1999). P3a and P3b from typical auditory and visual stimuli. *Clin. Neurophysiol.* 110, 24–30. doi: 10.1016/S0168-5597(98)00033-1
- Connolly, J. F., and Gruzelier, J. H. (1982). Amplitude and Latency Changes in the Visual Evoked Potential to Different Stimulus Intensities. *Psychophysiology* 19, 599–608. doi: 10.1111/j.1469-8986.1982.tb02510.x
- Creel, D. (1995). “Visually Evoked Potentials,” in *Webvision: The Organization of the Retina and Visual System*, eds H. Kolb, E. Fernandez, and R. Nelson (Salt Lake City (UT): University of Utah Health Sciences Center).
- Davis, H., Davis, P. A., Loomis, A. L., Harvey, E. N., and Hobart, G. (1939). Electrical reactions of the human brain to auditory stimulation during sleep. *J. Neurophysiol.* 2, 500–514. doi: 10.1152/jn.1939.2.6.500
- Davis, P. A. (1939). Effects of acoustic stimuli on the waking human brain. *J. Neurophysiol.* 2, 494–499. doi: 10.1152/jn.1939.2.6.494
- de Cheveigné, A., and Nelken, I. (2019). Filters: when, Why, and How (Not) to Use Them. *Neuron* 102, 280–293. doi: 10.1016/j.neuron.2019.02.039
- Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009
- Donchin, E. (1966). A Multivariate Approach to the Analysis of Average Evoked Potentials. *IEEE Transac. Biomed. Eng.* 13, 131–139. doi: 10.1109/TBME.1966.4502423
- Douibi, K., Le Bars, S., Lemontey, A., Nag, L., Balp, R., and Breda, G. (2021). Toward EEG-Based BCI Applications for Industry 4.0: challenges and Possible Applications. *Front. Hum. Neurosci.* 15:456. doi: 10.3389/fnhum.2021.705064
- Driver, J., and Noesselt, T. (2008). Multisensory Interplay Reveals Crossmodal Influences on ‘Sensory-Specific’. *Brain Regions, Neural Responses, and Judgments. Neuron* 57, 11–23. doi: 10.1016/j.neuron.2007.12.013
- Driver, J., and Spence, C. (2000). Multisensory perception: beyond modularity and convergence. *Curr. Biol.* 10, R731–R735. doi: 10.1016/S0960-9822(00)00740-5
- Fatourechi, M., Bashashati, A., Ward, R. K., and Birch, G. E. (2007). EMG and EOG artifacts in brain computer interface systems: a survey. *Clin. Neurophysiol.* 118, 480–494. doi: 10.1016/j.clinph.2006.10.019
- Finger, S. (2001). *Origins of Neuroscience: A History of Explorations Into Brain Function*. Oxford: Oxford University Press.
- Folstein, J. R., and Petten, C. V. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology* 45, 152–170. doi: 10.1111/j.1469-8986.2007.00602.x
- Gleiss, S., and Kayser, C. (2013). Eccentricity dependent auditory enhancement of visual stimulus detection but not discrimination. *Front. Integr. Neurosci.* 7:52. doi: 10.3389/fnint.2013.00052
- Gross, C. G. (1999). *Brain, Vision, Memory: Tales in the History of Neuroscience*. Cambridge: MIT Press.
- Guo, M., Jin, J., Jiao, Y., Wang, X., and Cichocki, A. (2019). Investigation of Visual Stimulus With Various Colors and the Layout for the Oddball Paradigm in Evoked Related Potential-Based Brain–Computer Interface. *Front. Comput. Neurosci.* 13:24. doi: 10.3389/fncom.2019.00024
- Hidaka, S., and Ide, M. (2015). Sound can suppress visual perception. *Sci. Rep.* 5:10483. doi: 10.1038/srep10483
- Hidaka, S., Manaka, Y., Teramoto, W., Sugita, Y., Miyauchi, R., Gyoba, J., et al. (2009). Alternation of Sound Location Induces Visual Motion Perception of a Static Object. *PLoS One* 4:e8188. doi: 10.1371/journal.pone.0008188
- Hill, D. (1958). Value Of The E.E.G in Diagnosis Of Epilepsy. *Brit. Med. J.* 1, 663–666.
- Holper, L., Muehleemann, T., Scholkmann, F., Eng, K., Kiper, D., and Wolf, M. (2010). Testing the potential of a virtual reality neurorehabilitation system during performance of observation, imagery and imitation of motor actions recorded by wireless functional near-infrared spectroscopy (fNIRS). *J. Neuroeng. Rehabil.* 7:57. doi: 10.1186/1743-0003-7-57
- Homan, R. W., Herman, J., and Purdy, P. (1987). Cerebral location of international 10–20 system electrode placement. *Electr. Clin. Neurophysiol.* 66, 376–382. doi: 10.1016/0013-4694(87)90206-9
- Hong, K.-S., and Khan, M. J. (2017). Hybrid Brain–Computer Interface Techniques for Improved Classification Accuracy and Increased Number of Commands: a Review. *Front. Neurobot.* 11:35. doi: 10.3389/fnbot.2017.00035
- Ide, M., and Hidaka, S. (2013). Tactile stimulation can suppress visual perception. *Sci. Rep.* 3:3453. doi: 10.1038/srep03453
- Iurilli, G., Ghezzi, D., Olcese, U., Lassi, G., Nazzaro, C., Tonini, R., et al. (2012). Sound-Driven Synaptic Inhibition in Primary Visual Cortex. *Neuron* 73, 814–828. doi: 10.1016/j.neuron.2011.12.026
- Katayama, J., and Polich, J. (1998). Stimulus context determines P3a and P3b. *Psychophysiology* 35, 23–33. doi: 10.1111/1469-8986.3510023
- Kazai, K., and Yagi, A. (2003). Comparison between the lambda response of eye-fixation-related potentials and the P100 component of pattern-reversal visual evoked potentials. *Cogn. Affect. Behav. Neurosci.* 3, 46–56. doi: 10.3758/CABN.3.1.46

- Kober, S. E., and Neuper, C. (2012). Using auditory event-related EEG potentials to assess presence in virtual reality. *Int. J. Hum.-Comput. Stud.* 70, 577–587. doi: 10.1016/j.ijhcs.2012.03.004
- Leeb, R., Sagha, H., Chavarriaga, R., Millán, J., and del, R. (2011). A hybrid brain-computer interface based on the fusion of electroencephalographic and electromyographic activities. *J. Neural Eng.* 8:025011. doi: 10.1088/1741-2560/8/2/025011
- Li, J., Yu, Z. L., Gu, Z., Wu, W., Li, Y., and Jin, L. (2018). A Hybrid Network for ERP Detection and Analysis Based on Restricted Boltzmann Machine. *IEEE Transac. Neural Syst. Rehab. Eng.* 26, 563–572. doi: 10.1109/TNSRE.2018.2803066
- Lines, C. R., Rugg, M. D., and Milner, A. D. (1984). The effect of stimulus intensity on visual evoked potential estimates of interhemispheric transmission time. *Exp. Brain Res.* 57, 89–98. doi: 10.1007/BF00231135
- Lotte, F., Faller, J., Guger, C., Renard, Y., Pfurtscheller, G., Lécuyer, A., et al. (2012). “Combining BCI with Virtual Reality: Towards New Applications and Improved BCI,” in *Towards Practical Brain-Computer Interfaces*, Chap. Millán, eds B. Allison, S. Dunne, R. Leeb, R. Del, and A. Nijholt (Berlin: Springer), 197–220. doi: 10.1007/978-3-642-29746-5_10
- Marchant, J. L., and Driver, J. (2013). Visual and Audiovisual Effects of Isochronous Timing on Visual Perception and Brain Activity. *Cereb. Cortex* 23, 1290–1298. doi: 10.1093/cercor/bhs095
- Marucci, M., Di Flumeri, G., Borghini, G., Sciaraffa, N., Scandola, M., Pavone, E. F., et al. (2021). The impact of multisensory integration and perceptual load in virtual reality settings on performance, workload and presence. *Sci. Rep.* 11:4831. doi: 10.1038/s41598-021-84196-8
- Meredith, M. A., Allman, B. L., Keniston, L. P., and Clemo, H. R. (2009). Auditory influences on non-auditory cortices. *Hearing Res.* 258, 64–71. doi: 10.1016/j.heares.2009.03.005
- Minguillon, J., Lopez-Gordo, M. A., and Pelayo, F. (2017). Trends in EEG-BCI for daily-life: requirements for artifact removal. *Biomed. Signal Proc. Control* 31, 407–418. doi: 10.1016/j.bspc.2016.09.005
- Nicolas-Alonso, L. F., and Gomez-Gil, J. (2012). Brain Computer Interfaces, a Review. *Sensors* 12, 1211–1279. doi: 10.3390/s120201211
- Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.-M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* 2011:156869. doi: 10.1155/2011/156869
- Oskskoei, M. A., Gan, J. Q., and Hu, H. (2009). “Adaptive schemes applied to online SVM for BCI data classification,” in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, (Piscataway: IEEE), 2600–2603. doi: 10.1109/IEMBS.2009.5335328
- Penfield, W., and Evans, J. (1935). The Frontal Lobe in man: a clinical study of maximum removals. *Brain* 58, 115–133. doi: 10.1093/brain/58.1.115
- Pfurtscheller, G., Allison, B., Bauernfeind, G., Brunner, C., Solis Escalante, T., Scherer, R., et al. (2010). The hybrid BCI. *Front. Neurosci.* 4:30. doi: 10.3389/fnpro.2010.00003
- Putze, F., Hesslinger, S., Tse, C.-Y., Huang, Y., Herff, C., Guan, C., et al. (2014). Hybrid fNIRS-EEG based classification of auditory and visual perception processes. *Front. Neurosci.* 8:373. doi: 10.3389/fnins.2014.00373
- Rebsamen, B., Guan, C., Zhang, H., Wang, C., Teo, C., Ang, M. H., et al. (2010). A Brain Controlled Wheelchair to Navigate in Familiar Environments. *IEEE Transac. Neural Syst. Rehab. Eng.* 18, 590–598. doi: 10.1109/TNSRE.2010.2049862
- Ritter, W., Simson, R., Vaughan, H. G., and Friedman, D. (1979). A brain event related to the making of a sensory discrimination. *Science* 203, 1358–1361. doi: 10.1126/science.424760
- Ritter, W., Simson, R., Vaughan, H. G., and Macht, M. (1982). Manipulation of event-related potential manifestations of information processing stages. *Science* 218, 909–911. doi: 10.1126/science.7134983
- Romei, V., Murray, M. M., Cappe, C., and Thut, G. (2009). Preperceptual and Stimulus-Selective Enhancement of Low-Level Human Visual Cortex Excitability by Sounds. *Curr. Biol.* 19, 1799–1805. doi: 10.1016/j.cub.2009.09.027
- Rozenkrants, B., and Polich, J. (2008). Affective ERP Processing in a Visual Oddball Task: arousal, Valence, and Gender. *Clin. Neurophysiol.* 119, 2260–2265. doi: 10.1016/j.clinph.2008.07.213
- Samar, V. J., Bopardikar, A., Rao, R., and Swartz, K. (1999). Wavelet Analysis of Neuroelectric Waveforms: a Conceptual Tutorial. *Brain Lang.* 66, 7–60. doi: 10.1006/brln.1998.2024
- Senkowski, D., Saint-Amour, D., Höfle, M., and Foxe, J. J. (2011). Multisensory interactions in early evoked brain activity follow the principle of inverse effectiveness. *NeuroImage* 56, 2200–2208. doi: 10.1016/j.neuroimage.2011.03.075
- Shams, L., Iwaki, S., Chawla, A., and Bhattacharya, J. (2005). Early modulation of visual cortex by sound: an MEG study. *Neurosci. Lett.* 378, 76–81. doi: 10.1016/j.neulet.2004.12.035
- Sharma, R., Joshi, S., Singh, K. D., and Kumar, A. (2015). Visual Evoked Potentials: normative Values and Gender Differences. *J. Clin. Diagn. Res.* 9, CC12–CC15. doi: 10.7860/JCDR/2015/12764.6181
- Stein, B. E., London, N., Wilkinson, L. K., and Price, D. D. (1996). Enhancement of Perceived Visual Intensity by Auditory Stimuli: a Psychophysical Analysis. *J. Cogn. Neurosci.* 8, 497–506. doi: 10.1162/jocn.1996.8.6.497
- Stige, S., Fjell, A. M., Smith, L., Lindgren, M., and Walhovd, K. B. (2007). The Development of Visual P3a and P3b. *Dev. Neuropsychol.* 32, 563–584. doi: 10.1080/87565640701361096
- Strotzer, M. (2009). One Century of Brain Mapping Using Brodmann Areas. *Clin. Neurophysiol.* 19, 179–186. doi: 10.1007/s00062-009-9002-3
- Sur, S., and Sinha, V. K. (2009). Event-related potential: an overview. *Ind. Psychiatry J.* 18, 70–73. doi: 10.4103/0972-6748.57865
- Talsma, D., Doty, T. J., and Woldorff, M. G. (2007). Selective Attention and Audiovisual Integration: is Attending to Both Modalities a Prerequisite for Early Integration? *Cereb. Cortex* 17, 679–690. doi: 10.1093/cercor/bhk016
- Tauscher, J.-P., Schottky, F. W., Grogoric, S., Bittner, P. M., Mustafa, M., and Magnor, M. (2019). “Immersive EEG: Evaluating Electroencephalography in Virtual Reality,” in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, (Piscataway: IEEE), 1794–1800. doi: 10.1109/VR.2019.8797858
- Tidoni, E., Gergondet, P., Kheddar, A., and Aglioti, S. M. (2014). Audio-visual feedback improves the BCI performance in the navigational control of a humanoid robot. *Front. Neurobot.* 8:20. doi: 10.3389/fnbot.2014.00020
- Vourvopoulos, A., Pardo, O. M., Lefebvre, S., Neureither, M., Saldana, D., Jahng, E., et al. (2019). Effects of a Brain-Computer Interface With Virtual Reality (VR) Neurofeedback: a Pilot Study in Chronic Stroke Patients. *Front. Hum. Neurosci.* 13:210. doi: 10.3389/fnhum.2019.00210
- Wang, C., Xiong, S., Hu, X., Yao, L., and Zhang, J. (2012). Combining features from ERP components in single-trial EEG for discriminating four-category visual objects. *J. Neural Eng.* 9:056013. doi: 10.1088/1741-2560/9/5/056013
- Wang, W. Y., Hu, L., Valentini, E., Xie, X. B., Cui, H. Y., and Hu, Y. (2012). Dynamic characteristics of multisensory facilitation and inhibition. *Cogn. Neurodyn.* 6, 409–419. doi: 10.1007/s11571-012-9197-x
- Wang, Y., Celebrini, S., Trotter, Y., and Barone, P. (2008). Visuo-auditory interactions in the primary visual cortex of the behaving monkey: electrophysiological evidence. *BMC Neurosci.* 9:79.
- Wiggins, I. M., and Hartley, D. E. H. (2015). A Synchrony-Dependent Influence of Sounds on Activity in Visual Cortex Measured Using Functional Near-Infrared Spectroscopy (fNIRS). *PLoS One* 10:e0122862. doi: 10.1371/journal.pone.0122862
- Xue, Z., Li, J., Li, S., and Wan, B. (2006). Using ICA to Remove Eye Blink and Power Line Artifacts in EEG. in *First International Conference on Innovative Computing*, *Inform. Contr.* 107–110. doi: 10.1109/ICICIC.2006.543
- Yahya, N., Musa, H., Ong, Z. Y., and Elamvazuthi, I. (2019). Classification of Motor Functions from Electroencephalogram (EEG) Signals Based on an Integrated Method Comprised of Common Spatial Pattern and Wavelet Transform Framework. *Sensors* 19:4878. doi: 10.3390/s19224878
- Yin, E., Zeyl, T., Saab, R., Chau, T., Hu, D., and Zhou, Z. (2015). A Hybrid Brain-Computer Interface Based on the Fusion of P300 and SSVEP Scores. *IEEE Transac. Neural Syst. Rehab. Eng.* 23, 693–701. doi: 10.1109/TNSRE.2015.2403270
- Zander, T. O., and Kothé, C. (2011). Towards passive brain-computer interfaces: applying brain-computer interface technology to human-machine systems in general. *J. Neural Eng.* 8:025005. doi: 10.1088/1741-2560/8/2/025005

Zander, T. O., Kothe, C., Jatzev, S., and Gaertner, M. (2010). ““Enhancing Human-Computer Interaction with Input from Active and Passive Brain-Computer Interfaces,”” in *Brain-Computer Interfaces: Applying our Minds to Human-Computer Interaction Human-Computer Interaction Series*, eds D. S. Tan and A. Nijholt (London: Springer), 181–199.

Zhang, S., McIntosh, J., Shadli, S. M., Neo, P. S.-H., Huang, Z., and McNaughton, N. (2017). Removing eye blink artefacts from EEG—A single-channel physiology-based method. *J. Neurosci. Methods* 291, 213–220. doi: 10.1016/j.jneumeth.2017.08.031

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Al Boustani, Weiß, Li, Meyer, Hiendlmeier, Rinklin, Menze, Hemmert and Wolfrum. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



OPEN ACCESS

EDITED BY

Elise Klein,
UMR8240 Laboratoire de Psychologie
du Développement et de l'Éducation
de l'enfant (LaPsyDÉ), France

REVIEWED BY

Shanna Kousaie,
University of Ottawa, Canada
Zhilong Xie,
Jiangxi Normal University, China

*CORRESPONDENCE

Jason Rothman
jason.rothman@uit.no

SPECIALTY SECTION

This article was submitted to
Cognitive Neuroscience,
a section of the journal
Frontiers in Human Neuroscience

RECEIVED 01 April 2022

ACCEPTED 06 July 2022

PUBLISHED 28 July 2022

CITATION

Pereira Soares SM, Prystauka Y,
DeLuca V and Rothman J (2022) Type
of bilingualism conditions individual
differences in the oscillatory dynamics
of inhibitory control.
Front. Hum. Neurosci. 16:910910.
doi: 10.3389/fnhum.2022.910910

COPYRIGHT

© 2022 Pereira Soares, Prystauka,
DeLuca and Rothman. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Type of bilingualism conditions individual differences in the oscillatory dynamics of inhibitory control

Sergio Miguel Pereira Soares^{1,2}, Yanina Prystauka³,
Vincent DeLuca³ and Jason Rothman^{3,4*}

¹Department of Linguistics, University of Konstanz, Konstanz, Germany, ²Language Development Department, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands, ³Department of Language and Culture, UiT the Arctic University of Norway, Tromsø, Norway, ⁴Nebrija Research Center in Cognition, University of Nebrija, Madrid, Spain

The present study uses EEG time-frequency representations (TFRs) with a Flanker task to investigate if and how individual differences in bilingual language experience modulate neurocognitive outcomes (oscillatory dynamics) in two bilingual group types: late bilinguals (L2 learners) and early bilinguals (heritage speakers—HSs). TFRs were computed for both incongruent and congruent trials. The difference between the two (Flanker effect vis-à-vis cognitive interference) was then (1) compared between the HSs and the L2 learners, (2) modeled as a function of individual differences with bilingual experience within each group separately and (3) probed for its potential (a)symmetry between brain and behavioral data. We found no differences at the behavioral and neural levels for the between-groups comparisons. However, oscillatory dynamics (mainly theta increase and alpha suppression) of inhibition and cognitive control were found to be modulated by individual differences in bilingual language experience, albeit distinctly within each bilingual group. While the results indicate adaptations toward differential brain recruitment in line with bilingual language experience variation overall, this does not manifest uniformly. Rather, earlier versus later onset to bilingualism—the bilingual type—seems to constitute an independent qualifier to how individual differences play out.

KEYWORDS

bi-/multilingualism, cognitive control, time-frequency representations (TFRs), brain oscillations, Flanker task

Introduction

Attaining competencies in and managing more than one language in a single mind is complex and dynamic. Because all languages in the mind maintain some level of activation, irrespective of apparent contextual need, there is a ubiquitous demand to manage them (*via* suppression and/or selection, Kroll et al., 2012). This mental juggling is argued to lead to adaptations in domain-general control, where cognitive and language control networks overlap (see e.g., Anderson et al., 2018a). While effects are

not always replicated (see [Lehtonen et al., 2018](#)), studies have shown that, at least under certain conditions, the brain adapts structurally, functionally and chemically to bilingual experience (e.g., [Stocco et al., 2014](#); [Abutalebi and Green, 2016](#); [Weekes et al., 2018](#); [DeLuca et al., 2019a](#); [Pliatsikas, 2019](#); [Grundy, 2020](#); [Pliatsikas et al., 2021](#)). And yet, the study of bilingualism and neurocognition has primarily focused on monolingual vs. bilingual (dichotomous) group comparisons across a variety of domains and tasks (see [Salig et al., 2021](#) for review). While such an approach has led to keen insights into the bilingual mind and brain, it has also resulted in the nature of individual-level variables across bilinguals themselves to not be adequately addressed ([Pliatsikas et al., 2020](#); [Salig et al., 2021](#)).

Over the past few years in particular, theoretical and empirical work has hypothesized and shown the usefulness of measuring and treating dual (multiple) language experiences as continuous variables (e.g., [Luk and Bialystok, 2013](#); [Li et al., 2014](#); [Bialystok, 2017](#); [Grundy et al., 2017](#); [DeLuca et al., 2019b, 2020](#); [Surrain and Luk, 2019](#); [Gullifer and Titone, 2020](#); [Di Pisa and Rothman, 2021](#); [Marian and Hayakawa, 2021](#)). In the real world, the shapes and forms with which exposure, experience and engagement with multiple languages dynamically present themselves over individual lifespans are nearly limitless. Age (of onset and time of testing), context (in and out of community immersion), linguistic proficiency and other seemingly categorizing proxies, while important factors with considerable explanatory coverage, are not the only variables that differentiate aggregates of bilinguals nor the individuals that comprise them (e.g., simultaneous vs. sequential bilinguals, early versus later child/adult bilinguals, special populations like translators and interpreters). Rather, variation in key factors falling within specific language history backgrounds across space and time (quantity and quality of input, intensity of exposure, patterns of language use, language switching, fluctuating dominance, language community size, linguistic social networks and more) delimit individual opportunities for linguistic bilingual engagement. Their equivalence and/or their potential impact on the outcomes we aim to measure cannot be taken for granted ([Leivada et al., 2021](#)).

Acknowledging and dealing empirically with the above reality has manifold consequences. Indeed, several recent models make distinct predictions regarding specific effects for duration and extent of engagement with bilingual experience ([Stocco et al., 2014](#); [Abutalebi and Green, 2016](#); [Grundy et al., 2017](#); [DeLuca et al., 2020](#); [Pliatsikas, 2020](#)). In this light, it seems reasonable to ponder the extent to which some of the discrepancies within the empirical record might be better explained in relation to the (non-)comparability of how important individual-level variables are distributed across participants ([Leivada et al., 2021](#)). If so, it could be simultaneously true that bilingualism affords no effects to the mind/brain under some conditions ([Paap et al., 2015](#)) while translating into considerable ones (along a continuum) given

distinct individual experience profiles. The overarching question and focus of research, thus, shifts to unpacking the constellations of experiences with dual (multiple) languages that give rise to effects ([Grundy, 2020](#)). However, if individual patterns of dual language engagement matter, it cannot be assumed that such patterns align symmetrically such that bilingual type (macro) categorizing factors do not intercede to affect how individual bilingual engagement patterns ultimately manifest. It is with this in mind that the present study is couched and contributory. Specifically, treating bilingualism as a continuous variable, we regress individual measures of bilingual engagement to understand how they might predict on-task neural dynamics while performing a Flanker task. However, bringing together two distinct types of bilinguals—early native (heritage speakers) and late(r) L2 learners—we further ask whether the age of onset of bilingualism interacts differentially with how proxies of bilingual engagement present in individual differences. Foreshadowing our results, while it is the case that individual patterns of dual language engagement predict individual differences in on-task TFRs in both bilingual type groups, the patterns are distinct within each group, suggesting that onset timing and/or overall duration of bilingualism has a moderating effect for how individual differences unfold.

Regarding the potential for experience related brain adaptations, heritage speaker bilingualism provides a unique and understudied test case. Heritage speakers (HSs) are early bilingual native speakers of a minority language (the heritage language) who grow up in a majority language context ([Rothman, 2009](#); [Montrul, 2016](#); [Polinsky, 2018](#)). HSs are typically, although not exclusively, studied and described in an adult state of linguistic knowledge ([Kupisch and Rothman, 2018](#)). Although degree of linguistic proficiency in a heritage language at the individual level varies considerably, partially overlapping with the outcome continuum reported for late(r) second language (L2) acquirers, the degree and contexts of exposure, age of acquisition, patterns of use and language switching, among many other factors definitively distinguish them. In both cases, however, experiential factors are deterministic for individual linguistic variation ([Polinsky and Scontras, 2020](#)). And yet, few studies in the neurocognition of bilingualism have looked specifically at HSs (or at least labeled and differentiated them as such).

Under the hypothesis that individual level engagement with bilingual language experience is ultimately the (most) deterministic factor for (degree of) bilingualism-induced domain-general neurocognitive effects, it is not clear if age-of-acquisition (AoA) should matter. Under what we refer to as the strong version of this hypothesis, AoA would not bring anything independent to bear, at least per se. In other words, there would be no implied potential for a maturational effect on how the brain deals with the mental exercise induced by managing more than one linguistic system. Ultimately, under such an approach, quantity and quality of engagement, if matched across subjects

of different bilingual types, would yield no differences, making them the singular driving forces behind adaptation at the individual level independent of AoA. Alternatively, under what we refer to as the weaker version of this hypothesis, AoA could matter for two partially exclusive reasons: (i) there is some type of maturational effect on how the brain adapts to the same quantity and quality of dual language experience or (ii) there is no maturational effect, but type of bilingualism entails greater and lesser likelihood for intensity of dual language experience itself. This would be the case, if, on average, HSs are more likely to have deeper experience with dual language management than later onset bilinguals. That is, earlier bilinguals might be more likely to have more opportunities for experiences that translate into increased neurocognitive adaptations. For example, on average, HSs might be much more likely to engage (deeply) in activities such as code-switching that are argued to be especially relevant for degree of neurocognitive effects (Green and Wei, 2014; Hofweber et al., 2020). And so, controlling for time of bilingualism—comparing 20-year-old HSs who have been bilingual for 20 years to 35-year-old L2 learner who also have been bilingual for 20 years—does not necessarily entail that AoA alone teases out the relevant factor of potential distinction precisely because over the same time period one or the other type of bilingual might be much more likely to have deeper dual language experience. As a result of either scenario of the weaker hypothesis, HSs could show different patterns as compared to L2 learners but for quite distinct reasons. Given what we now understand as the highly neuroplastic nature of the human brain over the lifespan (Fuchs and Flügge, 2014), we take the null hypothesis to be that there is no maturational effect on how the mind/brain will adapt to the same quantity and quality of dual language experience. However, we leave open the extent to which different bilingual types will, in their aggregates and indeed across the individuals that comprise them, differ in terms of empirical exponents of cognitive tasks likened to bilingualism-induced adaptations. Why? Precisely because it is highly likely that in the most common cases bilingual type will matter for quantity and quality of opportunities to engage (and thus change) the relevant underlying cognitive systems.

Despite the above provisos, it is likely, however, that HSs have been included in the aggregate young adult bilingual groups in published neurocognitive studies, collapsed with other types of non-native bilinguals. This is potentially problematic to the extent that age-of-acquisition (AoA) could itself be deterministic above and beyond individual differences with bilingual experiences or serve as a qualifier for how they manifest. By separating these two types of bilinguals, with sufficient numbers in each group, our aim is to directly address this possibility. HSs are not only more likely to have younger AoA, but their exposure to and engagement with the two languages are likely to be on average qualitatively distinct from other bilingual types while also showing a greater range of within

group variation (Pascual y Cabo and Rothman, 2012; Kupisch and Rothman, 2018). HSs in the European context at least, where English is ubiquitously taught at early ages and for which populations tend to have good competence when tested in young adulthood, are often multilinguals in the relevant sense in higher proportions than sequential L2 bilinguals. Let us imagine a potential scenario in the German context to understand why this should matter. A study might endeavor to examine the effects of bilingualism on property/domain X with young adult German L1 to English as an L2. If German monolingual learners of L2 English are combined with heritage speakers of Turkish—of which there are millions—also native L1 speakers and dominant in German to form the so-called bilingual group, deterministic variables that make them distinct are being overlooked or ignored. This reality, which we surmise is more common in practice than one might expect, would render relevant combined groupings an amalgamation of bilinguals and multilinguals. We do not (yet) know what differences, if any, juggling three or more languages confers above and beyond two. Moreover, the contexts and frequencies of how languages distribute in these bilinguals and multilinguals are not only likely to be quantitatively, but also qualitatively distinct. Collapsing them might be adding noise to the proverbial signal research endeavors to isolate. The present study, thus, takes these above provisos most seriously. This is not only one of the first studies to capitalize on and separate out HSs as a distinct group in the neurocognitive study of bilingualism, but we also examine HSs alongside a separate group of true additive L2 bilinguals. In our view, such an approach is not only beneficial to address our primary research questions, it is also in line with calls in bilingualism literatures for alternative group contrasts to sidestep the potential for a monolingual-to-bilingual comparative fallacy (e.g. Ortega, 2013; DeLuca et al., 2019b). Doing so enables us to investigate the relevant contribution and weighting of experience related variables specific to bilingualism and multilingualism and macro group level factors such as AoA, duration of being bilingual/multilingual and exposure/usage patterns in isolation as well as how they might interact with each other.

As mentioned above, bilingual experience has been shown to affect domain general cognitive control processes. The Flanker task, which we use herein, is a commonly used task to examine cognitive control including studies examining effects of bilingualism and the neural underpinnings thereof (see e.g., Van den Noort et al., 2019). However, to our knowledge, very few studies to date have examined the effects of bilingualism on the neural oscillatory dynamics related to inhibitory control (e.g., Calvo and Bialystok, 2021). Differently from the more commonly used EEG method of event related potentials (ERP) (see for reviews Grundy et al., 2017; Cespón and Carreiras, 2020), the analysis of time-frequency representations (TFR) decomposes the EEG signal into the frequency domain. This can be done at the pre-stimulus time window, as well as post-stimulus time of interest. This oscillatory activity is generated

from groups of neurons that synchronize and fire together tuned to internal and external stimuli (Engel et al., 2001; Buzsaki, 2006; Cohen, 2017). Oscillations are quantified by measuring frequency-specific power from the EEG/MEG signal (Varela et al., 2001). In humans, five frequency bands have been defined based on their initial clinical relevance: delta (1–2 Hz), theta (3–7 Hz), alpha (8–12 Hz), beta (13–30 Hz), and gamma (30–200 Hz) (Buzsaki, 2006). Decades of research in cognitive neuroscience have highlighted the role of rhythmic periodic brain activity in coordinating large-scale cortical networks to sustain cognitive processing and enable humans to pursue goal-directed behavior (Thut et al., 2012). Changes in environmental conditions need to be efficiently picked up by the executive functions system in order to allocate more attention and cognitive resources to a selected task, while at the same time being able to suppress distracting information (Botvinick et al., 2004). Conflict management and suppression of task-irrelevant information are typically related to top-down executive control mechanisms in the prefrontal cortex (PFC), particularly in the anterior cingulate cortex (ACC) (Botvinick et al., 2001; Fan et al., 2003; Helfrich and Knight, 2016; Hachahmet et al., 2021).

Consistently, EEG and magnetoencephalography (MEG) studies have linked power changes in prefrontal theta activity to temporal reorganization of neural networks coinciding with decision points, i.e., action monitoring and selection (Cavanagh et al., 2012). Similarly, a considerable amount of work has found midfrontal theta power and phase synchronization in frontal electrode clusters (ACC and PFC) associated with stimulus conflict detection and response monitoring (Pastötter et al., 2013; Cavanagh and Frank, 2014; Oehrns et al., 2014; Duprez et al., 2018; Brunetti et al., 2019; Pscherer et al., 2021). Although frontal theta appears to be the optimal candidate for regulating and modulating executive control functions, it is not the only one. Current research has demonstrated how the collective participation of multiple frequency bands (hence several underlying cognitive regulating phenomena) more reliably underpins preparation and implementation of cognitive control (Cooper et al., 2016). Specifically, delta has been found to regulate rule implementation and alpha motor response and anticipatory updating mechanisms (Cooper et al., 2016). Alpha has also been associated with the regulation of pre-stimulus proactive control mediated by the superior frontal cortex (SFC) (Freunberger et al., 2008; Suzuki et al., 2018). Furthermore, increases in alpha power are a proxy for the gating of contributions from task-irrelevant cortical areas (Jensen and Mazaheri, 2010; Jensen et al., 2012). Finally, beta has been associated with several higher cognitive processes, among others top-down selective attention (Siegel et al., 2011).

Studies in bilingual neurocognitive adaptations so far have mostly been based on functional magnetic resonance imaging ((f)MRI). Generally, dual language use leads to brain changes both anatomically and functionally (Del Maschio and Abutalebi, 2019; Pliatsikas, 2019), especially in older populations (Bialystok et al., 2012; Gallo et al., 2020). Furthermore, individual

language experiences differentially affect these brain adaptations (DeLuca et al., 2019b; Pliatsikas et al., 2020; Gallo et al., 2021). Related literature using EEG is limited and (almost) exclusively looks into related ERPs signatures (see Grundy et al., 2017; Cespón and Carreiras, 2020 for review). The strength and advantage of adopting EEG over (f)MRI is that it accommodates for the investigation of rich neural processing at the millisecond resolution level and can provide crucial information on both strength and timing of cognitive processes (Luck and Kappenman, 2011). In complement to ERPs, TFR is a welcome approach because it allows the capturing of multiple simultaneously occurring cognitive processes (Bastiaansen et al., 2011; Prystauka and Lewis, 2019).

While the use of oscillations to examine the neural underpinnings of language comprehension and sentence processing is increasing (see Prystauka and Lewis, 2019 for a review), this method has not been widely applied to look at potential bilingualism-induced effects on domain-general cognition. A series of resting-state EEG (rs-EEG) (a measure of the ongoing brain signal in a task-free context) studies have investigated immersive computerized (second) language learning paradigms (Prat et al., 2016, 2019). Results reveal correlations between learning outcomes and resting-state low-, mid- and high-beta power (Prat et al., 2016) and functional connectivity across all frequencies correlating with posttest memory retention and speech variance during learning (Prat et al., 2019). Although these two studies are done in the context of bilingual language learning, they highlight how oscillatory dynamics can be profitably employed in bilingualism research more generally to better understand underlying processes of neural computation serving language related functions.

Bice et al. (2020) were the first to apply neural oscillations with the intention to investigate if underlying functional brain adaptations can be predicted by bilingual experiences. Comparing rs-EEG data, they found that bilinguals exhibited higher alpha power and coherence in the alpha and beta frequency ranges over monolinguals, positively correlating with language background measures. Similarly, Pereira Soares et al. (2021) correlated rs-EEG with bilingual experience measures to investigate if and how determinants of bilingualism reshape the mind/brain. The findings revealed modulatory effects of age of second language acquisition on high beta and gamma power, whereas higher degree of use of the second language at home and in society contexts correlated with functional connectivity (mean coherence) in theta, alpha and gamma frequencies. In summary, the results of these two studies highlight the modulatory role of brain oscillations (measured in a task free context) in dual language scenarios and underline how these effects vary by degree of individual engagement with bilingualism experience factors over the lifespan.

The present study tests three hypotheses guided by the recent Unified Bilingual Experience Trajectories (UBET; DeLuca et al., 2020) framework. UBET makes detailed predictions about how four general components of bilingual

experience (intensity/diversity of use, language switching, relative proficiency, and duration of use) would variably drive adaptations in cognitive control and its neural underpinnings. As pertains specifically to adaptations in oscillatory dynamics, several of these predictions from the UBET framework are key: increased intensity and diversity of bilingual experience would positively correlate to theta power (in situations of cognitive control). This theta increase is predicted to stem from increased fronto-cortical recruitment to handle the higher control demands associated with this experience. Alternatively, prolonged duration of bilingual experience would relate to increased alpha suppression/desynchronization in situations requiring cognitive control, and decreased theta activity. This shift in oscillatory dynamics is related to a transition toward increased efficiency and automation of handling these existing control demands. Crucially, intensity of use would have an additive effect in this transition, that is, increased intensity of exposure would modulate the latency by which adaptations to efficiency manifest neurophysiologically. Accordingly, our predictions were as follows:

- (1) *Bilingual group types differences*: power differences at the group level between HSs and L2 learners, specifically increased theta activation for L2 learners and increased alpha suppression for HSs for interference suppression given the difference in duration commensurable with bilingual type.
- (2) *Individual differences predicted by engagement patterns*: collapsing the groups and regressing dual language engagement (i.e., how much and in what contexts they use both languages) as continuous variables, one might find correlations between power in theta (synchronization reflected in frontal electrode clusters) reflecting functional adaptations to increased control demands and alpha (suppression) with engagement above and beyond bilingual type specific effects, reflecting increased efficiency in handling control demands

and/or

If intensity over duration of engagement has an additive effect, we might expect different patterns of individual differences between HSs and L2 learners given inherent differences regarding language usage distribution. For example, one could expect all HSs to use a non-societal language at home (potentially the only context in which they use their heritage language). Thus, high use at home would not necessarily signal overall intensity at an equivalent level an equal score would for an L2 learner. In the latter case only, since the non-societal language is not expected in that context, a high score in home use is likely to denote a rather high level of intensity overall.

- (3) *Correspondence between neural dynamics and behavioral performance*: theta activation is predicted to correlate with faster reaction times in the Flanker paradigm. Alpha suppression is predicted to show a greater dissociation with reaction times, especially with prolonged duration of bilingual engagement, signifying increased efficiency of handling control demands.

Materials and methods

Participants

Data were collected from 60 bi-/multilingual participants (43 female), of which 32 were L2 learners outside of immersion (English in Germany) and 28 were early bilinguals (heritage speakers of Italian in Germany). The L2 learners spoke German as L1 and English as L2, whereas the HSs had Italian as their L1 and either acquired German simultaneously as their second L1 (2L1) or had acquired German from a very young age in Germany, below 4 (Meisel, 2011). The age that participants were exposed for the first time to bilingualism (the L2/2L1) (mean AoA for L2 = 9.4y; SD = 1.98y, mean AoA for 2L1 = 1.88y; SD = 1.7y) and their crucially contexts of bilingualism (immersion or not) differed, but age at time of testing did not (mean age for L2 = 24.65y; SD = 3.59y, mean age for 2L1 = 24.57y; SD = 3.41y). Although the majority of our participants were first exposed to the other language(s) at a young age, this does not exclude that meaningful variation afforded by the context of each individual's bilingual language use is washed out. Given the status of the L2/L3 (English) as *lingua franca*, timing of first exposure can be misleading, i.e., quantity, quality and intensity of exposure and use can vary even at such an early age especially outside of a native English immersion context. On the other hand, heritage bilingualism is, by definition, naturalistic. Provided that the home language is in active use in its domains, heritage bilingualism places, at least at young ages, individuals in a context of immersive opportunities with ample exposure and diversity of active use of the two languages. Especially over development/maintenance through young adulthood, however, it presents a plethora of mitigating circumstances for interindividual variation in linguistic proficiency outcomes and language engagement (e.g., Kupisch and Rothman, 2018; Polinsky, 2018). Notwithstanding differences from monolingual baselines, it is generally accepted that HS grammars are not only native and naturalistic, but also comprehensive, coherent, and universally compliant with natural language (Pascual y Cabo and Rothman, 2012; Rothman and Treffers-Daller, 2014; Lohndal et al., 2019; Polinsky and Scontras, 2020). And so, regardless of the reported early first exposure to English of the L2 group, there are considerable and important differences to the "earliness" of HSs linguistic onset and temporal exposure/engagement to bilingualism.

Since all participants are sampled from the same context in Germany where English is taught pervasively and early on, the heritage speakers have also been exposed to English, making them multilinguals. Their proficiency in English, in fact, did not differ from that of the L2 group. This is unsurprising since their trajectories with English over time are not expected to be, and were indeed not, different in the aggregate from the L2 learners in our sample. Thus, while all are at least bilingual, what distinguishes our groups is their native bilingualism with Italian (or not). In addition to more fine-grained measurements of linguistic exposure and engagement, described separately below, Socio-Economic Status (SES) was coded, from 0 to 4, based on the participant's mother's highest level of education (0 = lower than a high school diploma, 4 = postgraduate degree). The mean SES was 1.15 (SD = 1.22), and there were no differences across the groups ($t(55) = 1.99, p = 0.052$).

Background measures

Participants completed the Language and Social Background Questionnaire (LSBQ) (Anderson et al., 2018b), which documents language exposure and use throughout the lifetime in a wide range of settings and activities. The LSBQ factor calculator provides three different (weighted) composite scores derived from a subset of relevant questions: language use in the home environment (Home), language use in social contexts (Social), and language proficiency in the societal majority language (Proficiency). Regarding both Home and Social factors, the higher the score is, the greater the engagement in the non-societal language is. Alternatively, a lower score indicates more use and exposure with the societal language in a given context. As for Proficiency, higher scores reflect increased proficiency in the societal majority language (in our case German). Additionally, age-of-(onset) acquisition of the non-societal language (AoA) and length of exposure to the non-societal language (LoE) were also measured. We observed a mean score of 10.46 for Social (L2 learners = 9.84, SD = 8.69; HSs = 11.16; SD = 9.13), a mean score of 1.59 for Home (L2 learners = -5.66, SD = 3.01; HSs = 9.86; SD = 5.77), and a mean score of 0.71 (L2 learners = 0.70, SD = 1.88; HSs = 0.72; SD = 1.51) for Proficiency. Participants also completed the LexTALE (Lemhöfer and Broersma, 2012) to assess general English proficiency (L2 learners = 68.75%, SD = 12.40; HSs = 65.27%, SD = 12.46) (see [Supplementary Table 1](#) for all participants' metadata), for which the groups did not differ [$t(57) = 1.08, p = 0.28$].

Study procedure

The research procedures in this study were approved by the University of Konstanz Research Ethics Committee.

Before taking part in the experiment, participants gave written informed consent and confirmed no contraindication to the EEG investigation. Participants who presented a neurological condition (e.g. epilepsy, multiple sclerosis, etc.) were excluded from this study. Furthermore, participants were compensated for their time. First, participants completed the LSBQ. Then, for the EEG recording session, participants were fitted with an appropriate actiCap in accordance with the 10–20 system (Brain Products, Inc). The experiment was presented on a 17-inch screen. The Flanker task (Eriksen and Eriksen, 1974) was administered using Presentation software (Presentation®, Neurobehavioral Systems). Participants were presented with sets of five arrows and asked to specify the direction of the arrow in the middle. Surrounding arrows (flankers) were pointing either in the same direction (congruent condition: <<<<<) or in a different direction (incongruent condition: <<><<). Incongruent trials require participants to ignore the conflicting information coming from the surrounding arrows. Thus, comparatively, incongruent conditions involve greater recruitment of the executive control system.

The task procedure was first explained to the participants. A brief practice session preceded the experimental blocks to allow the participants to familiarize themselves with the task. The practice session consisted of a total of 12 trials, 4 congruent trials, 4 incongruent trials and 4 neutral arrows (only a single arrow was shown, pointing either left or right, this condition was not used in the EEG analysis). Participants were instructed to use their corresponding index finger to press left (on a button box—RB-740, Cedrus®) if the central arrow (or the single arrow in the neutral trials) pointed to the left, and to press right if the opposite was true. Participants were instructed to be as quick and accurate as possible. The experimental session consisted of two blocks of 120 trials each presented in a randomized order: 40 incongruent trials, 40 congruent trials and 40 neutral trials. Thus, the total number of trials was 240 (80 trials per condition). Each trial began with a fixation cross in the middle of the screen presented for a jittered duration of 400–1,600 ms (at randomized steps of 100 ms). The fixation cross allowed the participants both to focus their attention on the center of the screen and to reduce saccadic eye movements. Afterwards, a 200 ms baseline blank screen appeared followed by the stimulus, which was presented until the participants responded or for a maximum duration of 1,500 ms. An inter-trial interval (ITI) blank screen of 2,000 ms followed to avoid eventual carryover effects. The participants were instructed to take a short break between trial blocks. The EEG signal was continuously recorded from 32 Ag/AgCl scalp electrodes (LiveAmp32, Brain Products, Inc). AFz acted as the ground electrode and FCz as the online reference. The Fp1 and Fp2 electrodes, located on the forehead above the eyebrows, were employed to detect and monitor vertical and horizontal eye movements. Impedances were kept below 25 kΩs. Data were recorded with an online filter of 0.01–200 Hz and was amplified

TABLE 1 Summary of the behavioral analysis (mean Reaction Times and Accuracy) for the L2 learners ($n = 32$) and HSs ($n = 28$) for the three different conditions (congruent, incongruent, and neutral) in the Flanker task.

Condition	Group	meanRT (ms)	sdRT (ms)	meanAcc	sdAcc
Congruent	L2 learners	450	81.6	0.998	0.048
	HSs	436	89.1	0.998	0.042
Incongruent	L2 learners	560	123	0.984	0.124
	HSs	561	126	0.971	0.167
Neutral	L2 learners	434	80.2	0.994	0.079
	HSs	418	80.8	0.996	0.063

and continuously digitized at a 1000 Hz sampling rate using a Brain Vision LiveAmp amplifier.

Data pre-processing and time-frequency analysis

Offline processing of the data was done in two steps. First, in Brain Vision Analyzer (BVA) 2.0 (Brain Products, Inc), data were band-pass filtered from 0.1–45 Hz¹. Then, the signal was segmented from –750 to 1,250 ms around the stimulus onset. Next, an automatic independent component analysis (ICA) implemented in BVA was used to detect and eliminate eye movements and blinks. ICA was performed on the segmented dataset with 512 steps and an infomax (Gradient) restricted algorithm. A spheric spline topographic interpolation was employed if anything unusual (e.g., high noise, electrode picking up the heartbeat signal, etc.) happened to the electrodes during the recording (1 to maximal 3 electrodes per participant - total number of interpolated electrodes = 0.01% of the dataset). All trials were manually inspected for artifacts (drifts, excessive muscle artifact, blocking, etc.). Trial rejection resulted in the exclusion of 226 trials (0.016% of the data). The remaining epochs were baseline-corrected (-100 ms prior to stimulus onset) and then re-referenced to the averaged mastoids (TP9/10). Data were then exported for time frequency analysis using the Fieldtrip toolbox implemented in Matlab (Oostenveld et al., 2011). Only items correct at the behavioral level were used for further preprocessing and analysis.

The power spectrum was computed in the 2–45 Hz frequency range to accommodate the following frequency bands: 4–7 Hz (theta frequency), 8–12 Hz (alpha frequency), 13–20 Hz (low beta frequency), 21–30 Hz (high beta frequency) and 31–45 Hz (gamma frequency). To calculate the power spectrum, a 500 ms long stable moving window and a Hanning taper were employed. Power changes were computed in

steps of 50 ms and 1 Hz. Given the properties of the time-frequency analysis and the BVA to Fieldtrip export procedure, the resulting time-frequency representations (TFRs) contained data points between 500 ms prior to and 950 ms after the stimulus onset. These TFRs were expressed as a relative change from the –500 to –100 ms baseline period. Finally, TFRs were averaged for each subject and separately for each of the three conditions. For the remainder of the analysis, both at the brain and behavioral level, the interference effect was used, i.e., the difference between incongruent and congruent trials.

Statistical data analysis

Behavioral analysis

For the analysis of reaction times (RTs), RTs lower than 200 ms and non-accurate were excluded from further analysis. This led to the removal of 139 trials (0.97% of the trials). Repeated measures ANOVAs were performed for RTs (factors: Group [heritage speakers (HSs), L2 learners (L2)] \times Condition [incongruent, congruent, neutral]. For the accuracy analysis (Acc), generalized linear models from the binomial family were employed looking at the fit of condition and group on accuracy [glm(accuracy \sim condition*group, family = binomial)].

Time-frequency representation analysis

Analyses were performed on total oscillatory activity. Two sets of statistical analyses on the EEG data were performed in Fieldtrip: the first was aimed at comparing the interference effect between the two language groups (testing the first hypothesis, i.e., between group differences), the second was aimed at identifying the time-frequency clusters of interest for further analysis of individual differences (testing the second hypothesis, i.e., related to individual differences). A 1,000-randomizations cluster-based permutation approach (see Maris and Oostenveld, 2007) was used (two tailed dependent t-tests, cluster alpha = 0.05). The tests were run on the entire post-stimulus onset window with averaging per frequency band. A statistical

¹ For 5 participants the low cut-off was adjusted to 1 Hz in order to filter out recurrent skin artifacts.

threshold of $p < 0.025$ per tail was used to compute t -values for every electrode-time-frequency point and for corrected cluster-level significance.

Between-group comparison of the interference effect

To compare the interference effect between the two groups, we first subtracted the power spectrum for the congruent condition from the incongruent condition separately for the HS and L2 groups. We then ran a cluster-based permutation test comparing the resulting power spectra between them.

Identifying time-frequency clusters of interest

To select the time-frequency-electrode clusters of interest (scalp electrode region) for further analysis of individual differences, we compared power between the incongruent and congruent conditions (interference effect) in the frequency bands described above. For reasons discussed below, this process was done at the group level, i.e., for both groups separately, and at the sample level, i.e., collapsing all participants into one group. As can be seen in the figures below, at the group level, statistically significant condition differences were found within the following clusters: a positive theta cluster (300–600 ms) across fronto-central electrodes, a negative alpha cluster (600–950 ms) and positive alpha cluster (200–500 ms) over broad frontal electrodes, and a negative low beta cluster (350–950 ms) across posterior electrodes for the *L2 learners* (Figure 1). For the *HSs* (Figure 2) the following pattern emerged: a positive theta cluster (350–650 ms) across fronto-central electrodes, a broadly distributed negative alpha cluster (650–950 ms), a positive alpha cluster (300–500 ms) in fronto-central electrodes, and a broadly distributed negative low beta cluster (550–950 ms). At the *whole collapsed sample level* (both bilingual type groups included), the following clusters were found: a positive theta cluster (300–650 ms) across fronto-central electrodes, a broadly distributed negative cluster (600–950 ms) and a positive (350–500 ms) alpha cluster across fronto-central electrodes, a broadly distributed negative low beta cluster (450–950 ms) and a negative high beta cluster (750–950 ms) over central-posterior electrodes. These clusters were extracted and used for the computation of the average power-individual-frequency band.

Individual differences, language experience factors and brain interaction

A multiple regression analysis was conducted, extracting individual power values (for each frequency band of interest) correlated to the language variables derived from the LSBQ: non-societal language exposure and use at home (NSL-Home), non-societal language use in the society or community (NSL-Social), proficiency in the societal language (Proficiency), Age of L2 or 2L1 onset (AoA), and Length of exposure to the non-societal language (LoE). Biological age (Age) at time of testing

and LoE highly correlated ($r = 0.95$, $p < 0.001$). For this reason, we decided to drop LoE from our models. We also included Sex (male, female) and socio-economic status (SES) as covariates. All continuous variables included in this and following models were centered around the mean. Treatment coding was applied to categorical variables. The process was done in two parts. First, we correlated the language variables to the complete brain collapsed sample (linear regression models) (see Section **Identifying time-frequency clusters of interest** for clusters of interest). Then, we looked at individual correlations between these language experience factors and neural oscillations within each group separately (linear regression models) (see Section **Identifying time-frequency clusters of interest** for clusters of interest). All models were performed using the *lmrob* function from the *robustbase* package (Yohai, 1987; Koller and Stahel, 2011) on the statistical software R (R Core Team, 2021). Robust functions are designed to be more sensitive to outliers and suboptimal normal residuals, which is often the case with brain data.

Interaction brain and behavioral analysis

For this analysis, the interrelationship between brain oscillations and behavioral interference was investigated, i.e., how neural correlates predict response speed differences between incongruent and congruent trials (testing the third hypothesis). To do so, we used the data at the *whole collapsed sample level* (both bilingual type groups included) (see Section **Identifying time-frequency clusters of interest** for clusters of interest).

Results

Behavioral results

A summary of the behavioral results values (mean accuracy and reaction times) is shown below (Table 1). The accuracy analysis revealed that generally both groups performed better in the congruent conditions in comparison to the incongruent condition ($E = 1.91$, $SE = 0.44$, $p = 0.0002$ for the L2 learners and $E = 2.80$, $SE = 0.52$, $p = <0.0001$ for the HSs) and neutral ($E = 1.54$, $SE = 0.28$, $p = <0.0001$ for the L2 learners and $E = -1.99$, $SE = 0.36$, $p = <0.0001$ for the HSs). Differences between the groups were only found in the incongruent condition ($E = 0.62$, $SE = 0.20$, $p = 0.03$), where the L2 learners were more accurate. In the analysis on RTs, there was a main effect of condition [$F_{(2,174)} = 119.172$, $p = <0.0001$]. Both groups were faster in the congruent and neutral conditions in comparison to the incongruent condition (see Table 1), but there were no group differences in any of the conditions. Furthermore, no language group by condition interactions were found.

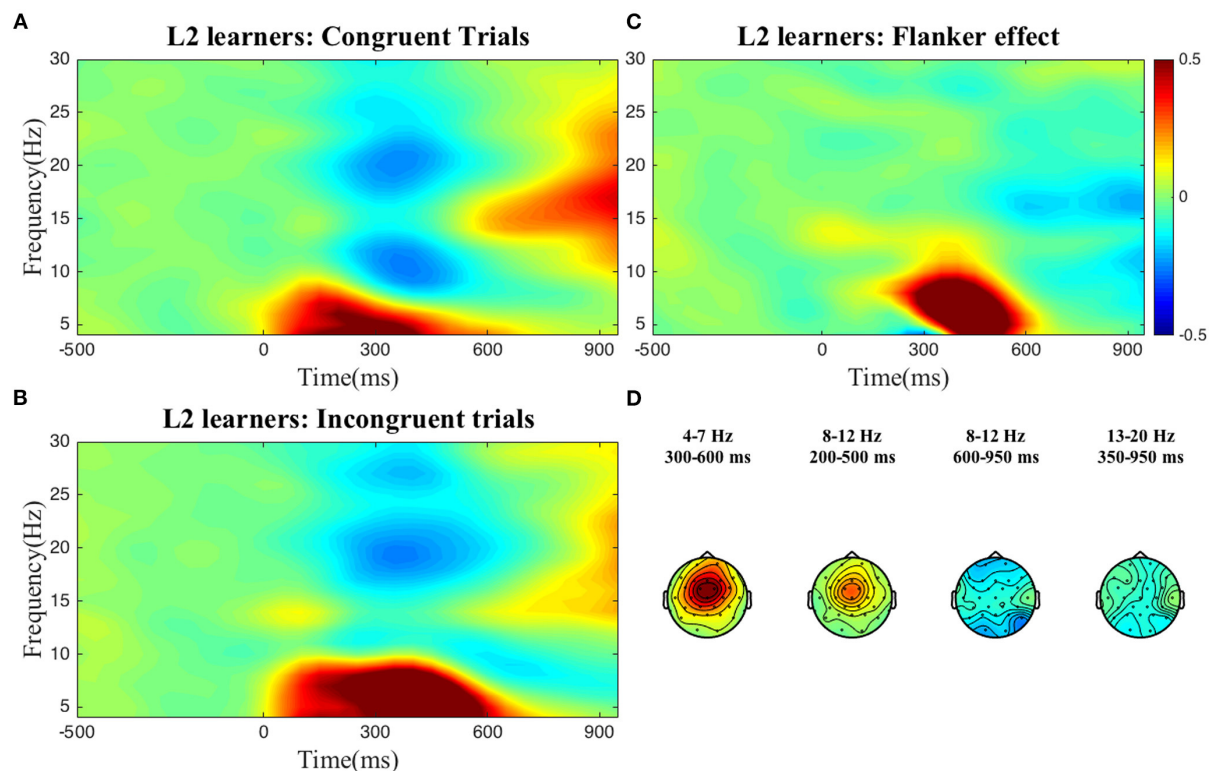


FIGURE 1

TFRs for the L2 learner group. Averaged time-frequency representations of individual conditions (A,B) and the Flanker effect (C) for the L2 learners for 4–30 Hz in a representative electrode, Cz and topographical plots (D) of the Flanker effect for the time-frequency clusters of interest. The Flanker effect was computed by subtracting power in the Congruent condition from power in the Incongruent condition. The color bar applies to both single electrode and topographical plots and indicates relative power change.

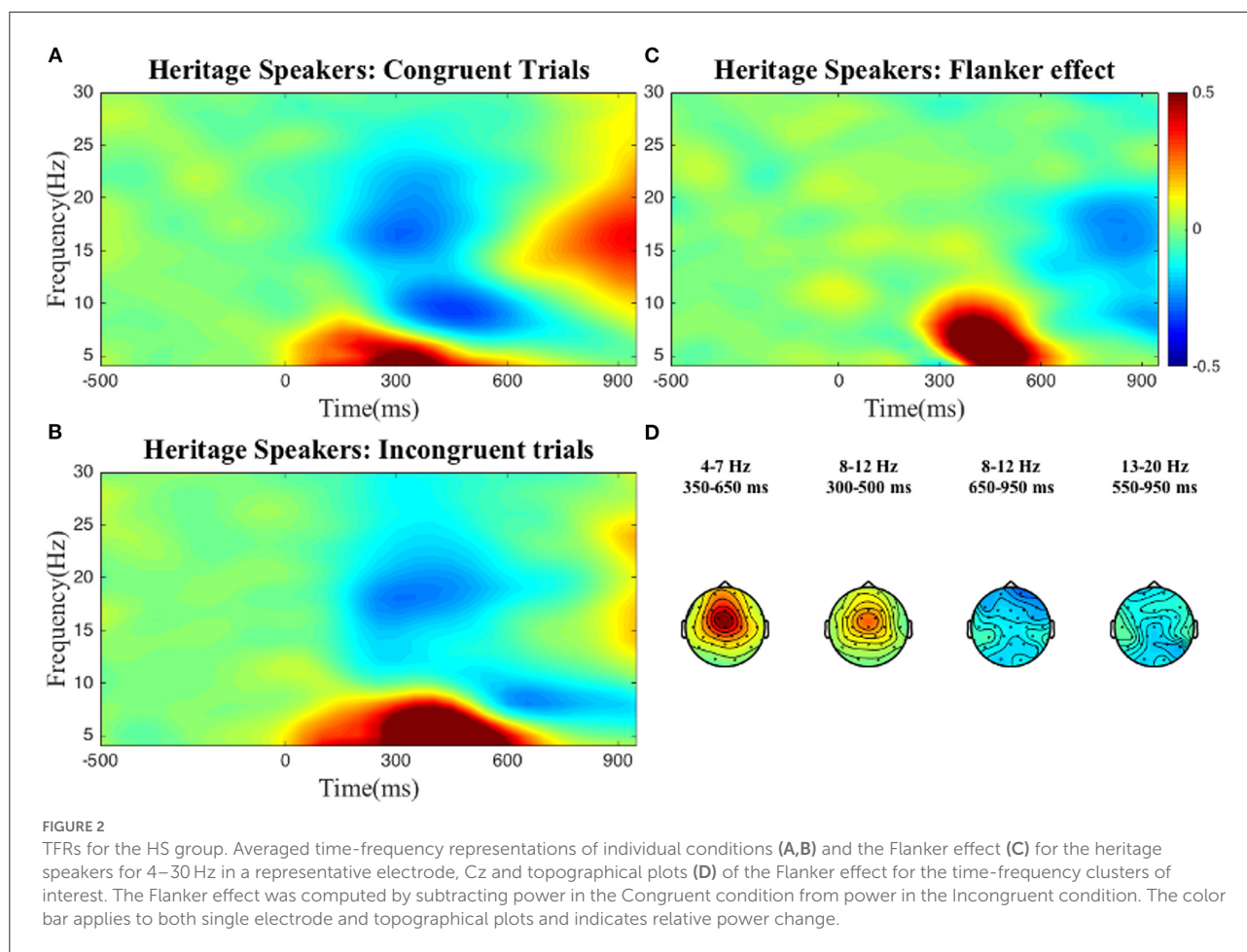
Time-frequency results; between-group comparison of the interference effect

We found no differences between the two groups, i.e., at the aggregate level HSs and L2 learners displayed very similar brain oscillatory patterns when looking at interference suppression in the Flanker task (see Figures 1, 2).

Individual differences, language experience factors and brain interaction

Although there do not seem to be any aggregate differences, this does not mean that individual level factors are not at play, whether in a collapsed group of all participants or across participants in each group separately. In principle, it is possible that each group contains participant samples with similar degrees of relevant experiential variation washing out any obvious effect. We cannot discount *a priori* that individual linguistic experience/engagement could trump any potential

effect AoA and/or -lingualism status (bi- vs. multilingualism) might confer independently. In other words, it might be the case that timing (onset and duration) of multiple language experience and/or quantity of languages (bilingual vs. multilingual) offer no explanatory value above and beyond individual engagement with multiple language use whenever one attains competence in more than one language. If so, this would mean that there are no caveats to the claim that linguistic engagement patterns are primarily deterministic for neurocognitive outcomes. After all, it is not only conceivable but indeed realistic to find adult L2 learners who are much more engaged in bilingual experiences than some childhood simultaneous bilinguals, even if the general trend and intuition pushes us to think in the opposite direction. If on the right track, there would be no reason to treat HSs and L2 learners distinctly. Of course, this is an empirical question. One, we submit, of significant importance that our design allows us to address. To test this, we first collapsed the two groups and ran multiple regressors to investigate how language usage variables predict neural outcomes overall. Some interesting results emerged significant, yet only for low beta. However, the question remained as to whether or not collapsing is the best



approach. If it turns out to be that group categorizing variables (e.g., AoA) brings something to bear independently, probing for individual differences related to linguistic experiences would still make sense to do regardless. However, doing so within each group separately should prove more meaningful.

Upon careful inspection of the data, it became clear that when collapsed together the two groups were clustering separately, instead of forming a homogenous bilingualism continuum (see Figure 3 for an example of this phenomenon). Therefore, the preferred analysis was the one that treated the groups separately. Moreover, upon treating the groups separately we note that significant effects are more pervasive, that is, not relegated only to low beta.

In order to examine whether language experience modulates the magnitude of power for each group, we again ran multiple regression analyses for each cluster. We report only significant results. Starting from the L2 learners, we found effects in alpha (negative cluster) and low beta (negative cluster) frequencies, and exclusively positive correlations with

the experience factors. The significant background-related factors that predicted power in the alpha cluster were AoA ($E = -0.02$, $t = -2.17$, $p = 0.04$) and Age ($E = 0.03$, $t = 2.33$, $p = 0.03$) (see Figure 4), whereas NSL-home ($E = 0.02$, $t = 2.47$, $p = 0.02$) significantly correlated with low beta (Figure 5).

For the HS group, we only found a negative correlation (Figure 6) between Age and theta power (positive cluster) ($E = -0.03$, $t = -2.52$, $p = 0.02$).

Brain and behavior interaction

In this analysis, we modeled RTs interference as predicted by brain-data*group. We found a significant interaction between group and alpha power (in the positive cluster) ($E = -0.99$, $t = -2.06$, $p = 0.04$) (Figure 7). Interestingly, the two groups show opposite patterns. While a positive correlation is found between alpha power and interference suppression effect for the L2 learners, the HSs exhibit a negative correlation: the more

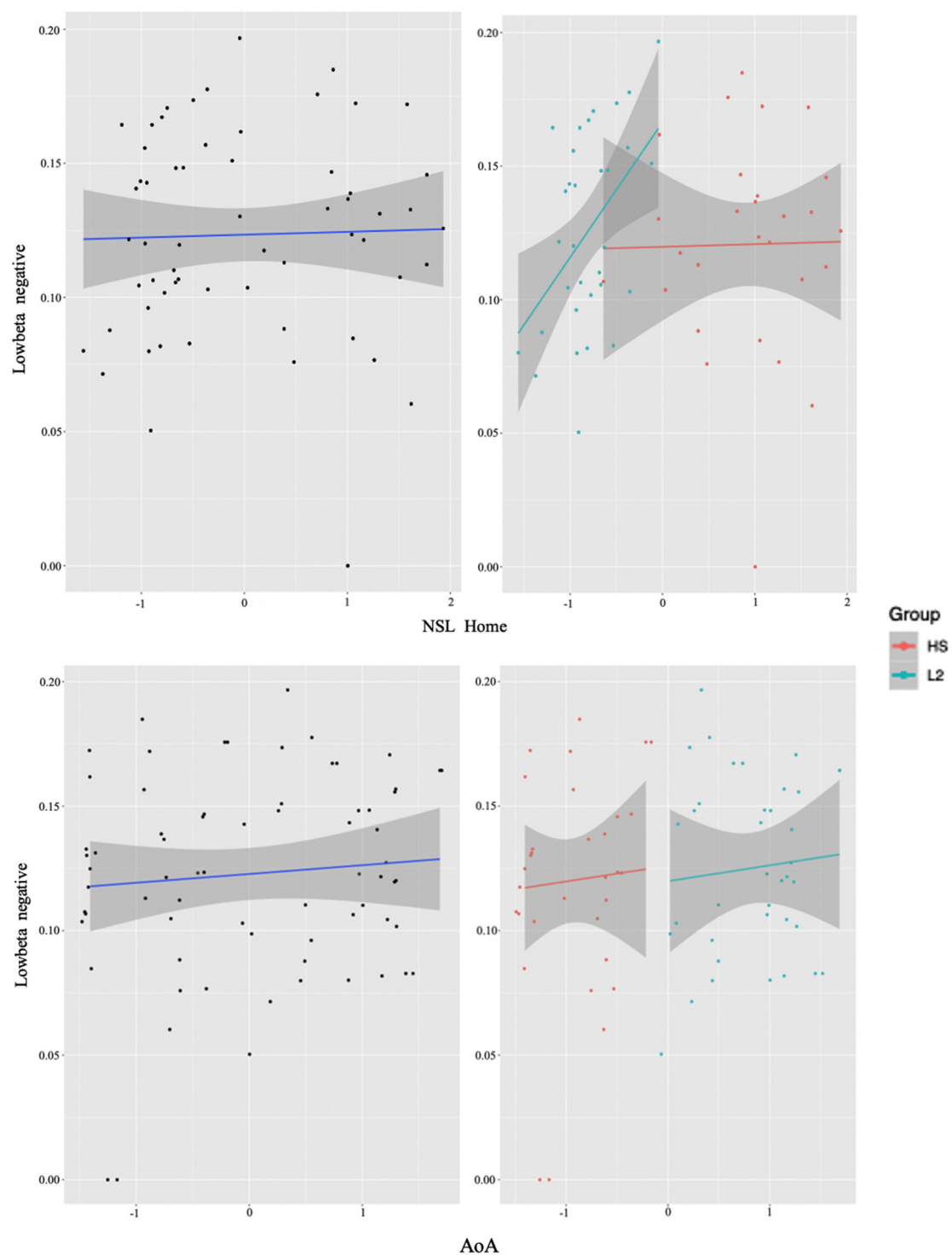


FIGURE 3

Collapsed group analysis results. The two figures on the left represent the correlation between experience factors (NSL Home top and AoA bottom) and low beta power from the whole collapsed dataset (units on both axes are normalized ones: x-axes age and y-axes power) (see Section **Identifying time-frequency clusters of interest** for clusters of interest). In the two right figures, the individual data points, representing the participants, have been sorted out into the two groups (HS for the heritage speakers of Italian and L2 for the L2 learners), where again the NSL Home interaction is on the top and AoA on the bottom. Notice how the two groups cluster separately, with little to no mixing between them. In the two right figures, the L2 learner group (L2) is represented in blue and the HS group (HS) in red.

alpha was recruited, the smaller the difference was between incongruent and congruent trials.

Discussion

In what follows, we unpack and interpret the above presented results in general and specifically in relation to our three hypotheses outlined in the introduction.

Between-group comparison of the interference effect

The first hypothesis anticipated the possibility that the L2 learner group would exhibit increased theta activation whereas, conversely, the HS group might show increased alpha suppression for inhibitory control as a result of the inherent difference in duration (and context) commensurable with bilingual type (DeLuca et al., 2020). The logic was as follows: adaptations to control demands (such as those associated with bilingual experience) are predicted to initially manifest in fronto-cortical regions and networks (Grundy et al., 2017; Pliatsikas, 2020), linked to increased theta band activation in tasks measuring processes related to inhibitory control (Cavanagh and Frank, 2014; Duprez et al., 2018; Brunetti et al., 2019; Pscherer et al., 2021). Given this, we might have expected that the shorter duration of bilingual experience associated with the L2 learner group to manifest as greater reliance on fronto-cortical circuits to handle inhibitory control demands, resulting in greater theta band activation. As discussed in the introduction, here we are referring to absolute (quantitative) as well as relative (qualitative) “time”, given that the L2 English is not within immersion, but the HS experience entails intensive early immersion in two languages by definition². Alternatively, the longer timeframe (and intensity) of exposure to (at least)

² As argued in the UBET framework (DeLuca et al., 2020), greater intensity of exposure, in isolation, should result in adaptations to increased control demands. However, in combination with prolonged duration of bilingual exposure, increased intensity would decrease the latency by which adaptations to efficiency in executive control would occur. Indeed, such trends have been seen in participant cohorts in intensive L2 immersion environments when compared to participants with potentially longer overall duration of use but lower intensity (see for comparison e.g., Burgaleta et al., 2016; Pliatsikas et al., 2017). Given our L2 learners are outside of immersion where English has a clear functional, but not wide societal dispersion, even for individuals who report a young English AoA by virtue of early classroom exposure such a dual language context is not equivalent to the intensity of two languages in a HS context where both languages are used across familial, social, and other contexts. And so, we tested the hypothesis that in our L2 learners compared to HSs this would play out empirically.

two languages inherent to the HS context would show a shift in reliance toward subcortical circuits to handle these control demands more efficiently (Grundy et al., 2017; Pliatsikas, 2020), which would result in stronger modulations in alpha power suppression (Mazaheri et al., 2014).

However, our findings show no such differences, disconfirming our first prediction. This, of course, does not mean that all bi-/multilinguals are the same. It is simply the case that our HSs, whose timing and intensity to bilingualism in early childhood tangibly differed, were not distinct in the aggregate from the L2 learners studied herein. Whereas, frameworks like UBET (DeLuca et al., 2020) and the *Bilingual Anterior to Posterior and Subcortical Shift* (BAPSS) (Grundy et al., 2017), from which we derived the predictions, anticipate such a distinction, neither make reference to the timeframe under which adaptations toward efficiency occur. Keeping in mind, then, that our L2 learners are not only proficient in English but have had significant time with English competence it could simply be that they are past a stage in bilingual development where increased reliance on fronto-cortical circuits is implicated. In other words, for all intents and purposes any potential effect duration of bilingualism could have had is surpassed in these groups. Without a monolingual comparison or an additional bilingual group with less L2 experience (beginners) we cannot tease the following three possibilities apart: (i) bilingualism has no such effect here (unlikely in light of the rest of the data unpacked below), (ii) there simply is no difference between L2 learners and HSs *par excellence*, that is, at any stage of development or (iii) such differences would obtain, but only when L2 learners are under a particular threshold of bilingual experience. Ultimately, it might not have been a fair question to ponder from the outset since HSs tested in adulthood are inherently at a mature state of linguistic bilingual development, whereas L2 learners at similar ages are at various stages of linguistic development—i.e., with HSs we are looking at neurocognitive effects of bilingual language maintenance as opposed to bilingual language development as one can do at various stages of proficiency with L2 learners.

Beyond the aspect of AoA, which admittedly is confounded in our HSs with number of languages as a consequence of the reality of HS bilingualism in the European context, one might ask if multilingualism over bilingualism evidences any differences as well. Recall that our design and specific sampling matched the HSs and L2 learners in English and German proficiency as well as dominance (they are all German dominant). Thus, linguistically what differentiated them was the presence or absence of an additional, heritage Italian, grammar in the mind. Given our results, one might conclude multilingualism over bilingualism and, more specifically, heritage language experience *per se* is not deterministic. Any such definitive conclusion at this point, however, would be

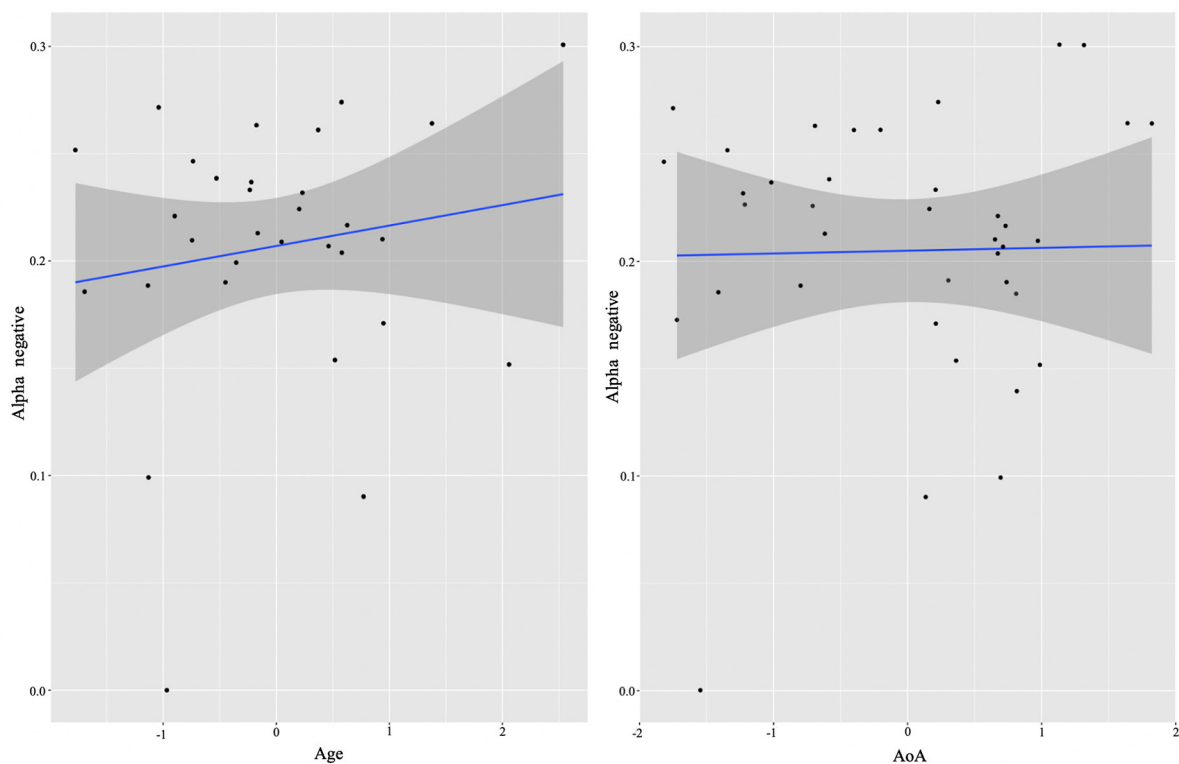
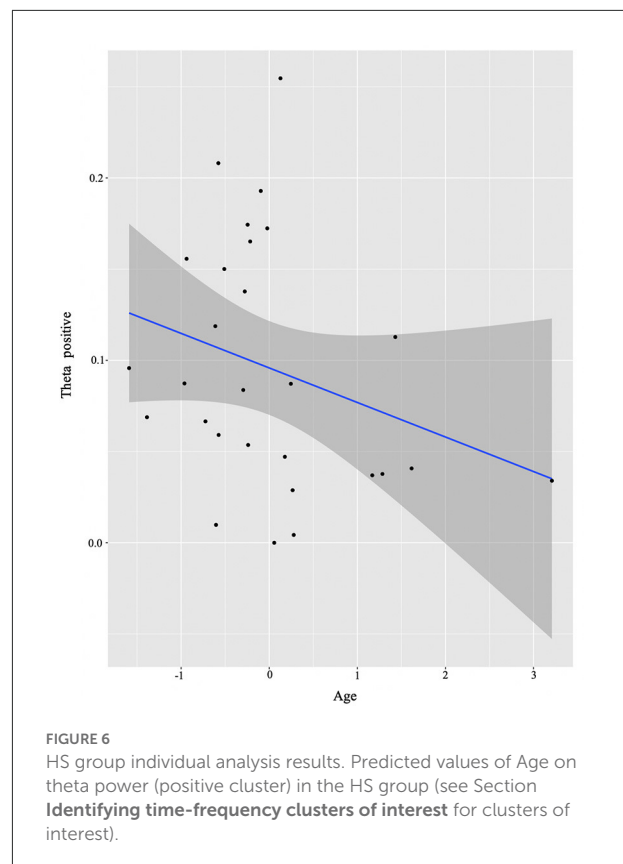
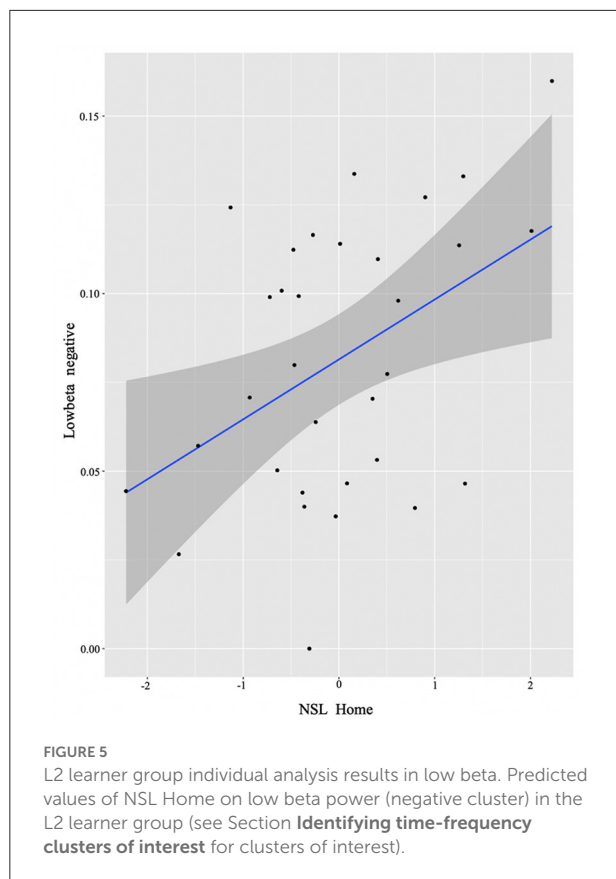


FIGURE 4
L2 learner group individual analysis results in alpha. Predicted values of Age (left) and AoA onset (right) on alpha power (negative cluster) in the L2 learner group (see Section **Identifying time-frequency clusters of interest** for clusters of interest).

precipitous based on this initial analysis alone. As we will unpack below, such a conclusion is, in fact, challenged by other aspects of the present data. Given the nature of our data—the mere reality of HS bilingualism in Europe where English is not the societal language but ubiquitous—we will not be able to tease apart AoA from multilingualism with our HSs. Notwithstanding, we suspect that AoA rather than multilingualism *per se* or an interaction of the two is the driving force of the differences noted in our results that we elaborate on below, not least in the context of other studies showing AoA effects on neurocognitive adaptations likened to bilingualism (e.g., Luk et al., 2011; Tao et al., 2011; Delcenserie and Genesee, 2017, see Berken et al. (2017) for a review of MRI-based AoA supportive evidence).

In truth, little is known about the processes of neurocognitive adaptations to multilingualism as a *departure* from bilingualism; that is, if and how the brain further adapts to the acquisition and management of a third (or more) language. It is intuitive, if not theoretically reasonable, to anticipate that adding more languages to the brain/mind systems would increase control and processing demands for

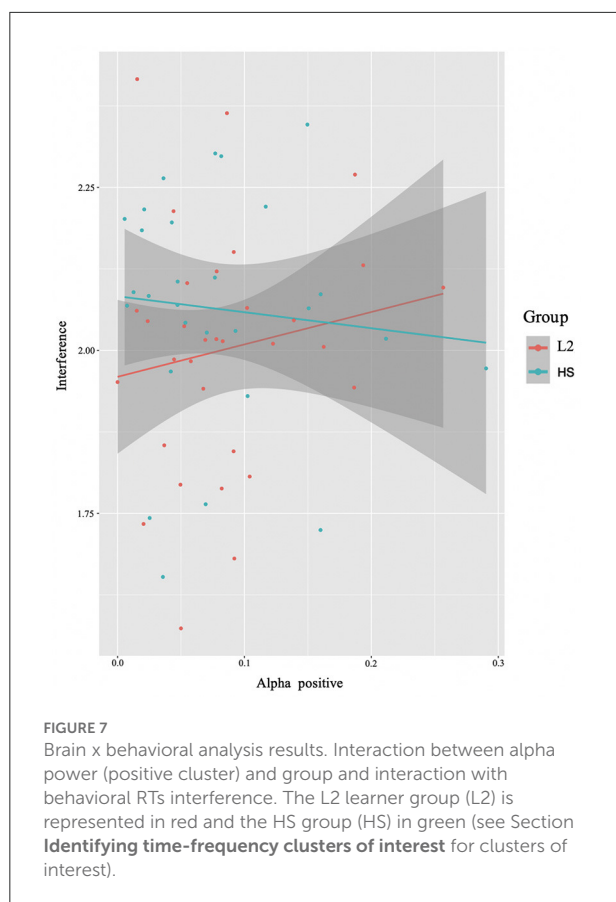
successful language selection and use. After all, there is no shortage of literature showing that multilingual acquisition and processing are distinct from bilingualism in linguistic domains (see for review Rothman et al., 2019). However, it is equally reasonable to expect that three or more languages would bring nothing or very minimal effects to bear on domain general neurocognitive adaptations specifically above and beyond two. This could be so either because the change in state in the relevant sense from bilingualism to multilingualism is much less severe than it is from monolingualism to (at least) bilingualism or because there are ceiling effects for relevant adaptations that are already met by managing two languages, at least under certain conditions. In other words, the brain might show highly differential adaptations when moving from a one-language system to a two-language system, but these adaptations may not be as required (or visible) when moving to a multi-language system. In this light, individual differences in language experiences, rather than number of languages spoken, are likely to drive patterns of multilingual neurocognitive adaptation (see Yee et al., 2022), to which we now turn.



Individual differences in language experiences

To test our second hypothesis regarding individual difference correlations to bilingual experience, we ran a cluster-based permutation analysis to explore the neural mechanisms accompanying conflict resolution, extracting power values per condition and regressing them with continuous measures of bilingual experience. For reasons discussed above, we first did this in a combined sample of the two groups, then by analyzing them separately. Results suggested that AoA of the L2/2L1 and usage of the non-societal language at home positively predicted low beta recruitment for the whole group (see Figure 3). However, when we plotted this while annotating individuals by group, it became visually apparent that only one of the two groups (the L2 learners) was driving the non-societal language effect and, overall, that the groups were not mixing into the continuum of individual experiences. Therefore, we will not discuss the results in terms of the combined group. Alternatively, we re-ran the analysis separating out the two original groups not least because previous work clearly indicates that both bilingual types, on their own, could potentially display within-group individual variation (Kupisch and Rothman, 2018; DeLuca et al., 2019a, 2020).

To the extent that this approach panned out, as in our analysis below, there is a combined epistemological and empirical note of importance. It can be true that bilingual experience is deterministic for outcomes of neurocognitive adaptations in general but teasing this out can be inadvertently obscured by combining proverbial apples and oranges. Combining distinct types of bilinguals, as in the whole group analysis, with crucial variables that make them incomparable at some level of apparent importance, such as AoA, could interact with how experience is differentially uptaken in each bilingual type and thus introduce noise. In other words, experience can matter generally, but not necessarily play out in the same ways depending on how experience interacts with other features. If so, we might expect to see that variation in experience predicts outcome differences regardless of bilingual type, but is appreciated only in comparison to peers within a singular type (at least for some outcomes whereas others might not show a type division). Conversely, it could be the case that experience only matters when certain other variables are true, for example, only if bilingualism commenced before a particular age. Maintaining group distinctions then, as warranted by our first pass analysis, will ultimately help us to determine this. Perhaps then, the order in which we approached these analyses should have been reversed. Examining the groups separately and



showing the same trends in each would have certainly warranted a follow up to see if the same held true when combining both groups: a lesson learned should also be a lesson shared.

Returning to the two-group regression analysis, let us remind the reader which clusters emerged for both the HS and L2 learner groups: (i) a positive theta cluster, (ii) an early positive alpha cluster, (iii) a late negative alpha cluster and (iv) a negative low beta cluster. These clusters are in line with previous literature on the oscillatory dynamics of domain general cognitive control. Increased theta power in the incongruent relative to the congruent conditions reflects greater conflict resolution demands (e.g., Nigbur et al., 2011). Both alpha decreases and increases in incongruent relative to congruent trials have been reported previously (McDermott et al., 2017). These increases and decreases have different neural generators. For example, McDermott et al. (2017) localized alpha decreases to the parietal and occipital cortices as well as the areas near the right temporoparietal junction (TPJ) with these latter areas being a part of a ventral attentional network believed to facilitate the detection of the stimulus, especially in the presence of unexpected or distracting stimuli. Alpha power increases were found in the part of the dorsal attention network which supports the selection of sensory stimuli and

their linking to the appropriate motor response (McDermott et al., 2017). Given the role of beta power for motor response preparation and execution (Engel and Fries, 2010; Grent-'t-Jong et al., 2013), the relative decrease of beta power for the incongruent relative to the congruent conditions could reflect conflict at the response level, that is simultaneous activation of both response directions in incongruent trials compared to the single direction in congruent trials.

Thus, qualitatively the HSs and L2 learners in their aggregates exhibited similar neural signatures of inhibitory control. However, in line with hypothesis 2, further analysis of individual differences revealed that these effects were differentially modulated by experiential factors for the two groups. For the L2 learners, a negative alpha cluster positively correlated with chronological Age (which, given the relatively small difference in age between the participants, can be considered as a proxy for the duration of exposure) and AoA. That is, the older the participants were at time of testing (hence the more exposure they had to the L2) and the older they were when they acquired their L2, the smaller the alpha power decrease accompanying conflict resolution was. Under the interpretation that reduced alpha suppression reflects more efficient task performance (Zhuang et al., 1997; Riečanský and Katina, 2010; Kielar et al., 2014), the Age effect suggests that participants who were exposed to their L2 longer demonstrated less effortful processing of incongruency in the Flanker task. This result is in agreement with the theoretical suggestion that longer duration of dual language use (which implies prolonged juggling of two language systems) confers adaptations that afford more efficient and automated processing of conflicting information (Grundy et al., 2017; Pliatsikas, 2020). However, as discussed in the introduction, there is also an account predicting that prolonged duration of bilingual experience would be associated with increased alpha suppression in posterior and subcortical areas in situations of control demands (DeLuca et al., 2020). This prediction is more in line with our finding of positive correlation between the AoA (given that it is often confounded with prolonged duration of bilingual use) and alpha decrease, whereby the earlier the participants learned their L2, the stronger alpha decreases are exhibited. Thus, the effects of Age and AoA are somewhat contradictory in our study and further work should more closely examine the contributions of these different but related factors.

Additionally, within the L2 group, the negative low-beta cluster positively correlated with the use of non-societal language at home, namely the more English is used at home, the smaller the relative beta power decrease is. Although at first glance this result might seem controversial, when considering the use of English at home in a society like Germany in younger populations, we can easily pinpoint factors of language-use divergence across individuals (some having international roommates or friends, playing online games, differential usage of social media and others). And so, similarly to the alpha

power results, such reduced beta power may indicate more efficient handling of conflicting information (Kielar et al., 2014) by individuals who more intensely use their L2. This is in line with the account suggesting that a change in control demands is expected with increased intensity of L2 use (DeLuca et al., 2020).

For the HSs, Age at time of testing negatively correlated with power in the theta frequency band, i.e., the older the participants were (hence, the more exposure to Italian and English, and thus prolonged exposure to multiple languages), the less theta power they recruited for processing incongruency. The reduction of theta power as a function of duration of exposure might signal a reverting to baseline levels of the functional recruitment patterns in fronto-cortical regions resulting from the increased efficiency in juggling their languages and a reduction in inhibitory control demands (DeLuca et al., 2020; Pliatsikas et al., 2021). So how can it be that given very similar brain signatures of conflict resolution between the two groups, we find these signatures to be differentially modulated by language experience factors? Here is where we believe we see an effect of HS bilingualism itself. Recall that the groups are controlled for context of learning and proficiency of English. However, the HSs experience with Italian under an immersion context, i.e., their earlier intensity to naturalistic bilingualism, drives differential outcomes.

The groups did not overlap on the continuous measures of AoA and the use of non-societal language at home. Thus, it could be that while these distinctions in experience do not lead to qualitative differences in which neural operations are recruited for processing conflict (at least the kind of conflict induced by the Flanker task), they do define the ways in which they interact with experiential factors. It could also be the case that when more than one language is acquired early enough and/or when the non-societal language use at home is more intense, there are threshold effects related to experiential factors (e.g., duration of exposure) differentially triggering conflict detection and the communication of the subsequent need for enhanced control (Cavanagh and Frank, 2014). When the onset of exposure to another language occurs past a certain threshold (e.g., as in our L2 learners) and non-societal language is minimally or not used at home, the duration of exposure may not have that same effect. However, in the case of later onset of exposure to L2 and minimal use of non-societal language at home, experiential factors modulate the engagement of attentional mechanisms required for managing the conflict as reflected in alpha oscillations and recruitment of the motor system reflected in beta oscillations.

The current study does not allow us to disentangle the effects of AoA and non-societal language use at home on the relationship between experiential factors and neural signatures of inhibitory control. It is possible that the differences observed between the two groups are driven by one or both of these measures, or that adaptations required by each are likely to be modulated by the other (DeLuca et al., 2020). Further work should address this issue by carefully selecting appropriate test

populations to address this query, for example, by adding a group of multilinguals with the same languages to our HS group but have acquired their L2 and L3 in adulthood (e.g., German natives who acquire English in the typical (early L2) case and are L3 learners of Italian with high proficiency).

Brain and behavioral interaction

As per hypothesis 3, we investigated the anticipatedly observable relationship between oscillatory dynamics and reaction time (RT) performance, inclusive of whether the pattern was modulated by differences in language experience (Figure 7). Recall here that the HSs and L2 learners showed differing patterns: the L2 learners showed a positive correlation between alpha power and RTs for inhibitory control, whereas the HSs showed a negative correlation between alpha power and task performance for this contrast.

The asymmetry in effects provides some support for the predictions of the UBET framework (DeLuca et al., 2020). Specifically, the decreased alpha power seen within the HS group can be interpreted as a measure of increased efficiency. Recall that the EEG data included within the model (see Section **Interaction brain and behavioral results**) came from a *positive* alpha cluster; that is, an increase in alpha power for incongruent relative to congruent trials. As increases in alpha per condition can be interpreted as inhibition of task-irrelevant information (see e.g., Van Diepen et al., 2019), the negative correlation here indicates that the HSs seem to rely *less* on inhibitory control processes to achieve similar reaction time performance for interference suppression.

The pattern seen in the L2 learners was not expected, but also suggests an adaptation to the nature of language control demands. The positive correlation between alpha power and reaction time indicates a shifting adaptation to inhibitory control costs that might not have (yet) reached a degree of efficiency noted in the HSs. This indicates a greater requirement/focus on specific cognitive networks related to language- and domain-general cognitive control. This effort then would lead to a slight slowing on trials requiring increased inhibitory control (such as the incongruent trials), resulting in larger interference effects at the level of RTs.

It is possible that with prolonged (or more intense) engagement with the L2, a broader cognitive control network would be engaged. Recall that the language demographics for the L2 learner group is characterized by L1-dominance (German) with relatively shorter exposure to the non-native language (see **Supplementary Table 1**). Given the decreased opportunities for engagement and the shorter time frame overall to engage in language control, this could reflect a system that is not fully optimized to handle the control demands (Grundy et al., 2017; Pliatsikas, 2020). It is possible that with increased intensity of use or prolonged duration of use, the patterns seen in this group

would more closely align with those seen in the HS group, but further research targeting individuals and groups with such demographic patterns is required to assess this.

Conclusion

Investigating EEG oscillations in a Flanker task, this study examined potential neurocognitive adaptations in two types of bilinguals. Motivated by inherent differences in duration (and entailed intensity by contextual distinctions) of bilingual experience, we hypothesized that our groups would diverge relative to each other: L2 learners would exhibit increased theta activation while HSs would show increased alpha suppression. However, between-group comparisons revealed no such distinctions. In line with the literature seeking to reveal—if existent in a given aggregate—and understand individual differences in bilingual language and executive control, we regressed various non-neural (e.g., demographic) variables to probe for prediction of individual differences in neural activation. When examining the Flanker effect separately within each group, we did find that both groups demonstrated (partial) signatures of cognitive control (specifically theta increase, initial alpha increase followed by later alpha decrease as well as low beta decrease). Indeed, the data revealed specific bilingual experiential factors correlated to modulatory effects in each group, but differentially so. Moreover, the two groups demonstrated distinct relationships between oscillatory dynamics and behavioral performance on the Flanker task (reaction times). Overall, insights from this study support the view that individual engagement matters in bilingualism and neurocognitive outcomes, but not necessarily in the same way across all bilingual(type)s. That said, it is also prudent to point out that our study is limited to a single cognitive task. That is, to the extent that Flanker is a good task to examine inhibitory control, the present can speak only to support our general conclusions for the relevant cognitive domain. And yet, claims of bilingualism—whether treated/envisaged as a continuous variable or not—potentially affecting executive functions are not limited to the domain of inhibitory control. And so, a modicum of caution needs to be applied in terms of the generalizability of even what the present data support: while we maintain the present study shows that degree of bilingual engagement matters—differentially so for distinct types of bilinguals—for neurocognitive outcomes, this does not mean it affords supportive evidence beyond the tested domain.

Data availability statement

The original contributions presented in the study are included in the article/**Supplementary material**, further inquiries can be directed to the corresponding author/s.

Ethics statement

This study involved human participants. It was reviewed and approved by the University of Konstanz Ethical Committee. The participants provided their written informed consent to participate in this study.

Author contributions

The present study formed part of a doctoral thesis completed by SMPS. SMPS contributed to the design and conducted the experiments. SMPS, YP, and VD performed the data-analyses. SMPS, YP, VD, and JR contributed to the writing of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This article was supported by generous funding to SMPS by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska Curie Grant Agreement No. 765556. JR, YP, and VD were funded by the AcqVA Auora Center grant. JR and YP received funding from the Tromsø Forskningsstiftelse (Tromsø Research Foundation) Grant No. A43484 and the Heritage-bilingual Linguistic Proficiency in their Native Grammar (HeLPiNG) (2019–2023).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnhum.2022.910910/full#supplementary-material>

SUPPLEMENTARY TABLE 1
Participants' metadata.

References

- Abutalebi, J., and Green, D. W. (2016). Neuroimaging of language control in bilinguals: Neural adaptation and reserve. *Bilingualism: Lang. Cogn.* 19, 689–698. doi: 10.1017/S1366728916000225
- Anderson, J. A., Chung-Fat-Yim, A., Bellana, B., Luk, G., and Bialystok, E. (2018a). Language and cognitive control networks in bilinguals and monolinguals. *Neuropsychologia* 117, 352–363. doi: 10.1016/j.neuropsychologia.2018.06.023
- Anderson, J. A., Mak, L., Chahi, A. K., and Bialystok, E. (2018b). The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. *Behav. Res. Methods* 50, 250–263. doi: 10.3758/s13428-017-0867-9
- Bastiaansen, M., Mazaheri, A., and Jensen, O. (2011). “Beyond ERPs: oscillatory neuronal,” in *The Oxford Handbook of Event-Related Potential Components*, eds S. J. Luck, and E. S. Kappenman (Oxford, UK: Oxford Library of Psychology), 31–50.
- Berken, J. A., Gracco, V. L., and Klein, D. (2017). Early bilingualism, language attainment, and brain development. *Neuropsychologia* 98, 220–227. doi: 10.1016/j.neuropsychologia.2016.08.031
- Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychol. Bull.* 143, 233–262. doi: 10.1037/bul0000099
- Bialystok, E., Craik, F. I., and Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends Cogn. Sci.* 16, 240–250. doi: 10.1016/j.tics.2012.03.001
- Bice, K., Yamasaki, B. L., and Prat, C. S. (2020). Bilingual Language Experience Shapes Resting-State Brain Rhythms. *Neurobiol. Lang.*, 1, 288–318. doi: 10.1162/nol_a_00014
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652. doi: 10.1037/0033-295X.108.3.624
- Botvinick, M. M., Cohen, J. D., and Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends Cogn. Sci.*, 8, 539–546. doi: 10.1016/j.tics.2004.10.003
- Brunetti, M., Zappasodi, F., Croce, P., and Di Matteo, R. (2019). Parsing the Flanker task to reveal behavioral and oscillatory correlates of unattended conflict interference. *Sci. Rep.* 9, 1–11. doi: 10.1038/s41598-019-50464-x
- Burgaleta, M., Sanjuán, A., Ventura-Campos, N., Sebastian-Galles, N., and Ávila, C. (2016). Bilingualism at the core of the brain. Structural differences between bilinguals and monolinguals revealed by subcortical shape analysis. *NeuroImage* 125, 437–445. doi: 10.1016/j.neuroimage.2015.09.073
- Buzsáki, G. (2006). *Rhythms of the Brain* (Oxford, UK: Oxford University Press).
- Calvo, N., and Bialystok, E. (2021). Electrophysiological signatures of attentional control in bilingual processing: Evidence from proactive interference. *Brain Lang.* 222, 105027. doi: 10.1016/j.bandl.2021.105027
- Cavanagh, J. F., and Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends Cogn. Sci.*, 18, 414–421. doi: 10.1016/j.tics.2014.04.012
- Cavanagh, J. F., Zambrano-Vazquez, L., and Allen, J. J. (2012). Theta lingua franca: A common mid-frontal substrate for action monitoring processes. *Psychophysiology* 49, 220–238. doi: 10.1111/j.1469-8986.2011.01293.x
- Cespón, J., and Carreiras, M. (2020). Is there electrophysiological evidence for a bilingual advantage in neural processes related to executive functions? *Neurosci Biobehav Rev.* 118, 315–330. doi: 10.1016/j.neubiorev.2020.07.030
- Cohen, M. X. (2017). “Neurophysiological oscillations and action monitoring,” in *The Wiley Handbook of Cognitive Control*, ed T. Egner (NJ: Wiley Blackwell), 242–258.
- Cooper, P. S., Darriba, Á., Karayanidis, F., and Barceló, F. (2016). Contextually sensitive power changes across multiple frequency bands underpin cognitive control. *NeuroImage* 132, 499–511. doi: 10.1016/j.neuroimage.2016.03.010
- Del Maschio, N., and Abutalebi, J. (2019). “Language organization in the bilingual and multilingual brain,” in *The Handbook of the Neuroscience of Multilingualism*, eds J. W. Schwieter, and M. Paradis (Hoboken, NJ: John Wiley & Sons Ltd), 197–213.
- Delcenserie, A., and Genesee, F. (2017). The effects of age of acquisition on verbal memory in bilinguals. *Int. J. Bilingualism*, 21, 600–616. doi: 10.1177/1367006916639158
- DeLuca, V., Miller, D., Pliatsikas, C., and Rothman, J. (2019a). “Brain adaptations and neurological indices of processing in adult second language acquisition: challenges for the critical period hypothesis,” in *The Handbook of the Neuroscience of Multilingualism*, eds J. W. Schwieter, and M. Paradis (John Wiley & Sons Ltd), 170–196.
- DeLuca, V., Rothman, J., Bialystok, E., and Pliatsikas, C. (2019b). Redefining bilingualism as a spectrum of experiences that differentially affects brain structure and function. *Proc. Natl. Acad. Sci.* 116, 7565–7574. doi: 10.1073/pnas.1811513116
- DeLuca, V., Segaert, K., Mazaheri, A., and Krott, A. (2020). Understanding bilingual brain function and structure changes? U Bet! A Unified Bilingual Experience Trajectory model. *J. Neurolinguistics* 56, 100930. doi: 10.1016/j.jneuroling.2020.100930
- Di Pisa, G., and Pereira Soares, S. M., and Rothman, J. (2021). Brain, mind and linguistic processing insights into the dynamic nature of bilingualism and its outcome effects. *J. Neurolinguistics* 58, 100965. doi: 10.1016/j.jneuroling.2020.100965
- Duprez, J., Gulbinaite, R., and Cohen, M. X. (2018). Midfrontal theta phase coordinates behaviorally relevant brain computations during response conflict. *BioRxiv* 1–38, 502716. doi: 10.1101/502716
- Engel, A. K., and Fries, P. (2010). Beta-band oscillations—Signalling the status quo? *Curr. Opin. Neurobiol.* 20, 156–165. doi: 10.1016/j.conb.2010.02.015
- Engel, A. K., Fries, P., and Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top-down processing. *Nat. Rev. Neurosci.*, 2, 704–716. doi: 10.1038/35094565
- Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* 16, 143–149. doi: 10.3758/BF03203267
- Fan, J., Flombaum, J. I., McCandliss, B. D., Thomas, K. M., and Posner, M. I. (2003). Cognitive and brain consequences of conflict. *Neuroimage* 18, 42–57. doi: 10.1006/nimg.2002.1319
- Freunberger, R., Klimesch, W., Griesmayr, B., Sauseng, P., and Gruber, W. (2008). Alpha phase coupling reflects object recognition. *Neuroimage* 42, 928–935. doi: 10.1016/j.neuroimage.2008.05.020
- Fuchs, E., and Flügge, G. (2014). Adult neuroplasticity: more than 40 years of research. *Neural Plast.* 2014, 541870. doi: 10.1155/2014/541870
- Gallo, F., Myachykov, A., Shtyrov, Y., and Abutalebi, J. (2020). Cognitive and brain reserve in bilinguals: field overview and explanatory mechanisms. *J. Cult. Cogn. Sci.* 4, 127–143. doi: 10.1007/s41809-020-00058-1
- Gallo, F., Novitskiy, N., Myachykov, A., and Shtyrov, Y. (2021). Individual differences in bilingual experience modulate executive control network and performance: Behavioral and structural neuroimaging evidence. *Bilingualism: Lang. Cogn.*, 24, 293–304. doi: 10.1017/S1366728920000486
- Green, D. W., and Wei, L. (2014). A control process model of code-switching. *Lang. Cogn. Neurosci.*, 29, 499–511. doi: 10.1080/23273798.2014.882515
- Grent-’t-Jong, T., Oostenveld, R., Jensen, O., Medendorp, W. P., and Praamstra, P. (2013). Oscillatory dynamics of response competition in human sensorimotor cortex. *NeuroImage* 83, 27–34. doi: 10.1016/j.neuroimage.2013.06.051
- Grundy, J. G. (2020). The effects of bilingualism on executive functions: An updated quantitative analysis. *J. Cult. Cogn. Sci.*, 4, 177–199. doi: 10.1007/s41809-020-00062-5
- Grundy, J. G., Anderson, J. A., and Bialystok, E. (2017). Neural correlates of cognitive processing in monolinguals and bilinguals. *Ann. N. Y. Acad. Sci.*, 1396, 183–201. doi: 10.1111/nyas.13333
- Gullifer, J. W., and Titone, D. (2020). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Lang. Cogn.*, 23, 283–294. doi: 10.1017/S1366728919000026
- Haciahmet, C. C., Frings, C., and Pastötter, B. (2021). Target Amplification and Distractor Inhibition: Theta Oscillatory Dynamics of Selective Attention in a Flanker Task. *Cogn., Affect. Behav. Neurosci.*, 21, 355–371. doi: 10.3758/s13415-021-00876-y
- Helfrich, R. F., and Knight, R. T. (2016). Oscillatory dynamics of prefrontal cognitive control. *Trends Cogn. Sci.*, 20, 916–930. doi: 10.1016/j.tics.2016.09.007
- Hofweber, J., Marinis, T., and Treffers-Daller, J. (2020). Experimentally induced language modes and regular code-switching habits boost bilinguals’ executive performance: Eviden|CE from a within-subject paradigm. *Front. Psychol.*, 11, 542326. doi: 10.3389/fpsyg.2020.542326
- Jensen, O., Bonnefond, M., and VanRullen, R. (2012). An oscillatory mechanism for prioritizing salient unattended stimuli. *Trends Cogn. Sci.* 16, 200–206. doi: 10.1016/j.tics.2012.03.002
- Jensen, O., and Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: Gating by inhibition. *Front. Hum. Neurosci.* 4, 186. doi: 10.3389/fnhum.2010.00186

- Kielar, A., Meltzer, J. A., Moreno, S., Alain, C., and Bialystok, E. (2014). Oscillatory responses to semantic and syntactic violations. *J. Cogn. Neurosci.*, 26, 2840–2862. doi: 10.1162/jocn_a_00670
- Koller, M., and Stahel, W. A. (2011). Sharpening wald-type inference in robust regression for small samples. *Comput. Stat. Data Anal.* 55, 2504–2515. doi: 10.1016/j.csda.2011.02.014
- Kroll, J. F., Dussias, P. E., Bogulski, C. A., and Kroff, J. R. V. (2012). Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. *Psychol. Learn. Motiv.* 56, 229–262. doi: 10.1016/B978-0-12-394393-4.00007-8
- Kupisch, T., and Rothman, J. (2018). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *Int. J. Bilingualism*, 22, 564–582. doi: 10.1177/1367006916654355
- Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., De Bruin, A., Antfolk, J., et al. (2018). Is bilingualism what it reveals and why it is important for formal linguistic theories. *Psychol. Bull.*, 144, 394–425. doi: 10.1037/bul0000142
- Leivada, E., Westergaard, M., Duñabeitia, J. A., and Rothman, J. (2021). On the phantom-like appearance of bilingualism effects on neurocognition: (How) should we proceed? *Bilingualism: Lang. Cogn.*, 24, 197–210. doi: 10.1017/S1366728920000358
- Lemhöfer, K., and Broersma, M. (2012). Introducing LexTALE: a quick and valid lexical test for advanced learners of English. *Behav. Res. Methods* 44, 325–343. doi: 10.3758/s13428-011-0146-0
- Li, P., Legault, J., and Litcofsky, K. A. (2014). Neuroplasticity as a function of second language learning: anatomical changes in the human brain. *Cortex*, 58, 301–324. doi: 10.1016/j.cortex.2014.05.001
- Lohndal, T., Rothman, J., Kupisch, T., and Westergaard, M. (2019). Heritage language acquisition: what it reveals and why it is important for formal linguistic theories. *Lang. Linguistics Compass* 13, e12357. doi: 10.1111/lnc3.12357
- Luck, S. J., and Kappenman, E. S. (2011). *The Oxford Handbook of Event-Related Potential Components* (Oxford, UK: Oxford university press).
- Luk, G., and Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *J. Cogn. Psychol.* 25, 605–621. doi: 10.1080/20445911.2013.795574
- Luk, G., De Sa, E. R. I. C., and Bialystok, E. (2011). Is there a relation between onset age of bilingualism and enhancement of cognitive control? *Bilingualism: Lang. Cogn.*, 14, 588–595. doi: 10.1017/S1366728911000010
- Marian, V., and Hayakawa, S. (2021). Measuring bilingualism: the quest for a “bilingualism quotient.” *Appl. Psycholinguist.* 42, 527–548. doi: 10.1017/S0142716420000533
- Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *J. Neurosci. Methods*, 164, 177–190. doi: 10.1016/j.jneumeth.2007.03.024
- Mazaheri, A., van Schouwenburg, M. R., Dimitrijevic, A., Denys, D., Cools, R., Jensen, O., et al. (2014). Region-specific modulations in oscillatory alpha activity serve to facilitate processing in the visual and auditory modalities. *NeuroImage* 87, 356–362. doi: 10.1016/j.neuroimage.2013.10.052
- McDermott, T. J., Wiesman, A. I., Proskovec, A. L., Heinrichs-Graham, E., and Wilson, T. W. (2017). Spatiotemporal oscillatory dynamics of visual selective attention during a flanker task. *NeuroImage*, 156, 277–285. doi: 10.1016/j.neuroimage.2017.05.014
- Meisel, J. M. (2011). *First and Second Language Acquisition: Parallels and Differences*. Cambridge University Press.
- Montrul, S. (2016). *The Acquisition of Heritage Languages* (Cambridge: Cambridge University Press).
- Nigbur, R., Ivanova, G., and Stürmer, B. (2011). Theta power as a marker for cognitive interference. *Clin. Neurophysiol.* 122, 2185–2194. doi: 10.1016/j.clinph.2011.03.030
- Oehrn, C. R., Hanslmayr, S., Fell, J., Deuker, L., Kremers, N. A., Do Lam, A. T., et al. (2014). Neural communication patterns underlying conflict detection, resolution, and adaptation. *J. Neurosci.*, 34, 10438–10452. doi: 10.1523/JNEUROSCI.3099-13.2014
- Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.*, 2011, 156869. doi: 10.1155/2011/156869
- Ortega, L. (2013). Ways Forward for a Bi/Multilingual Turn in SLA. In S. May (Ed.), *The Multilingual Turn: Implications for SLA, TESOL, and Bilingual Education*. UK: Routledge), 32–53.
- Paap, K. R., Johnson, H. A., and Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex* 69, 265–278. doi: 10.1016/j.cortex.2015.04.014
- Pascual y Cabo, D., and Rothman, J. (2012). The (il) logical problem of heritage speaker bilingualism and incomplete acquisition. *Appl. Linguist.* 33, 450–455. doi: 10.1093/applin/ams037
- Pastötter, B., Dreisbach, G., and Bäuml, K.-H. T. (2013). Dynamic adjustments of cognitive control: Oscillatory correlates of the conflict adaptation effect. *J. Cogn. Neurosci.* 25, 2167–2178. doi: 10.1162/jocn_a_00474
- Pereira Soares, S. M., Kubota, M., Rossi, E., and Rothman, J. (2021). Determinants of bilingualism predict dynamic changes in resting state EEG oscillations. *Brain Lang.* 223, 105030. doi: 10.1016/j.bandl.2021.105030
- Pliatsikas, C. (2019). “Multilingualism and brain plasticity,” in *The Handbook of the Neuroscience of Multilingualism*, eds J. W. Schwieter, and M. Paradis (Hoboken, NJ: Wiley- Blackwell), 230–251.
- Pliatsikas, C. (2020). Understanding structural plasticity in the bilingual brain: the dynamic restructuring model. *Bilingualism: Lang. Cogn.* 23, 459–471. doi: 10.1017/S1366728919000130
- Pliatsikas, C., DeLuca, V., Moschopoulou, E., and Saddy, J. D. (2017). Immersive bilingualism reshapes the core of the brain. *Brain Struct. Funct.* 222, 1785–1795. doi: 10.1007/s00429-016-1307-9
- Pliatsikas, C., DeLuca, V., and Voits, T. (2020). The many shades of bilingualism: Language experiences modulate adaptations in brain structure. *Lang. Learn.* 70, 133–149. doi: 10.1111/lang.12386
- Pliatsikas, C., Pereira Soares, S. M., Voits, T., DeLuca, V., and Rothman, J. (2021). Bilingualism is a long-term cognitively challenging experience that modulates metabolite concentrations in the healthy brain. *Sci. Rep.* 11, 1–12. doi: 10.1038/s41598-021-86443-4
- Polinsky, M. (2018). *Heritage Languages and their Speakers* (Cambridge, UK: Cambridge University Press).
- Polinsky, M., and Scontras, G. (2020). Understanding heritage languages. *Bilingualism: Lang. Cogn.* 23, 4–20. doi: 10.1017/S1366728919000245
- Prat, C. S., Yamasaki, B. L., Kluender, R. A., and Stocco, A. (2016). Resting-state qEEG predicts rate of second language learning in adults. *Brain Lang.* 157, 44–50. doi: 10.1016/j.bandl.2016.04.007
- Prat, C. S., Yamasaki, B. L., and Peterson, E. R. (2019). Individual differences in resting-state brain rhythms uniquely predict second language learning rate and willingness to communicate in adults. *J. Cogn. Neurosci.* 31, 78–94. doi: 10.1162/jocn_a_01337
- Prystauka, Y., and Lewis, A. G. (2019). The power of neural oscillations to inform sentence comprehension: A linguistic perspective. *Lang. Linguist. Compass* 13, e12347. doi: 10.1111/lnc3.12347
- Pscherer, C., Bluschke, A., Mückschel, M., and Beste, C. (2021). The interplay of resting and inhibitory control-related theta-band activity depends on age. *Human Brain Mapp.* 42, 3845–3857. doi: 10.1002/hbm.25469
- R Core Team (2021). *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria*. Available online at: <https://www.R-project.org/>
- Riečanský, I., and Katina, S. (2010). Induced EEG alpha oscillations are related to mental rotation ability: The evidence for neural efficiency and serial processing. *Neurosci. Lett.* 482, 133–136. doi: 10.1016/j.neulet.2010.07.017
- Rothman, J. (2009). Understanding the nature and outcomes of early bilingualism: romance languages as heritage languages. *Int. J. Bilingualism* 13, 155–163. doi: 10.1177/1367006909339814
- Rothman, J., Alonso, J. G., and Puig-Mayenco, E. (2019). *Third Language Acquisition and Linguistic Transfer* (Cambridge, UK: Cambridge University Press). doi: 10.1017/9781316014660
- Rothman, J., and Treffers-Daller, J. (2014). A Prolegomenon to the Construct of the Native Speaker: Heritage Speaker Bilinguals are Natives Too! *Appl. Linguist.* 35, 93–98. doi: 10.1093/applin/amt049
- Salig, L. K., Valdés Kroff, J. R., Slevc, L. R., and Novick, J. M. (2021). Moving from Bilingual Traits to States: understanding cognition and language processing through moment-to-moment variation. *Neurobiol. Lang.* 1–62. doi: 10.1162/nol_a_00046
- Siegel, M., Engel, A. K., and Donner, T. H. (2011). Cortical network dynamics of perceptual decision-making in the human brain. *Front. Human Neurosci.* 5, 21. doi: 10.3389/fnhum.2011.00021
- Stocco, A., Yamasaki, B., Natalenko, R., and Prat, C. S. (2014). Bilingual brain training: a neurobiological framework of how bilingual experience improves executive function. *Int. J. Bilingualism*, 18, 67–92. doi: 10.1177/1367006912456617

- Surrain, S., and Luk, G. (2019). Describing bilinguals: a systematic review of labels and descriptions used in the literature between 2005–2015. *Bilingualism: Lang. Cogn.* 22, 401–415. doi: 10.1017/S1366728917000682
- Suzuki, K., Okumura, Y., Kita, Y., Oi, Y., Shinoda, H., Inagaki, M., et al. (2018). The relationship between the superior frontal cortex and alpha oscillation in a flanker task: Simultaneous recording of electroencephalogram (EEG) and near infrared spectroscopy (NIRS). *Neurosci. Res.* 131, 30–35. doi: 10.1016/j.neures.2017.08.011
- Tao, L., Marzecová, A., Taft, M., Asanowicz, D., and Wodniecka, Z. (2011). The efficiency of attentional networks in early and late bilinguals: the role of age of acquisition. *Front. Psychol.*, 2, 123. doi: 10.3389/fpsyg.2011.00123
- Thut, G., Miniussi, C., and Gross, J. (2012). The functional importance of rhythmic activity in the brain. *Curr. Biol.* 22, R658–R663. doi: 10.1016/j.cub.2012.06.061
- Van den Noort, M., Struys, E., Bosch, P., Jaswetz, L., Perriard, B., Yeo, S., et al. (2019). Does the bilingual advantage in cognitive control exist and if so, what are its modulating factors? A systematic review. *Behav. Sci.* 9, 27. doi: 10.3390/bs9030027
- Van Diepen, R. M., Foxe, J. J., and Mazaheri, A. (2019). The functional role of alpha-band activity in attentional processing: The current zeitgeist and future outlook. *Curr. Opin. Psychol.*, 29, 229–238. doi: 10.1016/j.copsyc.2019.03.015
- Varela, F., Lachaux, J.-P., Rodriguez, E., and Martinerie, J. (2001). The brainweb: Phase synchronization and large-scale integration. *Nat. Rev. Neurosci.* 2, 229–239. doi: 10.1038/35067550
- Weekes, B. S., Abutalebi, J., Mak, H. K. F., Borsa, V., Soares, S. M. P., Chiu, P. W., et al. (2018). Effect of monolingualism and bilingualism in the anterior cingulate cortex: a proton magnetic resonance spectroscopy study in two centers. *Letras de Hoje.* 53, 5–12. doi: 10.15448/1984-7726.2018.1.30954
- Yee, J., DeLuca, V., and Pliatsikas, C. (2022). “The effects of multilingualism on brain structure, language control and language processing: insights from MRI,” in *The Cambridge Handbook of Third Language Acquisition and Processing*, eds J. Cabrelli, A. Chaouch-Orozco, J. González Alonso, S. M. Pereira Soares, E. Puig-Mayenco, and J. Rothman (Cambridge, UK: Cambridge University Press).
- Yohai, V. J. (1987). High breakdown-point and high efficiency robust estimates for regression. *Ann. Stat.* 15, 642–656. doi: 10.1214/aos/1176350366
- Zhuang, P., Toro, C., Grafman, J., Manganotti, P., Leocani, L., Hallett, M., et al. (1997). Event-related desynchronization (ERD) in the alpha frequency during development of implicit and explicit learning. *Electroencephalogr. Clin. Neurophysiol.* 102, 374–381. doi: 10.1016/S0013-4694(96)90630-7



OPEN ACCESS

EDITED BY

Joshua M. Carlson,
Northern Michigan University,
United States

REVIEWED BY

Lucy J. Troup,
University of the West of Scotland,
United Kingdom
Jacob Aday,
University of California, San Francisco,
United States

*CORRESPONDENCE

Yanmei Wang
ymwang@psy.ecnu.edu.cn

SPECIALTY SECTION

This article was submitted to
Cognitive Neuroscience,
a section of the journal
Frontiers in Human Neuroscience

RECEIVED 21 May 2022

ACCEPTED 28 June 2022

PUBLISHED 29 July 2022

CITATION

Wang Y, Tang Z, Zhang X and Yang L
(2022) Auditory and cross-modal
attentional bias toward positive natural
sounds: Behavioral and ERP evidence.
Front. Hum. Neurosci. 16:949655.
doi: 10.3389/fnhum.2022.949655

COPYRIGHT

© 2022 Wang, Tang, Zhang and Yang.
This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License](#)
(CC BY). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Auditory and cross-modal attentional bias toward positive natural sounds: Behavioral and ERP evidence

Yanmei Wang^{1,2*}, Zhenwei Tang¹, Xiaoxuan Zhang^{1,2} and Libing Yang¹

¹Shanghai Key Laboratory of Mental Health and Psychological Crisis Intervention, School of Psychology and Cognitive Science, East China Normal University, Shanghai, China, ²Shanghai Changning Mental Health Center, Shanghai, China

Recently, researchers have expanded the investigation into attentional biases toward positive stimuli; however, few studies have examined attentional biases toward positive auditory information. In three experiments, the present study employed an emotional spatial cueing task using emotional sounds as cues and auditory stimuli (Experiment 1) or visual stimuli (Experiment 2 and Experiment 3) as targets to explore whether auditory or visual spatial attention could be modulated by positive auditory cues. Experiment 3 also examined the temporal dynamics of cross-modal auditory bias toward positive natural sounds using event-related potentials (ERPs). The behavioral results of the three experiments consistently demonstrated that response times to targets were faster after positive auditory cues than they were after neutral auditory cues in the valid condition, indicating that healthy participants showed a selective auditory attentional bias (Experiment 1) and cross-modal attentional bias (Experiment 2 and Experiment 3) toward positive natural sounds. The results of Experiment 3 showed that N1 amplitudes were more negative after positive sounds than they were after neutral sounds, which further provided electrophysiological evidence that positive auditory information enhances attention at early stages in healthy adults. The results of the experiments performed in the present study suggest that humans exhibit an attentional bias toward positive natural sounds.

KEYWORDS

attentional bias, positive natural sounds, emotional cueing task, attention enhancement, event-related potentials

Introduction

Emotional attention refers to the tendency to give preference to processing emotional information over neutral information (Vuilleumier, 2005; Yiend, 2010; Pourtois et al., 2013; Gerdes et al., 2020). The ability to detect and respond to threats in the natural world is of vital importance for adaptive behavior, survival across species

(LeDoux, 1996; Silston and Mobbs, 2018; Kreutzmann et al., 2020). Whereas danger signals predict the onset of a potentially threatening event, safety signals indicate its non-occurrence, thereby inhibiting fear and stress responses. Organisms evolved specific nervous systems (e.g., amygdala and hypothalamus), associative learning processes (e.g., Pavlovian conditioning) or cognitive mechanisms that rapidly direct attention to threatening information and keep our attention focused on a potential threat as long as needed. Since threat-related attentional processes have fundamental survival value (Mathews and Mackintosh, 1998; Mogg and Bradley, 1998), and studies on threat-related attentional processes will be very helpful to reveal the cause of anxiety disorders (Sheppes et al., 2013). Most studies that explore the attentional bias of emotional information have focused on the cognitive and neural mechanisms of attentional bias toward threatening stimuli or negative information through the indirect behavioral index (e.g., response times) or the direct physiological/neurological index (e.g., eye movement, event-related brain potentials) (Dong et al., 2017; Schmidtdorf et al., 2018; Berggren and Eimer, 2021). In addition, a large number of studies have investigated attentional bias toward negative information in healthy individuals as well as in participants experiencing a variety of anxiety disorders (Bar-Haim et al., 2007). Extensive research shows that negative attentional bias is closely related to the formation and maintenance of anxiety symptoms (Bar-Haim et al., 2007; Van Bockstaele et al., 2014; Ren et al., 2020; Salahub and Emrich, 2020; Basanovic et al., 2021). The components of negative attentional bias can be divided into two mechanisms: (a) initial attentional orientation toward orienting to threatening stimuli (engagement bias or attention capture); (b) difficulty disengaging attention from threatening stimuli (disengagement bias) (Posner et al., 1987; Cisler and Koster, 2010). Initial attentional orientation influences attention selection during early, automatic processing stages, which occur before 150 ms, and disengaging attention from stimuli occurs after the stimulus has been selected during late processing stages, which occur after 250 ms (Weierich et al., 2008; Pool et al., 2016; Wang et al., 2019b; Gupta et al., 2021; Yuan et al., 2021).

Attentional bias for positive emotional stimuli

An increasing number of researchers have expanded the investigation into attentional biases toward positive emotional stimuli in healthy individuals (Wadlinger and Isaacowitz, 2008; Sali et al., 2014; Pool et al., 2016). Positive attentional bias means that individuals will preferentially pay more attention to positive stimuli or rewarding information than neutral stimuli (Brosch et al., 2008a; Wadlinger and Isaacowitz, 2008; Pool et al., 2016). Positive rewarding stimuli have been defined as stimuli that have a positive hedonic value that might elicit approach behaviors.

Positive rewarding and negative threatening stimuli are both affectively relevant; positive attentional bias thus has very important evolutionary implications for an organism's survival (Schultz, 2004; Anderson, 2016). If individuals prefer positive, rewarding information, they are more likely to recognize future benefits and have more positive beliefs and attitudes toward uncertainty, which can help them obtain survival resources or promote social status. Individuals experiencing positive attention biases may use selective attention as a tool to regulate their emotional experience during stressful circumstances and maintain a positive and stable mood, resulting in better mental health. Attention toward happy faces was positively correlated with positive mood and life satisfaction (Sanchez and Vazquez, 2014). A great deal of research has used different types of positive visual stimuli to explore positive attentional bias, including pictures of baby faces (Brosch et al., 2008a), happy faces (Joormann and Gotlib, 2007; Sanchez and Vazquez, 2014), pictures of food (Brignell et al., 2009; Tapper et al., 2010), positive words (Grafton and MacLeod, 2017), and stimuli features that are associated with reward outcomes (Anderson and Yantis, 2012, 2013; Pool et al., 2014; Sali et al., 2014). These studies consistently reported that, compared with neutral information, healthy individuals displayed attentional bias toward positive information. A recent meta-analysis systematically compared attentional biases for positive and neutral stimuli across 243 studies and showed that attention bias for positive stimuli among healthy individuals was larger in paradigms that measure early rather than late attentional processing, which suggests that positive attentional bias occurs rapidly and involuntarily (Pool et al., 2016). Moreover, it was also demonstrated that older adults exhibited attentional bias toward positive pictures at earlier stages in attention processing (Kennedy et al., 2020). However, another study investigated the time course of attentional bias toward positive visual stimuli (happy faces) using dot-probe paradigm, and revealed that happy faces were also associated with delayed disengagement during later stages of attentional processing (Torrence et al., 2017). Considering these inconsistent findings, it remains unclear at which stage positive emotion stimuli impacts attentional processing. Therefore, one of the purposes of the present study was to explore whether initial attention capture or subsequent disengagement was related to attentional bias toward positive information.

Attentional bias toward emotional auditory information

Humans can accurately perceive the world and navigate in complex environments, which can contain several different emotionally relevant (or irrelevant) cues from different sensory modalities. Vision and audition represent two important senses needed to navigate through space and time (Bell et al., 2019;

Choiniere et al., 2021; Peterson et al., 2022). It might be because the available, present-day technology is better suited for studying vision than for studying other modalities, there are more research on vision than on any other sensory modality (Hutmacher, 2019). Similarly, there were more studies on attentional bias toward visual emotional stimuli than on attentional bias toward emotional auditory information (e.g., Sheppes et al., 2013; Torrence et al., 2017). In fact, auditory sounds also convey emotionally relevant information and interact with vision to provide an appropriate judgment of the emotional qualities of a situation (Gerdes et al., 2014, 2020; Concina et al., 2019). Natural sounds in everyday life, such as screams, moans, cheers, applause, barks, and chirps, often contain emotional information and can carry biologically significant emotional information (Armony and LeDoux, 2010; Lepping et al., 2019; Wang et al., 2019b). For example, when we walk through busy streets, an alarm whistle can be a threatening cue that prompts an individual to make avoidance responses. Similarly, when we take a walk in the woods, the musical sounds made by birds can be rewarding signals and activate an organism to make approach-related behavior. If you hear the familiar call of friends, you will immediately look toward them with a smile. In other words, sounds serve as warnings or rewarding signals in daily life that can be critical for survival and fine-tune our actions (Harrison and Davies, 2013; Gerdes et al., 2020; McDougall et al., 2020). A recent study has investigated auditory attentional bias employing white noise as negative auditory cue (Wang et al., 2019a), but white noise lacks ecological validity. In fact, we need to detect and process natural sounds from real world more frequently in our everyday life (e.g., a barking sound of a dog or the sound of applause), resulting in avoidance-related or approach-related behaviors. Since natural sounds have been proved to have good ecological validity (Hu et al., 2017; Młynarski and McDermott, 2019), some researchers have suggested that studies that incorporate natural sounds should be encouraged (Sutherland and Mather, 2012; Koumura et al., 2019; Zhai et al., 2020; Zuk et al., 2020). Therefore, the present study aims to investigate whether participants exhibited attentional bias toward emotional natural sounds (positive and negative natural sounds).

The cross-modal attentional bias toward emotional auditory information

In everyday life, precise processing of temporal and spatial information needs to consider cross-modal interactions in endogenous spatial attention between vision and audition. Selective attention has traditionally been studied separately for different sensory modalities, with little direct contact between traditional research on “visual attention” or “auditory attention.” The issue of cross-modal interactions in spatial attention between vision and audition has been addressed

only more recently (Van Vleet and Robertson, 2006; Ariei and Marks, 2008; Blurton et al., 2015). Furthermore, recent studies have shown that negative auditory cues guide auditory or visual spatial allocation of attention (Peschard et al., 2017; Carlson et al., 2018; Wang et al., 2019a; Gerdes et al., 2020) or influence semantic processing (Gao et al., 2020). It has been found that emotional visual cues also modulate auditory spatial attention (Harrison and Woodhouse, 2016). This growing body of empirical evidence has proven that healthy participants exhibit auditory or cross-modal attentional bias toward negative auditory information. However, most studies have focused on whether or not positive visual stimuli modulated visual spatial attention (e.g., Pool et al., 2016), so it remains unclear whether positive natural sounds could modulate auditory or visual spatial allocation of attention. Therefore, the purpose of the present study was twofold: (a) to examine auditory and cross-modal attentional bias toward positive auditory information using natural sounds as positive cues; (b) to explore the cognitive and neural mechanisms underlying auditory and cross-modal positive attentional bias.

The mechanism of attentional bias toward emotional information

Attentional bias toward negative information is thought to be caused by two components: facilitated engagement with negative information (involuntary attentional capture by negative stimuli) and delayed disengagement from negative information (volitional delayed disengagement from negative information) (Bar-Haim et al., 2007). The former component is called engagement bias, which refers to the quick detection of threat stimuli as opposed to non-threat stimuli (Cisler et al., 2009). The latter component is disengagement bias, which refers to the fact that it is harder to direct attention away from threatening stimuli (Fox et al., 2001, 2002). A threat detection mechanism likely underlies facilitated attention, reflecting an automatic processing. Delayed disengagement from negative information might be related to attention control abilities or emotion regulation goals, reflecting a top-down processing (Cisler and Koster, 2010; Park et al., 2013; Basanovic et al., 2021). Similarly, the components of positive attentional bias might also be linked to two factors: facilitated attention for positive information, which occurs at an early stage, and disengagement from reward information, which might occur at a later stage of attention processing (Pool et al., 2016). According to evolutionary theory, both attentional engagement with positive information and maintenance of attention with positive information underlie attentional bias for positive stimuli. Orienting to positive stimuli automatically helps an individual detect reward information in the environment rapidly, pursue potential survival resources, and improve social status (Strauss and Allen, 2006; Vuilleumier, 2015; Gupta, 2019;

Gutiérrez-Cobo et al., 2019; Lockhofen et al., 2021). Delayed disengagement from positive information helps individuals keep their attention on positive stimuli to keep a positive and optimistic mood, accomplishing the goal of emotion regulation and mental health (Wadlinger and Isaacowitz, 2008; Demeyer and Raedt, 2013; Thoern et al., 2016; Booth and Sharma, 2020; Wadley et al., 2020). In the present study, we investigated these two possible explanations while comparing the early and later attentional mechanisms that underlie auditory/cross-modal attentional bias toward positive stimuli.

The present study

To reveal the subcomponents of attentional bias toward emotional stimuli, researchers have created several experimental paradigms centered on classical cognitive tasks. These experimental paradigms included the dot-probe task (e.g., Yiend, 2010), visual search task (e.g., Wieser et al., 2018), free viewing task (e.g., Dong et al., 2017; Navalón et al., 2021), emotional Stroop task (e.g., Phaf and Kan, 2007; Kaiser et al., 2017), and emotional cueing task (e.g., Victor et al., 2020). The emotional spatial cueing paradigm was modified from the exogenous cue-target task (Fox et al., 2001). Among these experimental paradigms, the emotional spatial cueing paradigm has been specifically designed to simultaneously measure initial orientation and difficulty with disengagement through a simple behavioral index (i.e., response latency) (Posner and Cohen, 1984; Pool et al., 2016; Preciado et al., 2017; Blicher and Reinholdt-Dunne, 2019; Wang et al., 2019a). However, recently, researchers have added a condition with two neutral cues furnishing a neutral baseline in the dot probe detection task (e.g., Carlson and Reinke, 2014), such modification becomes possible to measure initial orienting and later disengagement with the dot-probe task as well. Comparison of such a baseline with trials in which the target appears at the same location as the emotional cues reflects initial orienting, whereas comparison of the baseline with trials in which the target appears at the location opposite to the emotional cue reflects attention disengagement.

Some researchers also have proposed that using appropriate emotional stimuli and a behavior indicator could help reveal engagement bias and disengagement bias through the emotional spatial cueing paradigm (Fox et al., 2002; Schwerdtfeger and Derakshan, 2010; Imhoff et al., 2019; Wang et al., 2019a). In this task, participants direct their attention to a fixation point at the center of a screen. A cue (emotional or neutral) randomly appears on one side of the fixation point. Shortly after the cue offset, a target is presented either in the same location as the cue or on the opposite side. Participants are instructed to identify the location of the target as quickly and accurately as possible. Engagement bias is indicated through the facilitated detection of targets presented in the location occupied by an emotional cue (valid condition). Disengagement bias is indicated by slower

responses to targets presented on the opposite side to an emotional cue (invalid condition) compared with responses to a neutral cue. Although the reliability of the emotional cueing paradigm has not been assessed, the reliability of other tasks that measure attentional bias toward threatening information (such as the dot-probe task and emotional Stroop task) was proven to be low (Van Bockstaele et al., 2014). Therefore, in this study, we also adopted the emotional cueing task to explore the subcomponents of attentional bias toward positive auditory stimuli in healthy individuals.

To date, although most studies on attentional biases have applied behavioral paradigms, event-related brain potentials (ERPs) are particularly suited for examining attentional biases. ERPs can provide a temporally precise, direct measure of emotional attention and may detect subcomponents of positive attentional bias that are not evident in behavioral data. Recently, a growing number of studies have adopted high temporal resolution event-related potential techniques (ERP) to examine the neural correlates of attentional bias toward emotional information (Keil et al., 2007; Hao et al., 2015; Reutter et al., 2017; Berggren and Eimer, 2021; Carlson, 2021; Schindler et al., 2021). These studies reported that some ERP components were involved in the neural mechanisms of attentional bias toward emotional negative information or threatening stimuli. Specifically, evidence showed that negative faces amplified the N170 ERP component (Schindler et al., 2021), evoked larger N2pc component (Berggren and Eimer, 2021), and increased late positive potential (LPP) (Schindler et al., 2021) in healthy individuals. Also, some studies revealed that higher (i.e., more negative) N2pc amplitude was found for angry faces in social anxiety (Reutter et al., 2017; Wieser et al., 2018) or disgust faces (Yuan et al., 2019). Using visual oddball task, researchers found that previous depression was uniquely associated with greater P3 ERP amplitude following sad targets, reflecting a selective attention bias toward negative faces (Bistricky et al., 2014). Some research groups have suggested that both healthy and anxious populations displayed modulations of early ERP components, including the P1, N170, and N2pc, in response to threatening and emotional stimuli, suggesting that both typical and abnormal patterns of attentional bias were characterized by enhanced allocation of attention to threat and emotion at earlier stages of processing, and modulations of later components, such as the P3, reflecting conscious and evaluative processing of threat and emotion and disengagement difficulties at later stages of processing (Torrence and Troup, 2018; Gupta et al., 2019; Carlson, 2021). However, few studies have revealed the dynamic time course of attentional bias for positive auditory stimuli in healthy adults. It remains unclear whether modulations of early ERP components or late ERP components were associated with attentional bias toward positive natural sounds in healthy adults. Because positive emotions and negative emotions have different evolutionary functions for the survival of humans, it is important to examine the neural mechanism of attention bias for positive

emotional information. To date, only one study has explored the exact time course of early attention allocation toward positive visual stimuli using a dot-probe paradigm in healthy participants, and it found that N1 amplitudes after positive pictures were enhanced compared to those that followed after negative pictures, indicating enhanced attention engagement with positive visual information (Pintzinger et al., 2017). In addition, one recent study revealed that the amplitude of the N1 component was enlarged by emotional sounds relative to that evoked by neutral sounds (Folyi et al., 2016). Another study revealed that patients with internet gaming disorder (IGD) exhibited increased attentional bias toward visual gaming-related cues; specifically, higher LPP amplitudes were found for game-related cues in the IGD group (Kim et al., 2018). Therefore, one goal of the present study was to investigate the neural mechanism of attentional bias toward positive natural sounds employing event-related potential techniques.

With the above in mind, in three experiments, we adopted emotional natural sounds as cues and auditory stimuli (Experiment 1) or visual stimuli (Experiment 2) as targets to explore auditory attentional bias (Experiment 1) or cross-modal attentional bias (Experiment 2) toward positive sounds using an auditory emotional cueing paradigm. Considering that the ERP approach is one method that can be used to better understand the time course of attentional bias (Torrence and Troup, 2018), Experiment 3 was conducted to investigate the neural mechanism of cross-modal attentional bias toward positive sounds in healthy participants using event-related potential techniques. Our expectation was that visual attention would be preferentially oriented toward positive sounds rather than neutral sounds. For the ERP results, we hypothesized that some earlier components fell on the electrophysiological index of enhancement attention by positive auditory stimuli, including N1, N2pc, etc. The present study will provide neural evidence for the time course of attentional bias toward positive auditory information.

Experiment 1

Method

Participants

Fifty-four right-handed undergraduate students ($N_{\text{female}} = 42$, 18–24 years, $M_{\text{age}} = 20.73$ years, $SD_{\text{age}} = 2.91$) were recruited for monetary compensation. A power analysis (G*Power, Version 3.1) (Faul et al., 2007) estimated that 36 participants would be needed to achieve a power of 0.95 ($f = 0.25$, $\alpha = 0.05$, $\beta = 0.95$). We recruited a total of 58 participants, of which four participants were excluded from analyses for not appropriately following task instructions, resulting in a final sample of 54 participants. All participants had normal hearing and normal or corrected-to-normal

vision. All participants reported no history of neurological or psychological disorders. The present study was approved by the local research ethics committee (HR 310-2019). Each participant signed an informed consent form before the experiment.

Materials

Ten positive sounds, ten negative sounds, and ten neutral sounds that differed in valence norms were selected from the expanded version of the International Affective Digitized Sounds system (IADS-E, Yang et al., 2018) and presented *via* headphones (see Appendix A). The sounds were determined by using the Self-Assessment Manikin (SAM; Bradley and Lang, 1994). The criteria for choosing the sounds consisted of pleasant valence (mean SAM valence norm score- > 6.00) and neutral valence ($3.00 < \text{mean SAM valence norm score-} < 5.00$) on a 1–9 scale. The sounds were from the natural environment and had high ecological validity (see Appendix A for details of the selected IADS-E sounds). In the IADS-E system, these sounds last 6000 ms (consisting of three segments of the same sound repetition with a length of 2000 ms each). A 2000-ms segment was selected from each original sound on the basis that it was representative of the emotional content of the original sounds. The materials were rated for valence and arousal by an independent group of participants ($n = 49$) using 9-point rating scales (valence: 1 = very unpleasant, 9 = very pleasant; arousal: 1 = not at all arousing, 9 = very arousing). The results showed that positive sounds (mean valence = 6.94 ± 0.75) were rated as more pleasant than the neutral sounds (mean valence = 4.84 ± 0.60), $t(48) = 5.68$, $p < 0.001$, and negative sounds (mean valence = 2.96 ± 0.57) were rated as more unpleasant than the neutral sounds (mean valence = 4.84 ± 0.60). There was no significant difference between positive sounds (mean arousal = 6.13 ± 0.44) and negative sounds (mean arousal = 6.22 ± 0.81) with respect to arousal ratings, $t(48) = 1.02$, $p = 0.34$. Both positive and negative sounds were rated as higher arousal than neutral sounds (mean arousal = 4.78 ± 0.62).

The auditory target stimulus was a 100 ms neutral tone (150 Hz, pure tone “beep”). The mean dB level for all sounds was 65 dB, measured at the participants’ ear. Auditory stimuli were presented randomly to either the left or right ear using EDIFIER W800BT headphones. E-Prime 2.0 was used to control the experiment.

Procedures

Participants were seated 60 cm from the screen and completed twenty practice trials first using auditory stimuli not included in the main experiment. Each trial began with a central fixation cross that lasted 750 ms. After the offset of fixation, an auditory cue (a positive sound, a negative sound or a neutral sound) was presented for 2000 ms. The sound duration was chosen to ensure that the emotional content of the sound was fully processed by the participants before the target appeared.

Next, a neutral auditory target (pure tone) appeared for 100 ms. Participants were required to press the 'F' or 'J' key if the auditory target appeared on the left or right, respectively. If the auditory target appeared on the same side as the auditory cue, the auditory cue was regarded as valid; otherwise, the auditory cue was invalid. Participants had 2000 ms to respond after onset of the target. After response or after 2000 ms in the event of no response, an inter-trial interval of 2000 ms took place. All trials were presented in randomized order with an equal proportion of validly cued and invalidly cued trials (50%). In a validly cued trial, the target appeared on the side of the sound (positive/neutral); in an invalidly cued trial, it appeared on the opposite side of the sound (see **Figure 1A**).

We had a 3 (auditory cue valence: positive, negative vs. neutral) \times 2 (cue validity: valid cues vs. invalid cues) within-subjects design. Each of these conditions was repeated 60 times, resulting in 360 trials. The experiment contained two blocks. Each block comprised 180 trials: 30 positive-valid trials, 30 positive-invalid trials, 30 neutral-valid trials, 30 neutral-invalid trials, 30 negative-valid trials, and 30 negative-invalid trials. Each block was separated by a short rest break. All stimuli were randomly assigned to each condition, and the assignment was counterbalanced between participants.

Results

The mean error rate was 1.07%. These results were not analyzed further. Only reaction times on trials with correct responses were included in the analysis. Reaction times less than 150 ms and greater than 1500 ms were eliminated to eliminate premature responding to eliminate premature responding. RTs were submitted to a 3 (cue valence: positive vs. neutral vs. negative) \times 2 (cue validity: valid cues vs. invalid cues) ANOVA. The main effect of cue validity was not significant, $F(1, 53) = 0.14$, $p = 0.71$, and there was no significant main effect of cue valence, $F(2, 106) = 1.94$, $p = 0.15$. Importantly, the crucial Cue Valence \times Cue Validity interaction was significant, $F(2, 106) = 12.74$, $p < 0.001$, $\eta_p^2 = 0.20$. Further simple effect analysis of the two-way interaction revealed that RTs were shorter for positive auditory cues ($M = 505.22$ ms, $SD = 101.66$ ms) and negative auditory cues ($M = 524.73$ ms, $SD = 103.59$ ms) compared with neutral auditory cues ($M = 544.31$ ms, $SD = 117.77$ ms) in the valid condition. In the invalid condition, there was no significant difference between positive auditory cues ($M = 525.57$ ms, $SD = 117.71$ ms) and neutral auditory cues ($M = 518.25$ ms, $SD = 109.89$ ms), $t(53) = 0.82$, $p = 0.42$, negative auditory cues ($M = 534.86$ ms, $SD = 105.49$ ms) and neutral auditory cues ($M = 518.25$ ms, $SD = 109.89$ ms), $t(53) = 1.58$, $p = 0.12$ (see **Figure 1B**). For neutral cues, response times were slower in the valid condition ($M = 544.31$ ms, $SD = 117.77$ ms) relative to the invalid

condition ($M = 518.25$ ms, $SD = 109.89$ ms), $t(53) = 4.79$, $p < 0.001$, which suggested that the IOR (inhibition of return) effect was observed. For positive cues, response times were faster to valid cues ($M = 505.22$ ms, $SD = 101.66$) relative to invalid cues ($M = 525.57$ ms, $SD = 117.71$), $t(53) = -2.47$, $p < 0.05$.

Then, we calculated the IOR effect by subtracting the mean reaction time on invalid trials from the mean reaction time on valid trials. This meant that a positive IOR index indicated a reluctance to return attention to the previously attended location, while a negative IOR index indicated a facilitation effect for the valid location and the absence of IOR. To compare the magnitude of IOR effect for different auditory cues, a repeated-measures was performed for the IOR index. This analysis showed a significant main effect of cue valence, $F(2, 106) = 13.11$, $p < 0.001$, $\eta_p^2 = 0.38$. Follow-up analyses of the main effect indicated that both of the IOR effects for positive auditory cues ($M = -22.46$ ms, $SD = 8.29$ ms) and the IOR effects for negative auditory cues ($M = -7.54$ ms, $SD = 6.74$ ms) were significantly smaller compared to neutral auditory cues ($M = 26.06$ ms, $SD = 5.44$ ms). Taken together, the IOR effect was reduced when cues were natural positive sounds, which suggested that participants displayed positive attentional bias toward auditory information from the natural environment.

Discussion

Experiment 1 preliminarily explored attentional bias toward positive auditory stimuli from the natural environment in healthy participants and found that, like negative auditory sounds, positive auditory sounds also guide auditory spatial attention toward neutral sounds. The IOR effect was reduced more for positive auditory cues than neutral auditory cues. Furthermore, RTs were faster for positive natural sounds than for neutral sounds in valid condition. Given that the positive sounds are 2000 ms long, it is unclear whether these results indicate rapid capture, as there was a great deal of time after positive sound onsets during which engagement may begin. Therefore, we only concluded that healthy participants also displayed positive attentional bias toward positive natural sounds. In a recent study, researchers have revealed that healthy participants exhibited attentional bias toward aversive auditory stimuli (Wang et al., 2019a). The results of Experiment 1 further demonstrated the existence of attentional bias toward positive natural sounds in healthy individuals.

The above results are in agreement with previous studies on visual modal spatial cueing effects, which demonstrated that pleasant pictures influence the allocation of visual attention (Brosch et al., 2008a; Gable and Harmon-Jones, 2010; Peters et al., 2016; Pool et al., 2016; Kennedy et al., 2020). Our findings also provide evidence that confirms appraisal theories

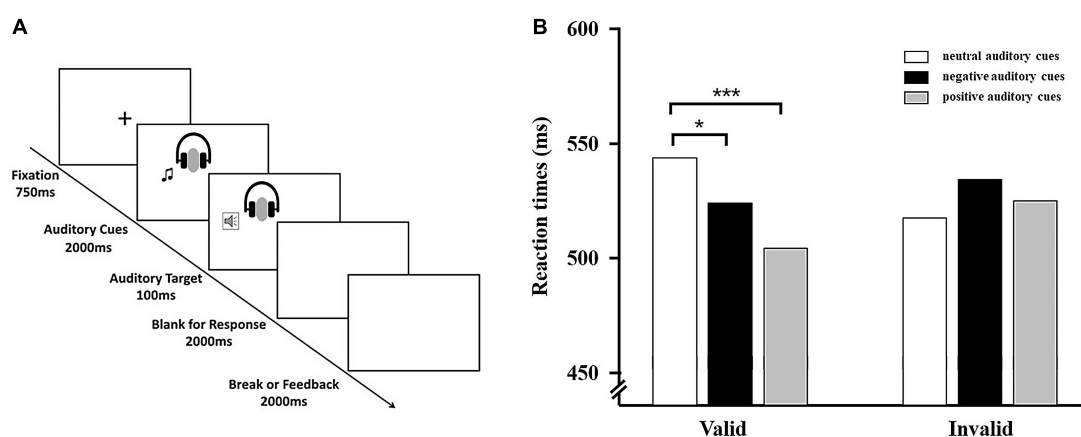


FIGURE 1

(A) Typical sequence of Experiment 1. Shown in the figure is an example of a valid trail. Emotional cues (positive, negative, or neutral) were presented equally often to the right or left ear and were followed by the target on the same (valid, 50%) or opposite side of space where the cue was presented (invalid, 50%). (B) Behavioral results from Experiment 1. Average reaction times for the six experimental conditions. Error bars indicate standard error (SE) of the mean. * $p < 0.05$; *** $p < 0.001$.

of emotion, which suggest that attention will be captured by stimuli that are relevant for the needs, goals, and well-being of the individual, irrespective of stimuli valence (Ellsworth and Scherer, 2003; Sander et al., 2005).

A cueing validity effect was found in emotional auditory conditions, while an inhibition of return effect (IOR) was found in the neutral auditory condition. Some research group has been investigated the neural correlates of attentional bias toward disgusting sound specifically in recent years (Zimmer et al., 2015, 2016). In their previous study, attention bias toward disgusting sounds was investigated using 1000 ms emotional sounds (aversive sounds vs. neutral sounds) as a cue, and it was found that an inter-stimulus interval (ISI) at 650–750 ms after neutral sound cues could induce IOR effects, but 1000 ms aversive sound cues plus an ISI at 50–750 ms was still found to induce cueing validity effects that prompted aversive sounds (Zimmer et al., 2019). Our results extended the previous findings that showed long SOA leads to IOR effects on neutral sounds, but the results of cueing validity effects on emotional sounds were still expanded to the longer stimulus onset asynchrony (SOA).

Experiment 1 investigated positive auditory bias in a single modality (auditory modality). However, in real-life environments, in addition to receiving the stimulation of a single sensory modality (such as the visual or the auditory channel), humans encounter simultaneous input from several different senses, such as vision and audition, when perceiving surrounding environmental changes, making appropriate behavioral responses, and coping with the complicated and changeable environment. Positive natural sounds also may carry biologically relevant emotional information that is useful to the mental and/or physical health of the individual. It remains unclear whether visual spatial attention is preferentially oriented

toward positive natural sounds and how the involvement of each of the two components of attention bias subcomponent (engagement bias and disengagement bias) impacts cross-modal positive attention. Thus, Experiment 2 was conducted to investigate how positive sounds modulate visual spatial attention and aimed to further demonstrate the cross-modal attentional bias toward positive natural sounds using a modified audio-visuospatial cueing task.

Experiment 2

Method

Participants

Based on prior studies on the modulation of spatial attention to visual targets by auditory stimuli (Harrison and Davies, 2013; Wang et al., 2019a; Evans, 2020) and a power analysis using G*Power Version 3.1.9.2 (Faul et al., 2007) conducted for a repeated measurement ANOVA to detect a medium interaction effect of $f = 0.25$ with a statistical power of 0.95 and a significance level of 0.05, we aimed to recruit a sample size that included a minimum of 36 participants.

Sixty-one participants participated in Experiment 2 after giving informed consent. The results of two participants were excluded from analyses because they did not follow task instructions appropriately, resulting in a final sample of 59 participants (mean age = 20.81 ± 4.31 years, 13 male). All subjects were right-handed with normal hearing and normal or corrected-to-normal vision. All participants reported no history of neurological or psychiatric disorders. The experiment was approved by the Ethics Committee of East China Normal University (HR 310-2019).

Materials

The emotional sounds materials were the same as those used in Experiment 1. The visual target was a white triangle (either downward or upward) presented on a black background at a 12-degree visual angle. The mean dB level for all sounds was 65 dB. Auditory stimuli were presented randomly to either the left or right ear over Sennheiser HD201 headphones.

Procedures

Participants were seated approximately 60 cm away from the monitor and introduced to the procedure. Each trial began with a fixation cross for 750 ms, followed by an auditory cue (positive, negative, or neutral) presentation that lasted for 2000 ms and appeared randomly in the left or right ear. Next, a visual target, a white triangle pointing upward or downward, appeared for 100 ms on the left or right side. During valid trials, the triangle appeared on the same side that was previously occupied by the emotional sound. During invalid trials, the triangle appeared on the opposite side of the auditory cue. Valid and invalid trials were presented in randomized order in equal proportions (50%). Participants were instructed to indicate the location (left or right) of the visual target (a white triangle) by pressing a key on the response keyboard ("F" or "J") when the triangle was in a certain orientation (pointing upward or downward, counterbalanced across participants). Participants had 2000 ms to respond. Participants were also asked to respond as quickly and accurately as possible (see [Figure 2A](#)).

We used a 3 (auditory cue valence: positive vs. negative vs. neutral) \times 2 (cue validity: valid cue vs. invalid cue) within-subjects design. The experiment contained two blocks. Each block included 180 trials: 30 positive-valid trials, 30 positive-invalid trials, 30 negative-valid trials, 30-negative-invalid trials, 30 neutral-valid trials, and 30 neutral-invalid trials. In total, the experiment consisted of 360 trials. After a practice phrase consisting of 20 trials, participants completed 360 experimental trials, of which 160 were valid trials and 160 invalid were trials. There was a 3 min break at the mid-point of the experimental trials.

Results

The mean error rate was 0.99%. These results were not analyzed further. Incorrect trials and RTs less than 150 ms and greater than 1500 ms were eliminated from the RT analysis. Response times were analyzed in a 3 \times 2 repeated measures ANOVA with the factors: cue valence (positive cue, neutral cue vs. negative cue) and cue validity (valid trials vs. invalid trials). Analysis revealed a significant main effect of cue valence, $F(2, 116) = 4.92$, $p < 0.01$, $\eta_p^2 = 0.08$, with faster reaction times

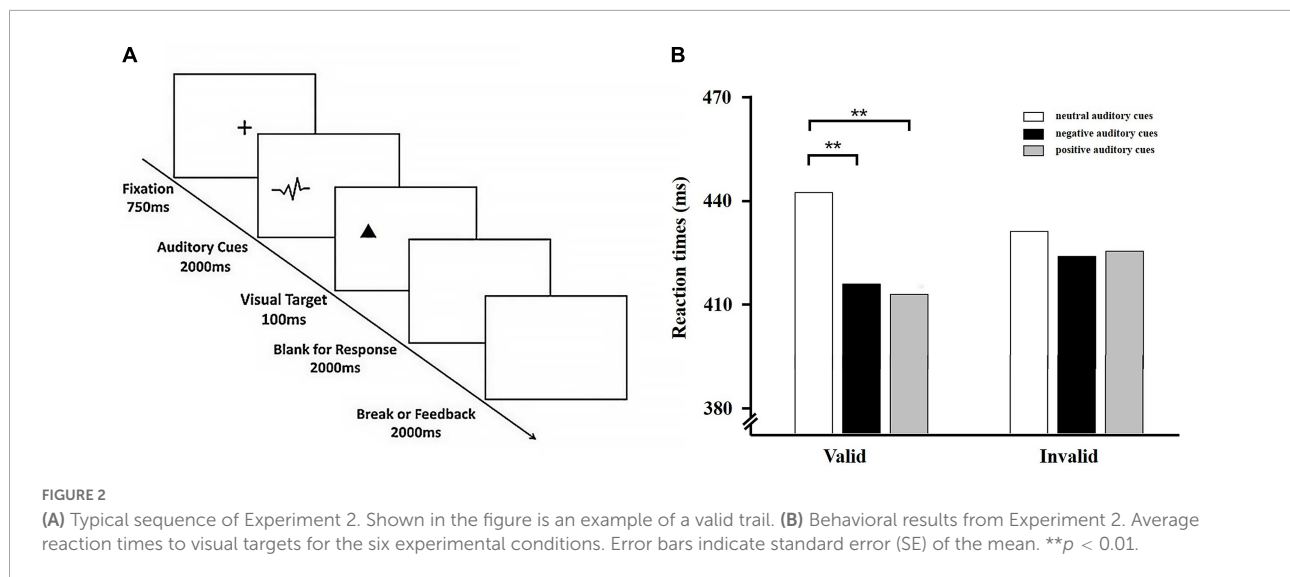
for positive auditory cues ($M = 415.13$ ms, $SD = 77.68$ ms) and negative auditory cues ($M = 415.77$ ms, $SD = 67.97$ ms) than for neutral auditory cues ($M = 430.34$ ms, $SD = 77.29$ ms). The main effect of cue validity was not significant, $F(1, 58) = 1.37$, $p > 0.05$. Importantly, there was also a significant Cue Valence \times Cue Validity interaction, $F(2, 116) = 8.10$, $p < 0.01$, $\eta_p^2 = 0.12$. Further simple effect analysis of the interaction effect revealed that RTs were faster for positive auditory cues ($M = 409.78$ ms, $SD = 78.29$ ms) and negative auditory cues ($M = 412.32$ ms, $SD = 64.95$ ms) than they were for neutral auditory cues ($M = 435.17$ ms, $SD = 79.86$ ms) in the valid condition, $F(2, 116) = 8.10$, $p < 0.01$, $\eta_p^2 = 0.15$. However, neither RTs for positive auditory cues nor RTs ($M = 420.48$ ms, $SD = 77.06$ ms) for negative auditory cues ($M = 419.21$ ms, $SD = 70.99$ ms) differed from RTs for neutral auditory cues ($M = 425.51$ ms, $SD = 74.74$ ms) in the invalid condition, $F(2, 116) = 1.18$, $p > 0.05$ (see [Figure 2B](#)).

For neutral auditory cues, response times were slower in the valid condition ($M = 435.17$ ms, $SD = 79.86$ ms) relative to the invalid condition ($M = 425.51$ ms, $SD = 74.74$ ms), $t(58) = 2.13$, $p < 0.05$, which suggested that the IOR effect was observed; however, for positive auditory cues, RTs were faster in the valid condition ($M = 409.78$ ms, $SD = 78.29$) relative to in the invalid condition ($M = 420.78$ ms, $SD = 77.06$ ms), $t(58) = -2.92$, $p < 0.01$, and also for negative cues, RTs were also faster in the valid condition ($M = 412.32$ ms, $SD = 64.95$ ms) than in the invalid condition ($M = 419.21$ ms, $SD = 70.99$ ms), $t(58) = -2.17$, $p < 0.05$.

Similar to Experiment 1, we also calculated the IOR effect by subtracting the mean reaction time on invalid trials from the mean reaction time on valid trials for different auditory cues. The one-way repeated measures for the magnitude of IOR effect revealed a significant main effect of cue valence, $F(2, 116) = 8.10$, $p < 0.01$, $\eta_p^2 = 0.12$. Follow-up analyses of the main effect indicated that both of the IOR effects for positive auditory cues ($M = -10.70$ ms, $SD = 28.17$ ms) and for negative auditory cues ($M = -6.89$ ms, $SD = 24.43$ ms) were significantly smaller compared to neutral auditory cues ($M = 9.66$ ms, $SD = 34.86$ ms). There was no significant difference in IOR effect between positive auditory cues and negative auditory cues. Taken together, the IOR effect was reduced when cues were emotional auditory cues relative to neutral auditory cues, which suggested that participants displayed cross-modal emotional attentional bias toward auditory information from the natural environment.

Discussion

Experiment 2 further demonstrated that emotional auditory cues (both positive and negative natural sounds) can guide visual spatial allocation of attention. It also provided preliminary evidence that positive natural sounds modulate spatial attention



to visual targets. The findings of Experiment 2 aligned with a recent study that examined spatial attention in 3-D space. It suggested that IOR size in the rewarded conditions was smaller than IOR size in the unrewarded condition and demonstrated that reward can attract much more attention in the near depth (Wang et al., 2022). Similarly, another study also demonstrated that social reward has the power to drive attention orienting behaviors (Hayward et al., 2018). Experiment 2 extended the findings of positive attentional bias that was found in auditory attention to the cross-modal domain. However, behavioral measures (reaction times) in Experiment 2 only provided an indirect measure of attentional processing. Event-related potentials (ERPs) are well-suited for investigating emotional attention bias because these measures allow for the examination of the time course of attention to emotional information with millisecond resolution (Kappenman et al., 2014, 2015). Researches on attentional bias toward visual emotional information have revealed that the modulations of early ERP components, including the P1, N170, and N2pc, suggesting that enhanced capture of attention by emotional stimuli at earlier stages of processing, and the modulations of later ERP components, such as the P3 or LPP components, indexing attention maintenance of emotion information at later stages of processing (Torrence and Troup, 2018; Gupta et al., 2019; Carlson, 2021). It remains unclear whether modulations of early ERP components or late ERP components were associated with attentional bias toward positive natural sounds in healthy adults. If individuals are sensitive to pursue potential survival resources or to obtain reward, healthy individuals tend to initially direct attentional resources toward positive natural sounds during early, automatic stages of processing adults, and might exhibit the modulations of earlier ERP components (e.g., N1, N2pc component). Or healthy adults tend to maintain their attention on positive stimuli to keep a positive and optimistic

mood, accomplishing the goal of emotion regulation and mental health, indexing the modulations of later ERP components (e.g., P3 component). In the auditory domain, the cue-elicited N1 component was reported to be modulated by the emotional state at an early stage (Herrmann and Knight, 2001; Folyi et al., 2012). If the attention of emotional sounds is enhanced at the early stage, one should expect an increased amplitude of the N1 component evoked by emotional sounds relative to that evoked by neutral sounds (Folyi et al., 2016). P3 amplitude is thought to measure attentional allocation or attentional maintenance, and thereby the relative salience of a stimulus to a particular individual (Gasbarri et al., 2007; Bistricky et al., 2014). If greater attentional resources were allocated toward emotional sounds rather than neutral sounds, we also expect greater P3 component evoked by emotional sounds. Thus, we conducted Experiment 3 to further explore the time course and neural substrates related to the processing of positive natural sounds using the event-related potential (ERP) technique in the same cross-modal emotional spatial cueing task used in Experiment 2. We expected that positive natural sounds would evoke an early attentional enhancement (e.g., the N1 component). Because the first two experiments demonstrated that healthy participants exhibited attention bias toward emotional sounds (both positive and negative), the purpose of Experiment 3 was undertaken mainly to investigate the neural mechanism of attentional bias toward positive natural sounds, so we only adopted positive and neutral natural sounds as auditory cues. We expected positive natural sounds could capture attention automatically, resulting in a greater N1 amplitude during earlier stages of attention processing in healthy individuals. In addition, in emotional spatial cueing task, positive auditory cues could facilitate attention to target detection in valid condition. Therefore, we also hypothesized that positive natural sounds would evoke greater P300 amplitude for validly cued targets relative to

neutral natural sounds, which reflected conscious and evaluative processing or attention maintenance of positive natural sounds.

Experiment 3

Participants

The current sample size was chosen based on previous studies focused on the ERP correlates of attentional bias (e.g., Gupta et al., 2021; Woltering et al., 2021). A power analysis (G*Power, Version 3.1) (Faul et al., 2007) estimated that 36 participants would be needed to achieve a power of 0.95 ($f = 0.25$, $\alpha = 0.05$, $\beta = 0.95$). Thirty-nine undergraduate students were recruited for monetary compensation. Three participants were excluded because of uncorrectable eye movement artifacts, and 36 subjects remained ($N_{\text{female}} = 20$, $M_{\text{age}} = 21.10$ years, $SD_{\text{age}} = 2.16$). All participants were right-handed and had normal or corrected-to-normal vision. None of the participants exhibited a history of psychiatric or neurological disorders. This study was approved by the Ethics Committee of East China Normal University (HR 310-2019), and all participants gave written informed consent.

Experimental materials and procedure

The materials and procedure in Experiment 3 were similar to those used in Experiment 2 with the following exceptions. Negative sounds were not included in Experiment 3. Ten positive sounds and ten neutral sounds were adopted in the experimental materials, resulting in twenty sounds in total. All sounds were cropped to a duration of 2000 ms.

A within-subject design of 2 (auditory cue valence: positive auditory cues vs. neutral auditory cues) \times 2 (cue validity: valid cues invalid cues) was used with 200 trials in each block (50 positive-valid trials; 50 positive-invalid trials; 50 neutral-valid trials; 50 neutral-invalid trials). The experiment contained two blocks. In total, the experiment consisted of 400 trials.

EEG recording

Experiments were conducted in a dimly illuminated, anechoic, and sound-proof room. EEG was recorded continuously using Brain Vision Recorder, BrainAmp DC and a 64-channel actiCAP (Brain Products GmbH, Gilching, Germany) with a sampling rate of 1000 Hz and a FCz reference. The left and right mastoid electrodes were used as offline-reference. Electrodes placed at the outer right canthi measured the horizontal electrooculogram (HEOG), and electrodes below the left eye measured the vertical electrooculogram

(VEOG). All electrode impedances were kept below 5 k Ω during data acquisition.

Event-related potentials-data analysis

Signal processing and analysis were performed in MATLAB using EEGLAB toolbox version 14.1 (Delorme and Makeig, 2004). First, EEG data were registered to the standard BESA head mold, and the bilateral mastoid was reset as the reference electrode (TP9 and TP10). Data were filtered using a low-pass filter of 0.1 Hz and a high-pass filter of 30 Hz. Independent component analysis (ICA, Delorme et al., 2007) was performed on each participant's data, and components that were clearly associated with eyeblinks or horizontal eye movements—as assessed by visual inspection of the waveforms and the scalp distributions of the components—were removed. Data exceeding ± 80 μ V were rejected, and remaining artifacts were manually removed, in total, 4.9% of the trials were excluded from further analyses. The remaining auditory cue-elicited epochs was segmented into epochs from 200 ms before and 2000 ms after the onset of auditory cue. Epochs were baseline corrected using the 200-ms pre-stimulus interval and averaged separately for the different valence conditions.

In the present study, electrode sites were chosen based on previous studies that focused on auditory spatial attention (Burra et al., 2019; Zimmer et al., 2019). We extracted the mean amplitude of the N1 component elicited by auditory cues, which was one of typical early auditory ERP components (Coch et al., 2005; Gädeke et al., 2013; Eddins et al., 2018; Topalidis et al., 2020; Rosburg and Mager, 2021) (in a cluster of six electrode sites (C1/C2, C3/C4, and T7/T8), covering the time from 130 to 200 ms after auditory cue onset). We also focused on the P300 component to visual targets in a cluster of five electrodes (CPz, CP1, CP2, CP3, and CP4) at a time window of 240–320 ms after the onset of visual targets.

Results

Behavioral performance

The average accuracy for visual targets was 98.9%, $SD = 0.92\%$. For the accuracy rates, a 2 (auditory cue valence: positive vs. neutral) \times 2 (cue validity: valid cues vs. invalid cues) repeated measures ANOVA was conducted. No significant main effect or interaction was found ($F_s < 1$, $p > 0.05$); therefore, these results were not analyzed further. Incorrect trials and RTs less than 150ms and greater than 1500 ms were eliminated from further analysis.

A 2 (auditory cue valence: positive vs. neutral) \times 2 (cue validity: valid cues vs. invalid cues) repeated measures ANOVA was performed for response times. The main effect of cue validity was not significant, $F(1, 35) = 3.93$, $p > 0.05$.

The main effect of cue valence was also not significant, $F(1, 35) = 0.96$, $p > 0.05 = 0.33$. The Cue Valence \times Cue Validity interaction was significant, $F(1, 35) = 10.89$, $p < 0.01$, $\eta_p^2 = 0.24$, suggesting faster responses for positive auditory cues ($M = 345.92$ ms, $SD = 88.76$ ms) relative to neutral auditory cues ($M = 357.96$ ms, $SD = 93.34$ ms) for valid trials, $t(35) = -3.88$, $p < 0.001$. No difference in RTs was observed between positive auditory cues ($M = 353.89$ ms, $SD = 92.62$ ms) and neutral cues ($M = 353.38$ ms, $SD = 90.68$ ms) for invalid trials, $t(35) = -0.13$, $p > 0.05 = 0.89$ (see **Figure 3**). We also calculated the IOR effect for different auditory cues. A t -test was conducted to compare the IOR effect difference between positive auditory cues and negative auditory cues. The analysis revealed that the IOR effect for positive auditory cues ($M = -2.94$ ms, $SD = 18.58$ ms) was significantly smaller than it was for neutral auditory cues ($M = 4.47$ ms, $SD = 15.48$ ms), $t(35) = -2.15$, $p < 0.05$. These findings were replicated with the results of Experiment 2, which revealed that healthy participants exhibited enhanced visual spatial attention to positive auditory cues.

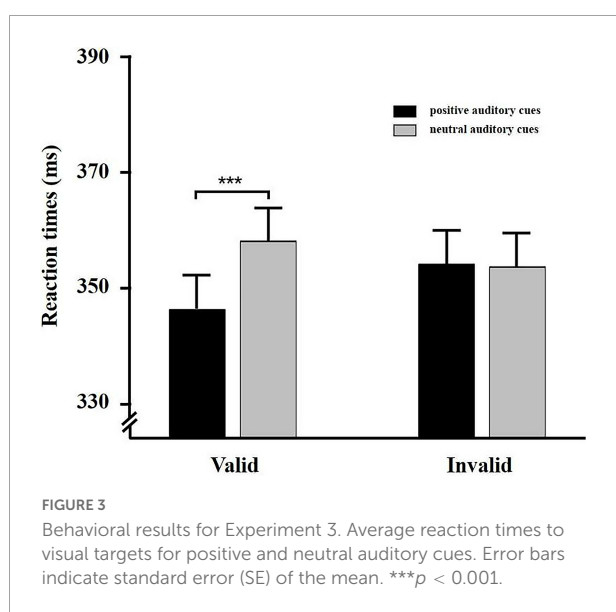
EEG results

Based on previous studies (Keil et al., 2014; Zimmer et al., 2015; Cai et al., 2020), we focused on the N2ac brain wave components (C3/4, CP5/CP6, FC5/FC6, and T7/T8 electrodes) and N1 (C1/C2, C3/C4, and T7/T8 electrodes) components evoked by emotional sound cues, and the calculation method of N2ac was contralateral wave minus ipsilateral wave deaveraging at left and right brain electrodes. A time window of 200–300 ms after the onset of auditory stimulation was selected for the N2ac component (Burra et al., 2019), and a time window of 130–200 ms was selected for the N1 component (Zimmer et al., 2015; Burra et al., 2019). However, we did not observe the N2ac component in the present experiment. This is most likely due

to our paradigm. An N2ac component is observed when cue and target are in close proximity. It is absent when stimuli are presented one at a time, as they were in the present study (Gamble and Luck, 2011; Luck, 2012, for a review; Zimmer et al., 2015). We also focused on the P300 component to visual targets in a cluster of five electrodes (CPz, CP1, CP2, CP3, and CP4) at a time window of 240–320 ms after the onset of visual targets.

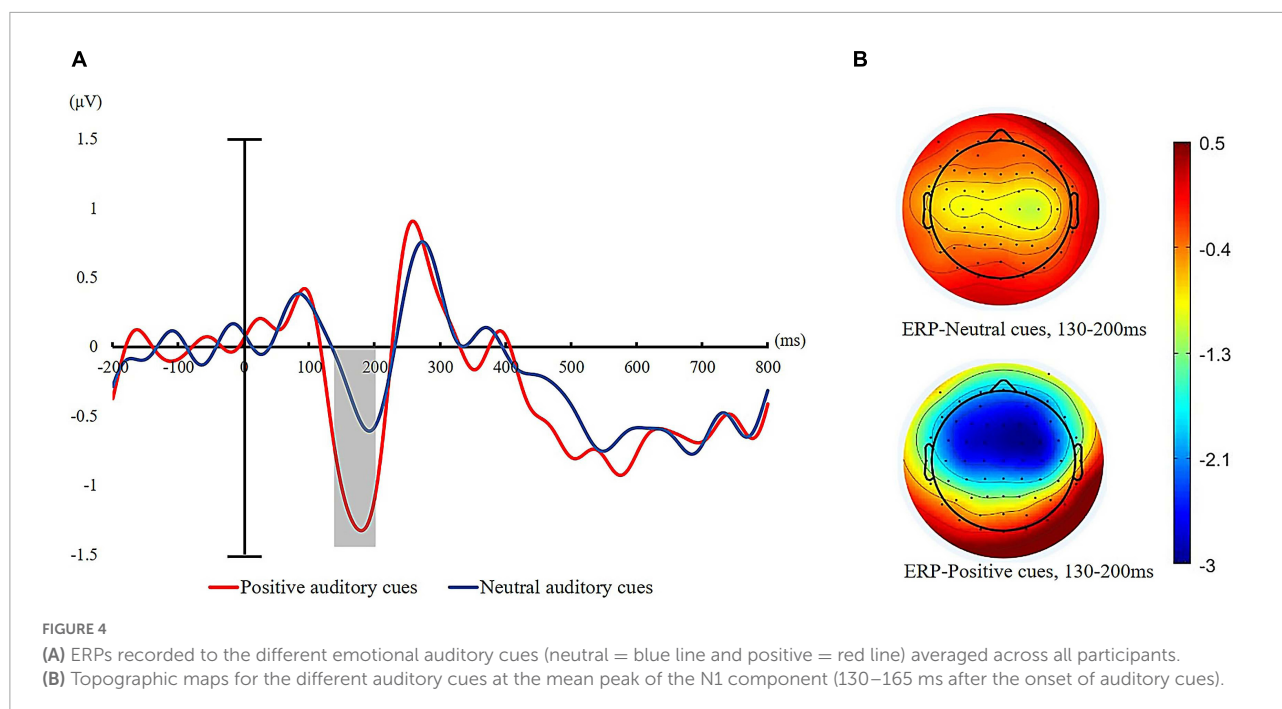
N1 amplitude in response to auditory cues

We calculated the averaged voltages for each type of auditory cue corresponding to a time window of 130–200 ms on N1 amplitude. For the N1 component elicited by auditory cues, a 2 (cue valence: positive auditory cue, neutral auditory cue) \times 2 (hemispheric laterality: contralateral vs. ipsilateral) repeated-measure ANOVA was conducted. Analysis of the mean N1 amplitudes following the onset of auditory cues revealed a significant main effect of cue valence, $F(1, 35) = 78.15$, $p < 0.001$, $\eta_p^2 = 0.69$. Follow-up analysis of the main effect revealed that the average N1 amplitude elicited by auditory cues was significantly more negative for positive sound ($M = -1.12$ μ V, $SD = 0.65$ μ V) relative to neutral sound ($M = -0.33$ μ V, $SD = 0.52$ μ V). The grand average N1 waves for the two auditory cues are displayed in **Figure 4**. There was also a significant main effect of hemispheric laterality, $F(1, 35) = 13.76$, $p < 0.01$, $\eta_p^2 = 0.26$. The N1 amplitude was significantly larger over the hemisphere contralateral to the sound location than the ipsilateral hemisphere (contralateral: $M = -0.83$ μ V, $SD = 0.72$ μ V; ipsilateral: $M = -0.44$ μ V, $SD = 0.68$ μ V). The interaction for cue valence \times cue validity was not significant, $F(35) = 0.35$, $p > 0.05$.



P300 amplitude in response to visual targets

We also calculated the mean amplitudes of P300 after the onset of visual targets at centroparietal sites (CPz, CP1, CP2, CP3, and CP4). Repeated measures ANOVAs of 2 (cue valence: positive vs. neutral) \times 2 (cue validity: valid cues vs. invalid cues) on mean amplitude of P300 revealed a significant main effect of cue valence, $F(1, 35) = 12.99$, $p < 0.01$, $\eta_p^2 = 0.27$, such that visual targets elicited greater P300 amplitudes for positive auditory cues ($M = 2.50$ μ V, $SD = 1.89$ μ V) than neutral auditory cues ($M = 1.95$ μ V, $SD = 1.51$ μ V). The main effect of cue validity was also significant, with the P300 amplitude significantly higher for invalid trials ($M = 2.58$ μ V, $SD = 2.71$) than for valid trials ($M = 1.88$ μ V, $SD = 1.87$ μ V), $F(1, 35) = 11.22$, $p < 0.001$, $\eta_p^2 = 0.24$. The interaction of cue validity and cue valence was also significant, $F(1, 35) = 4.44$, $p < 0.05$, $\eta_p^2 = 0.11$. Further simple effect analysis showed that



valid targets evoked a larger P300 for positive auditory cues compared to neutral auditory cues (positive: $M = 2.29 \mu V$, $SD = 1.98 \mu V$ vs. neutral: $M = 1.47 \mu V$, $SD = 1.75 \mu V$), $t(35) = 5.06$, $p < 0.001$. However, the mean amplitude of P300 evoked by invalid targets was not statistically distinguishable between positive auditory cues ($M = 2.71 \mu V$, $SD = 1.81 \mu V$) and neutral auditory cues ($M = 2.43 \mu V$, $SD = 1.61 \mu V$) (see **Figure 5**).

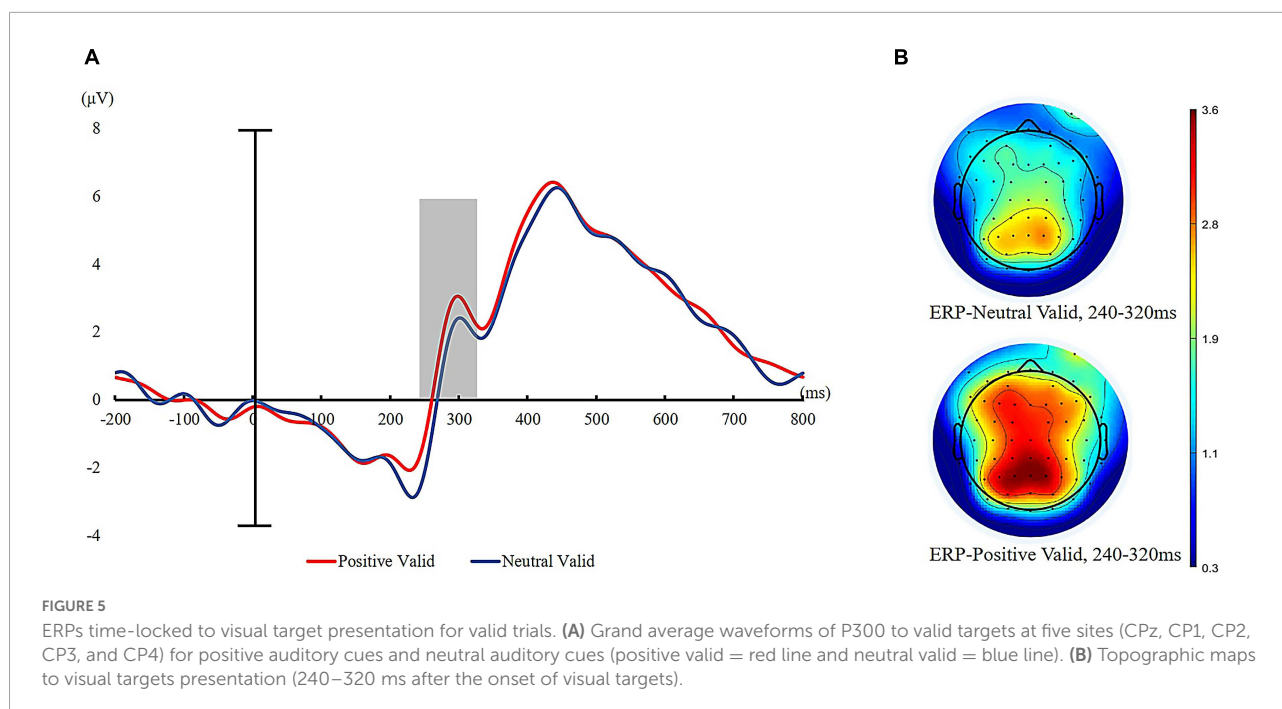
Discussion

Experiment 3 replicated and extended the results of Experiment 2. In Experiment 3, behavioral results again found that healthy individuals exhibited attentional bias toward positive auditory emotional stimuli, as indicated by faster responses to positive auditory cues than neutral auditory cues in the valid condition. Furthermore, the ERP results of Experiment 3 found an increased amplitude of N1 triggered by positive auditory cues relative to that triggered by neutral auditory cues, suggesting that enhanced attention with positive natural sounds compared with neutral natural sounds at a perceptual stage of auditory processing. In addition, we found that the amplitude of P300 evoked by visual targets was significantly larger for positive auditory cues than neutral auditory cues on valid trials, illustrating that those positive emotional sounds increased attention allocation to subsequent same-location visual target stimuli. Therefore, the converging behavioral and electrophysiological evidence of Experiment 3 suggested that healthy individuals exhibited cross-modal attentional bias

toward positive natural sounds, and this positive auditory attentional bias was related to facilitated attentional engagement with positive natural sounds.

General discussion

Across three experiments that combined behavioral and event-related potential techniques, the present study investigated auditory attentional bias (Experiment 1, auditory spatial cueing task) and cross-modal attentional bias (Experiment 2 and Experiment 3, cross-modal spatial cueing task) toward positive sounds from natural environment and further revealed the neural time course of attentional bias toward positive auditory information (Experiment 3). The behavioral results of these three experiments consistently demonstrated that positive auditory cues can modulate auditory/visual spatial allocation of attention. The electrophysiological findings of Experiment 3 suggest that attentional bias toward positive natural sounds occurs rapidly and during early stages of attention processing, which reflected the amplitude of N1 component. This result aligns with the view that processing of emotional stimuli can be facilitated by rapid attention (Yiend, 2010; Folyi et al., 2016; Pool et al., 2016). In addition, the results of Experiment 3 also showed that the amplitudes of P300 component was enlarged for valid targets when cued by positive sounds than by neutral sounds, which might indicate positive natural sounds can also facilitate conscious attention maintenance of positive information at later stages of attention processing.



Attentional bias for emotional stimuli has attracted considerable interest in neuroscience (Vuilleumier, 2005; Yuan et al., 2019) and psychology (Yiend, 2010; Van Bockstaele et al., 2014). Initially, experimental research in both fields mainly focused on negative attentional bias in healthy individuals (e.g., Preciado et al., 2017) or anxiety individuals (e.g., Sheppes et al., 2013). The preferential allocation of selective attention to threatening stimuli can promote the efficient detection of potential dangers in the environment (Öhman and Mineka, 2001; Pourtois et al., 2013). However, attentional bias toward positive stimuli also plays an important role in human reproduction, maintaining mental health, and coping with stress (Dandeneau et al., 2007; Moors, 2014; Anderson, 2016; Thoern et al., 2016). Until recently, most studies that examine the impact of positive emotional stimuli on spatial attention have focused on the visual domain (Brosch et al., 2008b; Gable and Harmon-Jones, 2010). In addition to having spatial or physical characteristics, sounds can also transport object-specific information that may help individuals identify or detect visual objects. The acoustic channel is particularly well suited for signifying warning signals (Haas and van Erp, 2014). Therefore, we conducted the present study with the aim of exploring how positive natural sounds modulate auditory and visual spatial attention.

The behavioral results of these three experiments consistently revealed that response times were faster for positive auditory cues than for neutral auditory cues in the valid condition, and the magnitude of IOR was reduced in positive trials relative to neutral trials. Our findings confirmed that positive natural sounds were preferentially processed

over neutral auditory information and guide auditory or visual spatial attention. These results aligned with previous findings on negative auditory attentional bias (Asutay and Västfjäll, 2015; Wang et al., 2019a,b; Bonmassar et al., 2020) or negative visual attention bias (Harrison and Davies, 2013; Gerdes et al., 2014, 2020; Burra et al., 2019). In the first two experiments, we also found that participants exhibited a similar response pattern to positive and negative sounds; in other words, participants showed attentional bias toward emotional sounds. Therefore, attentional bias for these two kinds of emotional stimuli (positive rewarding and negative threatening stimuli) should be similar (Pool et al., 2016). The present study provided initial evidence suggesting that humans exhibited an attentional bias toward positive natural sounds. Similar to positive visual information, positive natural sounds (e.g., chirpings of birds; sounds of gurgling water) can also carry significant emotional information (such as potential survival resources or opportunities, indicating novelty and diversity in nature), resulting in faster responses to auditory/visual targets.

Behavioral measures of attentional bias toward positive sounds failed to clarify the time course of the positive attentional bias. Therefore, Experiment 3 explored the ERP component of positive attentional bias and found that positive natural sounds elicited larger N1 amplitudes (130–200 ms after emotional auditory cues onset) relative to neutral sounds, suggesting an involuntary orienting of attention toward positive auditory cues. Recent reviews on the neural correlates of threat-related attentional bias have shown that both healthy and anxious populations exhibited facilitated engagement with negative stimuli, as shown by the greater amplitude of early ERP

components, including the N1 component (Gupta et al., 2019). Previous studies that focused on auditory attention have shown that the increased amplitude of N1 component was generally considered to be an increase in attentional resources to acoustic cue stimuli (Hillyard et al., 1973; Lange et al., 2003; Doricchi et al., 2020). Based on previous findings on the N1 component, the present study indicated that the modulation of positive sounds on spatial attention emerged at quite early stages of attention.

Our results aligned with the findings of previous studies on the neural mechanism of attentional bias toward emotional visual information (Santesso et al., 2008; Pintzinger et al., 2017). It also has been shown that the N1 amplitude is larger for both pleasant and unpleasant images compared to neutral images (Olofsson et al., 2008; Foti et al., 2009; Grassini et al., 2019) and for cannabis-related cues in cannabis use disorders (Ruglass et al., 2019). Moreover, some studies that focused on auditory attention have shown that the increased amplitude of N1 component was generally considered to signify enhanced attention engagement with auditory cues (Hillyard et al., 1973; Woldorff et al., 1987; Lange et al., 2003; Doricchi et al., 2020). Evidence showed that the amplitude of the N1 component was enlarged for infant high-distress cries in mothers (Maupin et al., 2019), for task-relevant tones (Alho et al., 1994; Lange, 2013), and in the difficult auditory detection task (Sabri et al., 2006). Thus, the enlarged N1 amplitude in Experiment 3 might reflect enhancement of auditory attention captured by positive auditory stimuli in normal populations, which aligns with a recent study focused on attentional bias toward positive visual information (Pintzinger et al., 2017). Further, these results suggested that healthy adults preferred to process rewarding stimuli involuntarily within the auditory field. The findings of the present study are also in agreement with an extensive body of literature on multi-modal spatial cueing effects, which demonstrates that cues in one sensory channel can guide spatial attention allocation in another modality (Zimmer et al., 2015; Pierce et al., 2018; Gerdes et al., 2020).

In Experiment 3, visual targets elicited larger P300 amplitudes for positive auditory cues compared to neutral auditory cues on valid trials. Previous studies have shown that the amplitude of P300 was related to attentional resource allocation to emotional stimuli (Yee and Miller, 1994; Kessels et al., 2010; Willner et al., 2015), attentional disengagement with threat (Baik et al., 2018), the maintenance of attention (Bauer, 2021). Experiment 3 found that P300 amplitudes were higher for visual targets after positive auditory cues than those for visual targets after neutral auditory cues in valid trials, but not in invalid trials. Therefore, it is reasonable to assume that positive auditory cues facilitated attention resource allocation to visual targets preceded by positive auditory cues. Taken together, the findings of Experiment 3 revealed electrophysiological evidence that attention allocation to positive auditory information occurred at earlier stages in healthy adults, which was consistent with previous findings

(Nelson and Hajcak, 2017; Ferry and Nelson, 2021). More importantly, in a recent meta-analytic study, researchers systematically compared attentional bias toward positive visual stimuli across 243 studies, and they suggested that attentional bias for positive visual information occurred rapidly and involuntarily during early stages of attention processing (Pool et al., 2016). To our knowledge, this is the first study to provide converging evidence that positive auditory information enhances both initial and sustained attention in healthy individuals. We also expanded the previous studies conducted on visual stimuli to auditory stimuli of natural word. This may help to transfer our findings of cross-modal influences on visual attention to more natural settings.

There are a few limitations. First, effect sizes for observed experimental effects are relatively small, one explanation is that the number of trials is low. Future studies should use a larger number of trials to examine this question. Another possible explanation is the effect size of attentional bias toward positive information was rather small. One recent meta-analysis showed that healthy individuals existed reward-related attention distraction with a small effect size (Rusz et al., 2020). Another meta-analytic investigation has reported that the effect size of the positive attentional bias using spatial cueing task was smaller compared with rapid serial visual presentation paradigm (Pool et al., 2016). Second, we did not include negative sounds in Experiment 3. It would be helpful to compare the ERP correlates of negative with those of positive sounds. Nonetheless, we emphasize that the main goal of the current study is on the differential attentional bias toward positive natural sound vs. the neutral sound distinction and how these effects are influenced in auditory and cross-modal setting.

Future studies can be refined in the following aspects. First, according to circumplex theories of emotion, all emotions are underlain by two orthogonal dimensions: valence and arousal (Russell, 1980; Anderson, 2005). Based on this assumption, positive emotional stimuli influence attentional selection contingent on their potential to elicit emotional arousal. Thus, a potential direction for future research involves exploring whether the size of positive attentional bias could be predicted by different arousing positive natural sounds. Second, only healthy adults were included in the present study. It may be extremely useful to examine the cognitive and neural mechanism of attention bias toward positive auditory emotional stimuli in clinical samples (e.g., depression, anxiety and dysphoric individuals), because doing so could help us understand the mechanisms that underline the etiology and maintenance of emotion disorders (Elgersma et al., 2018; Sadek et al., 2020). Third, emotion appraisal theory proposes that the psychological mechanism that drives emotional attention is the detection of a stimulus that is relevant to the observer's concerns and goals (Sander et al., 2005). Therefore, future studies need to manipulate the stimulus to make it relevant for the participants' concern in order to examine attentional bias toward positive auditory stimuli. For example, the sound

of laughter from intimate friends has a much higher level of relevance to the participant than the laughter of strangers, resulting a larger attentional bias. Fourth, the present study only investigated the modulational effect of positive environmental sounds on subsequent neutral visual or auditory target stimuli. In order to comprehensively investigate cross-modal attention bias toward positive emotional stimuli, further studies are needed to specifically investigate whether positive visual cues (or simultaneously presented auditory cues) might modulate auditory attention. Finally, future investigations are also needed to develop a more nuanced understanding of how these effects generalize across discrete emotions (e.g., disgust, sadness, joy, fear, etc.) rather than just comparison general positive/negative understanding as of now.

Conclusion

In summary, to our best knowledge, the present study is the first to investigate the mechanism of attentional bias toward positive natural sounds in healthy individuals. In three experiments, both behavioral and electrophysiological data converge on the central finding that normal individuals exhibited attention enhancement with positive auditory emotional stimuli, and the results of the N1 component in Experiment 3 provide direct electrophysiological evidence that attention bias toward positive auditory stimuli occurs very rapidly, involuntarily, and during the early stages of the attentional process. In general, our results aligned with evidence that there is an attentional bias toward positive visual information and further demonstrated that positive auditory information is processed preferentially.

Data availability statement

The original data presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

References

- Alho, K., Teder, W., Lavikainen, J., and Näätänen, R. (1994). Strongly focused attention and auditory event-related potentials. *Biol. Psychol.* 38, 73–90. doi: 10.1016/0301-0511(94)90050-7
- Anderson, A. K. (2005). Affective influences on the attentional dynamics supporting awareness. *J. Exp. Psychol. General* 134, 258–281. doi: 10.1037/0096-3445.134.2.258
- Anderson, B. A. (2016). Social reward shapes attentional bias. *Cogn. Neurosci.* 7, 30–36. doi: 10.1080/17588928.2015.1047823
- Anderson, B. A., and Yantis, S. (2012). Value-driven attentional and oculomotor capture during goal-directed, unconstrained viewing. *Attention Percept. Psychophys.* 74, 1644–1653. doi: 10.3758/s13414-012-0348-2
- Anderson, B. A., and Yantis, S. (2013). Persistence of value-driven attentional capture. *J. Exp. Psychol. Hum. Percept. Perform.* 39, 6–9. doi: 10.1037/a0030860
- Arieh, Y., and Marks, L. E. (2008). Cross-modal interaction between vision and hearing: a speed-accuracy analysis. *Percept. Psychophys.* 70, 412–421. doi: 10.3758/pp.70.3.412

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of East China Normal University (HR 310-2019). The patients/participants provided their written informed consent to participate in this study.

Author contributions

YW and ZT conceived of the project, designed the experiments, implemented the experiments, and collected the data. YW, ZT, XZ, and LY wrote the manuscript. All authors approved it for publication.

Funding

This work was supported by the Philosophy and Social Science Foundation of Shanghai (2019BSH010), the Humanity and Social Science Fund of Ministry of Education of China (20YJA190008), the Research Project of Shanghai Science and Technology Commission (20dz2260300), and The Fundamental Research Funds for the Central Universities.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Armony, J. L., and LeDoux, J. (2010). "Emotional responses to auditory stimuli," in *The Oxford Handbook of Auditory Science: The Auditory Brain*, Vol. 2, eds A. Rees and A. R. Palmer (New York, NY: Oxford University Press.), 479–505. doi: 10.1093/oxfordhb/9780199233281.013.0019
- Asutay, E., and Västfjäll, D. (2015). Negative emotion provides cues for orienting auditory spatial attention. *Front. Psychol.* 6:618. doi: 10.3389/fpsyg.2015.00618
- Baik, S. Y., Jeong, M., Kim, H. S., and Lee, S. H. (2018). ERP investigation of attentional disengagement from suicide-relevant information in patients with major depressive disorder. *J. Affect. Disord.* 225, 357–364. doi: 10.1016/j.jad.2017.08.046
- Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., and van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and non-anxious individuals: a meta-analytic study. *Psychol. Bull.* 133, 1–24. doi: 10.1037/0033-2909.133.1.1
- Basanovic, J., Kaiko, I., and MacLeod, C. (2021). Change in attentional control predicts change in attentional bias to negative information in response to elevated state anxiety. *Cogn. Therapy Res.* 45, 111–122. doi: 10.1007/s10608-020-10176-3
- Bauer, L. (2021). Temporal instability in brain activation: a novel paradigm for evaluating the maintenance of attention among substance dependent patients. *Psychopharmacology* 238, 2937–2946. doi: 10.1007/s00213-021-05909-5
- Bell, L., Wagels, L., Neuschaefer-Rube, C., Fels, J., Gur, R. E., and Konrad, K. (2019). The cross-modal effects of sensory deprivation on spatial and temporal processes in vision and audition: a systematic review on behavioral and neuroimaging research since 2000. *Neural Plast.* 2019:9603469. doi: 10.1155/2019/9603469
- Berggren, N., and Eimer, M. (2021). The role of trait anxiety in attention and memory-related biases to threat: an event-related potential study. *Psychophysiology* 58:e13742. doi: 10.1111/psyp.13742
- Bistricky, S. L., Atchley, R. N., Ingram, R., and O'Hare, A. (2014). Biased processing of sad faces: an ERP marker candidate for depression susceptibility. *Cogn. Emot.* 28, 470–492. doi: 10.1080/02699931.2013.837815
- Blicher, A., and Reinholdt-Dunne, M. (2019). Components of attentional bias to threat in clinically anxious children: an experimental study using the emotional spatial cueing paradigm. *Cogn. Therapy Res.* 43, 884–892. doi: 10.1007/s10608-019-10008-z
- Blurton, S. P., Greenlee, M. W., and Gondan, M. (2015). Cross-modal cueing in audiovisual spatial attention. *Attent. Percept. Psychophys.* 77, 2356–2376. doi: 10.3758/s13414-015-0920-7
- Bonmassar, C., Widmann, A., and Wetzel, N. (2020). The impact of novelty and emotion on attention-related neuronal and pupil responses in children. *Dev. Cogn. Neurosci.* 42:100766. doi: 10.1016/j.dcn.2020.100766
- Booth, R. W., and Sharma, D. (2020). Attentional control and estimation of the probability of positive and negative events. *Cogn. Emot.* 34, 553–567. doi: 10.1080/02699931.2019.1657382
- Bradley, M. M., and Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* 25, 49–59. doi: 10.1016/0005-7916(94)90063-9
- Brignell, C., Griffiths, T., Bradley, B. P., and Mogg, K. (2009). Attentional and approach biases for pictorial food cues. influence of external eating. *Appetite* 52, 299–306. doi: 10.1016/j.appet.2008.10.007
- Brosch, T., Sander, D., Pourtois, G., and Scherer, K. R. (2008a). Beyond fear: rapid spatial orienting toward positive emotional stimuli. *Psychol. Sci.* 19, 362–370. doi: 10.1111/j.1467-9280.2008.02094.x
- Brosch, T., Grandjean, D., Sander, D., and Scherer, K. R. (2008b). Behold the voice of wrath: cross-modal modulation of visual attention by anger prosody. *Cognition* 106, 1497–1503. doi: 10.1016/j.cognition.2007.05.011
- Burra, N., Kerzel, D., Munoz Tord, D., Grandjean, D., and Ceravolo, L. (2019). Early spatial attention deployment toward and away from aggressive voices. *Soc. Cogn. Affect. Neurosci.* 14, 73–80. doi: 10.1093/scan/nsy100
- Cai, W., Wang, L., Chen, T., Zhao, S., Feng, C., and Feng, W. (2020). Auditory attentional biases in young males with physical stature dissatisfaction. *Psychophysiology* 57:e13635. doi: 10.1111/psyp.13635
- Carlson, J. M. (2021). A systematic review of event-related potentials as outcome measures of attention bias modification. *Psychophysiology* 58:e13801. doi: 10.1111/psyp.13801
- Carlson, J. M., Conger, S., and Sterr, J. (2018). Auditory distress signals potentiate attentional bias to fearful faces: evidence for multimodal facilitation of spatial attention by emotion. *J. Nonverbal Behav.* 42, 417–426. doi: 10.1007/s10919-018-0282-7
- Carlson, J. M., and Reinke, K. S. (2014). Attending to the fear in your eyes: facilitated orienting and delayed disengagement. *Cogn. Emot.* 28, 1398–1406. doi: 10.1080/02699931.2014.885410
- Choiniere, J. N., Neenan, J. M., Schmitz, L., Ford, D. P., Chapelle, K. E. J., Balanoff, A. M., et al. (2021). Evolution of vision and hearing modalities in theropod dinosaurs. *Science* 372, 610–613. doi: 10.1126/science.abe7941
- Cisler, J. M., and Koster, E. H. W. (2010). Mechanisms of attentional biases towards threat in anxiety disorders: an integrative review. *Clin. Psychol. Rev.* 30, 203–216.
- Cisler, J. M., Olatunji, B. O., and Lohr, J. M. (2009). Disgust, fear, and the anxiety disorders: a critical review. *Clin. Psychol. Rev.* 29, 34–46. doi: 10.1016/j.cpr.2008.09.007
- Coch, D., Sanders, L. D., and Neville, H. J. (2005). An event-related potential study of selective auditory attention in children and adults. *J. Cogn. Neurosci.* 17, 605–622. doi: 10.1162/0898929053467631
- Concina, G., Renna, A., Grosso, A., and Sacchetti, B. (2019). The auditory cortex and the emotional valence of sounds. *Neurosci. Biobehav. Rev.* 98, 256–264. doi: 10.1016/j.neubiorev.2019.01.018
- Dandeneau, S. D., Baldwin, M. W., Baccus, J. R., Sakellaropoulos, M., and Pruessner, J. C. (2007). Cutting stress off at the pass: reducing vigilance and responsiveness to social threat by manipulating attention. *J. Pers. Soc. Psychol.* 93, 651–666. doi: 10.1037/0022-3514.93.4.651
- Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009
- Delorme, A., Westerfield, M., and Makeig, S. (2007). Medial prefrontal theta bursts precede rapid motor responses during visual selective attention. *J. Neurosci.* 27, 11949–11959. doi: 10.1523/JNEUROSCI.3477-07.2007
- Demeyer, I., and Raedt, R. D. (2013). Attentional bias for emotional information in older adults: the role of emotion and future time perspective. *PLoS One* 8:e65429. doi: 10.1371/journal.pone.0065429
- Dong, Y., De, B. A., Yu, L., and Zhou, R. (2017). Eye-movement evidence of the time-course of attentional bias for threatening pictures in test-anxious students. *Cogn. Emot.* 31, 781–790. doi: 10.1080/02699931.2016.1152953
- Doricchi, F., Pellegrino, M., Marson, F., Pinto, M., Caratelliet, L., Cestari, V., et al. (2020). Deconstructing reorienting of attention: cue predictiveness modulates the inhibition of the no-target side and the hemispheric distribution of the P1 response to invalid targets. *J. Cogn. Neurosci.* 32, 1046–1060. doi: 10.1162/jocn_a_01534
- Eddins, A. C., Ozmeral, E. J., and Eddins, D. A. (2018). How aging impacts the encoding of binaural cues and the perception of auditory space. *Hear. Res.* 369, 79–89. doi: 10.1016/j.heares.2018.05.001
- Elgersma, H. J., Koster, E. H. W., van Tuijl, L. A., Hoekzema, A., Hoekzema, A., Penninx, B. W. J. H., et al. (2018). Attentional bias for negative, positive, and threat words in current and remitted depression. *PLoS One* 13:e0205154. doi: 10.1371/journal.pone.0205154
- Ellsworth, P. C., and Scherer, K. R. (2003). "Appraisal processes in emotion," in *Handbook of Affective Sciences*, eds R. J. Davidson, H. H. Goldsmith, and K. R. Scherer (Oxford: Oxford University Press), 572–595.
- Evans, K. K. (2020). The role of selective attention in cross-modal interactions between auditory and visual features. *Cognition* 196:104119. doi: 10.1016/j.cognition.2019.104119
- Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (2007). GPower 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/bf03193146
- Ferry, R. A., and Nelson, B. D. (2021). Tactile P300 to unpredictable electric shocks: association with anxiety symptoms, intolerance of uncertainty, and neuroticism. *Biol. Psychol.* 162:108094. doi: 10.1016/j.biopsycho.2021.108094
- Folyi, T., Feher, B., and Horvath, J. (2012). Stimulus-focused attention speeds up auditory processing. *Int. J. Psychophysiol.* 84, 155–163. doi: 10.1016/j.ijpsycho.2012.02.001
- Folyi, T., Liesefeld, H. R., and Wentura, D. (2016). Attentional enhancement for positive and negative tones at an early stage of auditory processing. *Biol. Psychol.* 114, 23–32. doi: 10.1016/j.biopsycho.2015.12.001
- Foti, D., Hajcak, G., and Dien, J. (2009). Differentiating neural responses to emotional pictures: evidence from temporal-spatial PCA. *Psychophysiology* 46, 521–530. doi: 10.1111/j.1469-8986.2009.00796.x
- Fox, E., Russo, R., Bowles, R., and Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *J. Exp. Psychol. General* 130, 681–700. doi: 10.1037/0096-3445.130.4.681
- Fox, E., Russo, R., and Dutton, K. (2002). Attentional bias for threat: evidence for delayed disengagement from emotional faces. *Cogn. Emot.* 16, 355–379.

- Gable, P. A., and Harmon-Jones, E. (2010). Late positive potential to appetitive stimuli and local attentional bias. *Emotion* 10, 441–446. doi: 10.1037/a0018425
- Gädeke, J. C., Föcker, J., and Röder, B. (2013). Is the processing of affective prosody influenced by spatial attention? an ERP study. *BMC Neurosci.* 14:14. doi: 10.1186/1471-2202-14-14
- Gamble, M. L., and Luck, S. J. (2011). N2ac: an ERP component associated with the focusing of attention within an auditory scene. *Psychophysiology* 48, 1057–1068. doi: 10.1111/j.1469-8986.2010.01172.x
- Gao, C., Wedell, D. H., and Shinkareva, S. V. (2020). Crossmodal negativity bias in semantic processing. *Emotion* Online ahead of print. doi: 10.1037/emo0000918
- Gasbarri, A., Arnone, B., Pompili, A., Pacitti, F., Pacitti, C., and Cahill, L. (2007). Sex-related hemispheric lateralization of electrical potentials evoked by arousing negative stimuli. *Brain Res.* 1138, 178–186. doi: 10.1016/j.brainres.2006.12.073
- Gerdas, A. B. M., Alpers, G. W., Braun, H., Köhler, S., Nowak, U., and Treiber, L. (2020). Emotional sounds guide visual attention to emotional pictures: an eye-tracking study with audio-visual stimuli. *Emotion* 21, 679–692. doi: 10.1037/emo0000729
- Gerdas, A. B. M., Wieser, M. J., and Alpers, G. W. (2014). Emotional pictures and sounds: a review of multimodal interactions of emotion cues in multiple domains. *Front. Psychol.* 5:1351. doi: 10.3389/fpsyg.2014.01351
- Grafton, B., and MacLeod, C. (2017). A positive perspective on attentional bias: positive affectivity and attentional bias to positive information. *J. Happiness Stud.* 18, 1029–1043. doi: 10.1007/s10902-016-9761-x
- Grassini, S., Railo, H., Valli, K., Revonsuo, A., and Koivisto, M. (2019). Visual features and perceptual context modulate attention towards evolutionarily relevant threatening stimuli: electrophysiological evidence. *Emotion* 19, 348–364. doi: 10.1037/emo0000434
- Gupta, R. (2019). “Positive emotions have a unique capacity to capture attention,” in *Emotion and Cognition*, ed. N. Srinivasan (Cambridge, MA: Elsevier Academic Press). doi: 10.1016/b978-0-12-810902-016-9761-x
- Gupta, R. S., Kujawa, A., and Vago, D. R. (2019). The neural chronometry of threat-related attentional bias: event-related potential (ERP) evidence for early and late stages of selective attentional processing. *Int. J. Psychophysiol.* 146, 20–42. doi: 10.1016/j.ijpsycho.2019.08.006
- Gupta, R. S., Kujawa, A., and Vago, D. R. (2021). A preliminary investigation of ERP components of attentional bias in anxious adults using temporospatial principal component analysis. *J. Psychophysiol.* Online ahead of print. doi: 10.1027/0269-8803/a000275
- Gutiérrez-Cobo, M. J., Luque, D., Most, S. B., Fernández-Berrocal, P., and Le Pelley, M. E. (2019). Reward and emotion influence attentional bias in rapid serial visual presentation. *Quarterly J. Exp. Psychol.* 72, 2155–2167. doi: 10.1177/1747021819840615
- Haas, E. C., and van Erp, J. B. F. (2014). Multimodal warnings to enhance risk communication and safety. *Safety Sci.* 61, 29–35. doi: 10.1016/j.ssci.2013.07.011
- Hao, H. T., Schröger, E., and Kotz, S. A. (2015). Selective attention modulates early human evoked potentials during emotional face-voice processing. *J. Cogn. Neurosci.* 27, 798–818. doi: 10.1162/jocn_a_00734
- Harrison, N. R., and Davies, S. J. (2013). Modulation of spatial attention to visual targets by emotional environmental sounds. *Psychol. Neurosci.* 6, 247–251. doi: 10.1162/jocn.2009.21110
- Harrison, N. R., and Woodhouse, R. (2016). Modulation of auditory spatial attention by visual emotional cues: differential effects of attentional engagement and disengagement for pleasant and unpleasant cues. *Cogn. Process.* 17, 205–211. doi: 10.1007/s10339-016-0749-6
- Hayward, D. A., Pereira, E. J., Otto, R., and Ristic, J. (2018). Smile! social reward drives attention. *J. Exp. Psychol. Hum. Percept. Perform.* 44, 206–214. doi: 10.1037/xhp0000459
- Herrmann, C. S., and Knight, R. T. (2001). Mechanisms of human attention: event-related potentials and oscillations. *Neurosci. Biobehav. Rev.* 25, 465–476. doi: 10.1016/S0149-7634(01)00027-6
- Hillyard, S. A., Hink, R. F., Schwent, V. L., and Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science* 182, 177–180. doi: 10.1126/science.182.4108.177
- Hu, X., Guo, L., Han, J., and Liu, T. (2017). Decoding power-spectral profiles from fMRI brain activities during naturalistic auditory experience. *Brain Imaging Behav.* 11, 253–263. doi: 10.1007/s11682-016-9515-8
- Hutmacher, F. (2019). Why is there so much more research on vision than on any other sensory modality? *Front. Psychol.* 10:2246. doi: 10.3389/fpsyg.2019.02246
- Imhoff, R., Lange, J., and Germar, M. (2019). Identification and location tasks rely on different mental processes: a diffusion model account of validity effects in spatial cueing paradigms with emotional stimuli. *Cogn. Emot.* 33, 231–244. doi: 10.1080/02699931.2018.1443433
- Joormann, J., and Gotlib, I. H. (2007). Selective attention to emotional faces following recovery from depression. *J. Abnorm. Psychol.* 116, 80–85. doi: 10.1037/0021-843X.116.1.80
- Kaiser, D., Jacob, G. A., Domes, G., and Arntz, A. (2017). Attentional bias for emotional stimuli in borderline personality disorder: a meta-analysis. *Psychopathology* 49, 383–396.
- Kappenman, E. S., Farrens, J. L., Luck, S. J., and Proudfit, G. H. (2014). Behavioral and ERP measures of attentional bias to threat in the dot-probe task: poor reliability and lack of correlation with anxiety. *Front. Psychol.* 5:1368. doi: 10.3389/fpsyg.2014.01368
- Kappenman, E. S., MacNamara, A., and Proudfit, G. H. (2015). Electrooculographic evidence for rapid allocation of attention to threat in the dot-probe task. *Soc. Cogn. Affect. Neurosci.* 10, 577–583. doi: 10.1093/scan/nsu098
- Keil, A., Bradley, M. M., Junghöfer, M., Rüssmann, T., Lowenthal, W., and Lang, P. J. (2007). Cross-modal attention capture by affective stimuli: evidence from event-related potentials. *Cogn. Affect. Behav. Neurosci.* 7, 18–24. doi: 10.3758/cabn.7.1.18
- Keil, A., Debener, S., Gratton, G., Junghöfer, M., Kappenman, E. S., Luck, S. J., et al. (2014). Committee report: publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography. *Psychophysiology* 51, 1–21. doi: 10.1111/psyp.12147
- Kennedy, B. L., Huang, R., and Mather, M. (2020). Age differences in emotion-induced blindness: positivity effects in early attention. *Emotion* 20, 1266–1278. doi: 10.1037/emo0000643
- Kessels, L. T. E., Ruiter, R. A. C., and Jansma, B. M. (2010). Increased attention but more efficient disengagement: neuroscientific evidence for defensive processing of threatening health information. *Healthy Psychol.* 29, 346–354. doi: 10.1037/a0019372
- Kim, S. N., Kim, M., Lee, T. H., Lee, J., Park, S., Kim, D., et al. (2018). Increased attentional bias toward visual cues in internet gaming disorder and obsessive-compulsive disorder: an event-related potential study. *Front. Psychiatry* 9:315. doi: 10.3389/fpsyg.2018.00315
- Koumura, T., Terashima, H., and Furukawa, S. (2019). Cascaded tuning to amplitude modulation for natural sound recognition. *J. Neurosci.* 39, 5517–5533. doi: 10.1523/JNEUROSCI.2914-18.2019
- Kreutzmann, J. C., Jovanovic, T., and Fendt, M. (2020). Infralimbic cortex activity is required for the expression but not the acquisition of conditioned safety. *Psychopharmacology* 237, 2161–2172. doi: 10.1007/s00213-020-05527-7
- Lange, K. (2013). The ups and downs of temporal orienting: a review of auditory temporal orienting studies and a model associating the heterogeneous findings on the auditory N1 with opposite effects of attention and prediction. *Front. Hum. Neurosci.* 7:263. doi: 10.3389/fnhum.2013.00263
- Lange, K., Rösler, F., and Röder, B. (2003). Early processing stages are modulated when auditory stimuli are presented at an attended moment in time: an event-related potential study. *Psychophysiology* 40, 806–817. doi: 10.1111/1469-8986.00081
- LeDoux, J. E. (1996). *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. New York, NY: Simon & Schuster.
- Lepping, R. J., Bruce, J. M., Gustafson, K. M., Hu, J., Martin, L. E., Savage, C. R., et al. (2019). Preferential activation for emotional Western classical music versus emotional environmental sounds in motor, interoceptive, and language brain areas. *Brain Cogn.* 136:103593. doi: 10.1016/j.bandc.2019.103593
- Lockhofen, D. E. L., Hübner, N., Hemdan, F., Sammer, G., Henare, D., Schubö, A., et al. (2021). Differing time courses of reward-related attentional processing: an EEG source-space analysis. *Brain Topogr.* 34, 283–296. doi: 10.1007/s10548-021-00827-3
- Luck, S. J. (2012). “Electrophysiological correlates of the focusing of attention with complex visual scenes: N2pc and related ERP components,” in *The Oxford Handbook of ERP Components*, eds S. J. Luck and E. S. Kappenman (New York, NY: Oxford University Press).
- Mathews, A., and Mackintosh, B. (1998). A cognitive model of selective processing in anxiety. *Cogn. Therapy Res.* 22, 539–560. doi: 10.1023/A:1018738019346
- Maupin, A. N., Rutherford, H. J. V., Landi, N., Potenza, M. N., and Mayes, L. C. (2019). Investigating the association between parity and the maternal neural response to infant cues. *Soc. Neurosci.* 14, 214–225. doi: 10.1080/17470919.2017.1422276

- McDougall, S., Edworthy, J., Sinimeri, D., Goodliffe, J., Bradley, D., and Foster, J. (2020). Searching for meaning in sound: learning and interpreting alarm signals in visual environments. *J. Exp. Psychol. Appl.* 26, 89–107. doi: 10.1037/xap0000238
- Mogg, K., and Bradley, B. P. (1998). A cognitive-motivational analysis of anxiety. *Behav. Res. Ther.* 36, 809–848.
- Moors, A. (2014). Flavors of appraisal theories of emotion. *Emot. Rev. Emot. Rev.* 6, 303–307. doi: 10.1177/1754073914534477
- Młynarski, W., and McDermott, J. H. (2019). Ecological origins of perceptual grouping principles in the auditory system. *PNAS* 116, 25355–25364. doi: 10.1073/pnas.1903887116
- Navalón, P., Serrano, E., Almansa, B., Perea, M., Benavent, P., Domínguez, A., et al. (2021). Attentional biases to emotional scenes in schizophrenia: an eye-tracking study. *Biol. Psychol.* 160:108045. doi: 10.1016/j.biopsycho.2021.108045
- Nelson, B. D., and Hajcak, G. (2017). Defensive motivation and attention in anticipation of different types of predictable and unpredictable threat: a startle and event-related potential investigation. *Psychophysiology* 54, 1180–1194. doi: 10.1111/psyp.12869
- Öhman, A., and Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychol. Rev.* 108, 483–522. doi: 10.1037/0033-295X.108.3.483
- Olofsson, J. K., Nordin, S., Sequeira, H., and Polich, J. (2008). Affective picture processing: an integrative review of ERP findings. *Biol. Psychol.* 77, 247–265. doi: 10.1016/j.biopsycho.2007.11.006
- Park, G., Van Bavel, J. J., Vasey, M. W., and Thayer, J. F. (2013). Cardiac vagal tone predicts attentional engagement to and disengagement from fearful faces. *Emotion* 13, 645–656. doi: 10.1037/a0032971
- Peschard, V., Gilboa-Schechtman, E., and Philippot, P. (2017). Selective attention to emotional prosody in social anxiety: a dichotic listening study. *Cogn. Emot.* 31, 1749–1756. doi: 10.1080/02699931.2016.1261012
- Peters, M. L., Vieler, J. S. E., and Lautenbacher, S. (2016). Dispositional and induced optimism lead to attentional preference for faces displaying positive emotions: an eye-tracker study. *J. Positive Psychol.* 11, 258–269. doi: 10.1080/17439760.2015.1048816
- Peterson, J. M., Bala, A. D. S., and Dassonville, P. (2022). Differential visual and auditory effects in a crossmodal induced roelofs illusion. *J. Exp. Psychol. Hum. Percept. Perform.* 48, 232–245. doi: 10.1037/xhp0000983
- Phaf, R. H., and Kan, K. J. (2007). The automaticity of emotional stroop: a meta-analysis. *J. Behav. Ther. Exp. Psychiatry* 38, 184–199. doi: 10.1016/j.jbtep.2006.10.008
- Pierce, A. M., McDonald, J. J., and Green, J. J. (2018). Electrophysiological evidence of an attentional bias in cross-modal inhibition of return. *Neuropsychologia* 114, 11–18. doi: 10.1016/j.neuropsychologia.2018.04.007
- Pintzinger, N. M., Pfabigan, D. M., Vienna, V., Kryspin-Exner, I., Lamm, C., and Vienna, A. (2017). Temperament differentially influences early information processing in men and women: preliminary electrophysiological evidence of attentional biases in healthy individuals. *Biol. Psychol.* 122, 69–79. doi: 10.1016/j.biopsycho.2016.07.007
- Pool, E., Brosch, T., Delplanque, S., and Sander, D. (2014). Where is the chocolate? rapid spatial orienting toward stimuli associated with primary rewards. *Cognition* 130, 348–359. doi: 10.1016/j.cognition.2013.12.002
- Pool, E., Brosch, T., Delplanque, S., and Sander, D. (2016). Attentional bias for positive emotional stimuli: a meta-analytic investigation. *Psychol. Bull.* 142, 79–106. doi: 10.1037/bul0000026
- Posner, M. I., and Cohen, Y. (1984). Components of visual orienting. attention and performance X. *Control Lang. Processes* 32, 531–556.
- Posner, M. I., Inhoff, A. W., Friedrich, F. J., and Cohen, A. (1987). Isolating attentional systems—a cognitive-anatomical analysis. *Psychobiology* 15, 107–121.
- Pourtois, G., Schettino, A., and Vuilleumier, P. (2013). Brain mechanisms for emotional influences on perception and attention: what is magic and what is not. *Biol. Psychol.* 92, 492–512. doi: 10.1016/j.biopsycho.2012.02.007
- Preciado, D., Munneke, J., and Theeuwes, J. (2017). Was that a threat? attentional biases by signals of threat. *Emotion* 17, 478–486. doi: 10.1037/emo0000246
- Ren, L., Yang, Z., Cui, L., Jin, Y., Ma, Z., Zhang, Q., et al. (2020). The relations among worry, meta-worry, intolerance of uncertainty and attentional bias for threat in men at high risk for generalized anxiety disorder: a network analysis. *BMC Psychiatry* 20:452. doi: 10.1186/s12888-020-02849-w
- Reutter, M., Hewig, J., Wieser, M. J., and Osinsky, R. (2017). The N2pc component reliably captures attentional bias in social anxiety. *Psychophysiology* 54, 519–527. doi: 10.1111/psyp.12809
- Rosburg, T., and Mager, R. (2021). The reduced auditory evoked potential component N1 after repeated stimulation: refractoriness hypothesis vs. habituation account. *Hearing Res.* 400:108140. doi: 10.1016/j.heares.2020.10.8140
- Ruglass, L. M., Shevorykin, A., Dambreville, N., and Melara, R. D. (2019). Neural and behavioral correlates of attentional bias to cannabis cues among adults with cannabis use disorders. *Psychol. Addict. Behav.* 33, 69–80. doi: 10.1037/adb0000423
- Russell, J. A. (1980). A circumplex model of affect. *J. Pers. Soc. Psychol.* 39, 1161–1178. doi: 10.1037/h0077714
- Rusz, D., Le Pelley, M. E., Kompier, M. A. J., Mait, L., and Bijleveld, E. (2020). Reward-driven distraction: a meta-analysis. *Psychol. Bull.* 146, 872–899. doi: 10.1037/bul0000296
- Sabri, M., Liebenthal, E., Waldron, E. J., Medler, D. A., and Binder, J. R. (2006). Attentional modulation in the detection of irrelevant deviance: a simultaneous ERP/fMRI Study. *J. Cogn. Neurosci.* 18, 689–700. doi: 10.1162/jocn.2006.18.5.689
- Sadek, S. A., Daniel, M. R., and Langdon, P. E. (2020). Attentional bias toward negative and positive pictorial stimuli and its relationship with distorted cognitions, empathy, and moral reasoning among men with intellectual disabilities who have committed crimes. *Aggress. Behav.* Online ahead of print. doi: 10.1002/ab.21908
- Salahub, C., and Emrich, S. M. (2020). Fear not! anxiety biases attentional enhancement of threat without impairing working memory filtering. *Cogn. Affect. Behav. Neurosci.* 20, 1248–1260.
- Sali, A. W., Anderson, B. A., and Yantis, S. (2014). The role of reward prediction in the control of attention. *J. Exp. Psychol. Hum. Percept. Perform.* 40, 1654–1664. doi: 10.1037/a0037267
- Sanchez, A., and Vazquez, C. (2014). Looking at the eyes of happiness: positive emotions mediate the influence of life satisfaction on attention to happy faces. *J. Positive Psychol.* 9, 435–448. doi: 10.1080/17439760.2014.910827
- Sander, D., Grandjean, D., and Scherer, K. R. (2005). A systems approach to appraisal mechanisms in emotion. *Neural Networks* 18, 317–352. doi: 10.1016/j.neunet.2005.03.001
- Santesso, D. L., Meuret, A. E., Hofmann, S. G., Mueller, E. M., Ratner, K. G., Roesch, E. B., et al. (2008). Electrophysiological correlates of spatial orienting towards angry faces: a source localization study. *Neuropsychologia* 46, 1338–1348. doi: 10.1016/j.neuropsychologia.2007.12.013
- Schindler, S., Bruchmann, M., Krasowski, C., Moeck, C., and Straube, T. (2021). Charged with a crime: the neuronal signature of processing negatively evaluated faces under different attentional conditions. *Psychol. Sci.* 32, 1311–1324. doi: 10.1177/0956797621996667
- Schmidtendorf, S., Wiedau, S., Asbrand, J., Tuschen-Caffier, B., and Heinrichs, N. (2018). Attentional bias in children with social anxiety disorder. *Cogn. Therapy Res.* 42, 273–288. doi: 10.1007/s10608-017-9880-7
- Schultz, W. (2004). Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioral ecology. *Curr. Opin. Neurobiol.* 14, 139–147. doi: 10.1016/j.conb.2004.03.017
- Schwerdtfeger, A., and Derakshan, N. (2010). The time line of threat processing and vagal withdrawal in response to a self-threatening stressor in cognitive avoidant copers: evidence for vigilance-avoidance theory. *Psychophysiology* 47, 786–795. doi: 10.1111/j.1469-8986.2010.00965.x
- Sheppes, G., Luria, R., Fukuda, K., and Gross, J. J. (2013). There's more to anxiety than meets the eye: isolating threat-related attentional engagement and disengagement biases. *Emotion* 13, 520–528. doi: 10.1037/a0031236
- Silston, B., and Mobbs, D. (2018). Detecting and responding to threats in the natural world. *Psychol. Inquiry* 29, 28–31. doi: 10.1080/1047840X.2018.1435708
- Strauss, G. P., and Allen, D. N. (2006). The experience of positive emotion is associated with the automatic processing of positive emotional words. *J. Positive Psychol.* 1, 150–159. doi: 10.1080/17439760600566016
- Sutherland, M. R., and Mather, M. (2012). Negative arousal amplifies the effects of saliency in short-term memory. *Emotion* 12, 1367–1372. doi: 10.1037/a0027860
- Tapper, K., Pothos, E. M., and Lawrence, A. D. (2010). Feast your eyes: hunger and trait reward drive predict attentional bias for food cues. *Emotion* 10, 949–954. doi: 10.1037/a0020305
- Thoern, H. A., Grueschow, M., Ehler, U., Ruff, C. C., and Kleim, B. (2016). Attentional bias towards positive emotion predicts stress resilience. *PLoS One* 11:e0148368. doi: 10.1371/journal.pone.0148368
- Topalidis, P., Zinchenko, A., Gädeke, J. C., and Föcker, J. (2020). The role of spatial selective attention in the processing of affective prosodies in congenitally blind adults: an ERP study. *Brain Res.* 1739:146819. doi: 10.1016/j.brainres.2020.146819

- Torrence, R. D., and Troup, L. J. (2018). Event-related potentials of attentional bias toward faces in the dot-probe task: a systematic review. *Psychophysiology* 55, 1–20. doi: 10.1111/psyp.13051
- Torrence, R. D., Wylie, E., and Carlson, J. M. (2017). The time-course for the capture and hold of visuospatial attention by fearful and happy faces. *J. Nonverbal Behav.* 41, 139–153.
- Van Bockstaele, B., Verschuere, B., Tibboel, H., De Houwer, H., Crombez, G., and Koster, E. H. W. (2014). A review of current evidence for the causal impact of attentional bias on fear and anxiety. *Psychol. Bull.* 140, 682–721. doi: 10.1037/a0034834
- Van Vleet, T. M., and Robertson, L. C. (2006). Cross-modal interactions in time and space: auditory influence on visual attention in hemispatial neglect. *J. Cogn. Neurosci.* 18, 1368–1379. doi: 10.1162/jocn.2006.18.8.1368
- Victeur, Q., Huguet, P., and Silvert, L. (2020). Attentional allocation to task-irrelevant fearful faces is not automatic: experimental evidence for the conditional hypothesis of emotional selection. *Cogn. Emot.* 34, 288–301. doi: 10.1080/02699931.2019.1622512
- Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends Cogn. Sci.* 9, 585–594. doi: 10.1016/j.tics.2005.10.011
- Vuilleumier, P. (2015). Affective and motivational control of vision. *Curr. Opin. Neurol.* 28, 29–35. doi: 10.1097/WCO.0000000000000159
- Wadley, G., Smith, W., Koval, P., and Gross, J. J. (2020). Digital emotion regulation. *Curr. Direct. Psychol. Sci.* 29, 412–418. doi: 10.1177/0963721420920592
- Wadlinger, H. A., and Isaacowitz, D. M. (2008). Looking happy: the experimental manipulation of a positive visual attention bias. *Emotion* 8, 121–136. doi: 10.1037/1528-3542.8.1.121
- Wang, A., Lu, F., Gao, W., Zhang, T., and Zhang, M. (2022). Reward weakened inhibition of return (IOR) in the near depth plane. *Perception* 51, 114–130. doi: 10.1177/03010066211073855
- Wang, Y., Xiao, R. Q., Luo, C., and Yang, L. (2019b). Attentional disengagement from negative natural sounds for high-anxious individuals. *Anxiety Stress Coping* 32, 298–311. doi: 10.1080/10615806.2019.1583539
- Wang, Y., Xiao, R., and Luo, C. (2019a). Mechanisms underlying auditory and cross-modal emotional attentional biases: engagement with and disengagement from aversive auditory stimuli. *Motivation Emot.* 43, 354–369.
- Weierich, M. R., Treat, T. A., and Hollingworth, A. (2008). Theories and measurement of visual attentional processing in anxiety. *Cogn. Emot.* 22, 985–1018.
- Wieser, M. J., Hambach, A., and Weymar, M. (2018). Neurophysiological correlates of attentional bias for emotional faces in socially anxious individuals – evidence from a visual search task and N2pc. *Biol. Psychol.* 132, 192–201. doi: 10.1016/j.biopsycho.2018.01.004
- Willner, C. J., Gatzke-Kopp, L. M., Segalowitz, S. J., Bierman, K. L., and Greenberg, M. T. (2015). Relevance of a neurophysiological marker of attention allocation for children's learning-related behaviors and academic performance. *Dev. Psychol.* 15, 1148–1162. doi: 10.1037/a0039311
- Woldorff, M., Hansen, J. C., and Hillyard, S. A. (1987). Evidence for effects of selective attention in the mid-latency range of the human auditory event-related potential. *Electroencephalography Clin. Neurophysiol.* 40, 146–154.
- Woltering, S., Chen, S., and Jia, Y. (2021). Neural correlates of attentional bias to food stimuli in obese adolescents. *Brain Topogr.* 34, 182–191. doi: 10.1007/s10548-020-00812-2
- Yang, W., Makita, K., Nakao, T., Kanayama, N., Machizawa, M. G., Sasaoka, T., et al. (2018). Affective auditory stimulus database: an expanded version of the international affective digitized sounds (IADS-E). *Behav. Res. Methods* 50, 1415–1429. doi: 10.3758/s13428-018-1027-6
- Yee, C. M., and Miller, G. A. (1994). A dual-task analysis of resource allocation in dysthymia and anhedonia. *J. Abnorm. Psychol.* 103, 625–636. doi: 10.1037//0021-843x.103.4.625
- Yiend, J. (2010). The effects of emotion on attention: a review of attentional processing of emotional information. *Cogn. Emot.* 24, 3–47. doi: 10.1080/02699930903205698
- Yuan, J., Tian, Y., Huang, X., Fan, H., and Wei, X. (2019). Emotional bias varies with stimulus type, arousal and task setting: meta-analytic evidences. *Neurosci. Biobehav. Rev.* 107, 461–472. doi: 10.1016/j.neubiorev.2019.09.035
- Yuan, J., Zhang, Q., and Cui, L. (2021). Disgust face captures more attention in individuals with high social anxiety when cognitive resources are abundant: evidence from N2pc. *Neuropsychologia* 151:107731. doi: 10.1016/j.neuropsychologia.2020.107731
- Zhai, X., Khatami, F., Sadeghi, M., He, F., Read, H., Stevenson, I., et al. (2020). Distinct neural ensemble response statistics are associated with recognition and discrimination of natural sound textures. *PNAS* 117, 31482–31493. doi: 10.1073/pnas.2005644117
- Zimmer, U., Höfler, M., Koschutnig, K., and Ischebeck, A. (2016). Neuronal interactions in areas of spatial attention reflect avoidance of disgust, but orienting to danger. *NeuroImage* 134, 94–104. doi: 10.1016/j.neuroimage.2016.03.050
- Zimmer, U., Keppel, M. T., Poglitsch, C., and Ischebeck, A. (2015). ERP evidence for spatial attention being directed away from disgusting locations. *Psychophysiology* 52, 1317–1327.
- Zimmer, U., Rosenzopf, H., Poglitsch, C., and Ischebeck, A. (2019). ERP-study on the time course of disgust-motivated spatial avoidance. *Biol. Psychol.* 144, 20–27. doi: 10.1016/j.biopsycho.2019.02.007
- Zuk, N. J., Teoh, E. S., and Lalor, E. C. (2020). EEG-based classification of natural sounds reveals specialized responses to speech and music. *NeuroImage* 210:116558. doi: 10.1016/j.neuroimage.2020.116558

Appendix

TABLE A1 List of emotional auditory al sounds used in the current study.

Number	Description	Valence
206	shower	neutral
262	yawn	neutral
278	walking	neutral
318	sound of zipper	neutral
372	wing shaking	neutral
442	temple bell	neutral
483	clock	neutral
730	typing	neutral
829	page turning	neutral
922	washing clothes	neutral
111	music box	positive
123	sea	positive
151	robin	positive
172	brook	positive
200	valley	positive
298	gurgling water	positive
351	applause	positive
419	bird song	positive
1038	gurgling water	positive
1385	baby laugh	positive
255	vomit	negative
273	crash	negative
276	scream	negative
292	male scream	negative
293	bomb	negative
296	women crying	negative
432	shooting1	negative
434	shooting2	negative
602	thunderstorm	negative
626	explosion	negative



OPEN ACCESS

EDITED BY

Gaoxiang Ouyang,
Beijing Normal University, China

REVIEWED BY

Yue Gu,
Tianjin University of Technology, China
Kazuki Hyodo,
Physical Fitness Research Institute,
Meiji Yasuda Life Foundation of Health
and Welfare, Japan

*CORRESPONDENCE

Raimundo da Silva Soares Jr.
raimundo.silva@ufabc.edu.br

SPECIALTY SECTION

This article was submitted to
Cognitive Neuroscience,
a section of the journal
Frontiers in Human Neuroscience

RECEIVED 07 March 2022

ACCEPTED 27 June 2022

PUBLISHED 12 August 2022

CITATION

da Silva Soares R Jr, Ambriola Oku AY,
Barreto CSF and Ricardo Sato J (2022)
Applying functional near-infrared
spectroscopy and eye-tracking in a
naturalistic educational environment
to investigate physiological aspects
that underlie the cognitive effort
of children during mental rotation
tests.
Front. Hum. Neurosci. 16:889806.
doi: 10.3389/fnhum.2022.889806

COPYRIGHT

© 2022 da Silva Soares, Ambriola Oku,
Barreto and Ricardo Sato. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Applying functional near-infrared spectroscopy and eye-tracking in a naturalistic educational environment to investigate physiological aspects that underlie the cognitive effort of children during mental rotation tests

Raimundo da Silva Soares Jr.^{1,2*},
Amanda Yumi Ambriola Oku¹, Cândida S. F. Barreto³ and
João Ricardo Sato¹

¹Center for Mathematics, Computation and Cognition, Universidade Federal do ABC, São Bernardo do Campo, Brazil, ²Graduate Program in Neuroscience and Cognition, Federal University of ABC, São Bernardo do Campo, Brazil, ³South African National Research Foundation Research Chair, Faculty of Education, University of Johannesburg, Johannesburg, South Africa

Spatial cognition is related to academic achievement in science, technology, engineering, and mathematics (STEM) domains. Neuroimaging studies suggest that brain regions' activation might be related to the general cognitive effort while solving mental rotation tasks (MRT). In this study, we evaluate the mental effort of children performing MRT tasks by measuring brain activation and pupil dilation. We use functional near-infrared spectroscopy (fNIRS) concurrently to collect brain hemodynamic responses from children's prefrontal cortex (PFC) and an Eye-tracking system to measure pupil dilation during MRT. Thirty-two healthy students aged 9–11 participated in this experiment. Behavioral measurements such as task performance on geometry problem-solving tests and MRT scores were also collected. The results were significant positive correlations between the children's MRT and geometry problem-solving test scores. There are also significant positive correlations between dorsolateral PFC (dlPFC) hemodynamic signals and visuospatial task performances (MRT and geometry problem-solving scores). Moreover, we found significant activation in the amplitude of deoxy-Hb variation on the dlPFC and that pupil diameter increased during the MRT, suggesting that both physiological responses are related to mental effort processes during the visuospatial task. Our findings indicate that children with more mental effort under the task performed better. The multimodal approach to monitoring students' mental effort can be of great interest in providing objective feedback on cognitive resource conditions and advancing our comprehension of the neural mechanisms that underlie cognitive effort. Hence, the ability to detect

two distinct mental states of rest or activation of children during the MRT could eventually lead to an application for investigating the visuospatial skills of young students using naturalistic educational paradigms.

KEYWORDS

spatial cognition, cognitive effort, pupillometry, fNIRS, naturalistic experimentation, educational research

Introduction

Spatial cognition is usually described as “the capacity to generate, retain, retrieve, and transform well-structured visual images in mind” (Lohman, 1994). This ability to understand and manipulate the objects’ spatial characteristics has been associated with academic achievement in science, technology, engineering, and mathematics (STEM) domains (Wai et al., 2009; Khine, 2016). Psychology research has often measured spatial cognition via visuospatial tasks, including the mental rotation task (MRT) paradigm (Shepard and Metzler, 1971, 1988; Bruce and Hawes, 2015; Sladky et al., 2016). MRT requires the subject to determine whether pairs of figures exhibited are the same or mirror images. To accomplish this task, the subject creates a mental image and rotates it into a different orientation, which is related to spatial reasoning, visuospatial ability, and resembles everyday situations that require spatial orientation and the capacity to manipulate objects mentally (Guillot et al., 2012).

In cognitive neuroscience, the MRT has been applied to different ages and neural development stages. Studies with adolescents and adults indicate a relationship between MRT and mathematical achievement (Kytälä and Lehto, 2008; Tolar et al., 2009; Wei et al., 2012). Visuospatial cognition studies on children highlight the relevance of spatial ability in many scientific domains and support intervention to develop spatial skills (Webb et al., 2007; Tzuriel and Egozi, 2010). However, there are very few studies that investigate the brain activity of children during MRT (Roberts and Bell, 2000; Wu et al., 2020). Indeed, many studies examine the neural substrate associated with MRT and its relationship to adults’ learning process (Jordan et al., 2001; Harris and Miniussi, 2003; Zacks et al., 2003; Zacks, 2008). Neuroimaging studies have shown that some brain regions, including the superior parietal [Brodmann’s area (BA) 40], premotor cortex (BA 6), and the frontal cortices [dorsolateral prefrontal cortex (dlPFC) BAs 9 and 46; ventrolateral prefrontal cortex (vlPFC) BA 44], are involved with bilateral activation regions during the MRT (Cohen et al., 1996; Hugdahl et al., 2006; Zacks, 2008; Wu et al., 2020). Other studies found that the modulation of activity in the frontal cortex reflects the participant’s cognitive effort, i.e., mental resources required under cognitive demands (Prat et al., 2007;

Ayaz et al., 2010, 2012a,b; Fishburn et al., 2014; Causse et al., 2017). For example, Causse et al. (2017) demonstrated that the functional near-infrared spectroscopy (fNIRS) was sensitive to task difficulty variations. Also, the prefrontal activation intensity sheds insight on the level of mental effort, indicating the amount of cognitive workload required by demanding tasks. Importantly, there are age-related differences in cognitively demands of mental processing and performance (McCabe et al., 2010). Older adults tend to express a higher effortful reaction and lower performance on cognitively demanding activity (Hess and Ennis, 2012; Westbrook et al., 2013), suggesting that poor performance could be a sign of depletion or reduced cognitive resources. It seems that greater levels of mental effort are related to lower performance (Hess and Ennis, 2012). Yet, little is known about how schoolchildren’s cognitive effort relates to students’ performance (Chevalier, 2018), and physiological measurements could be helpful to address this matter.

Eye-tracking (ET) is another reliable technological tool for research in the educational context. It provides a non-invasive and real-time measurement of pupil sizes from subjects during cognitive tasks (Lum et al., 2017; McGarrigle et al., 2017). Pupillometry studies indicate that pupil diameter changes in response to mental activity and increases with task difficulty in children (Boersma et al., 1970; Karatekin, 2004; Karatekin et al., 2007; Chatham et al., 2009; Johnson et al., 2014; Rojas-Libano et al., 2019) and adults (Hess and Polt, 1964; Beatty and Kahneman, 1966; Granholm et al., 1996). Pupillometry measures during cognitive demanding mental arithmetic tasks suggest that pupil dilation may indicate mental activity, which includes mental arithmetic tasks (Hess and Polt, 1964; Ahern and Beatty, 1979; Jainta and Baccino, 2010; Klingner et al., 2011) and spatial ability tasks (Buckley et al., 2018; Campbell et al., 2018; Bauer et al., 2021a,b). For example, Toth and Campbell (2019) showed that the cognitive demand of the MRT elicited increases in participants’ pupil diameters (Toth and Campbell, 2019). It is well known that the noradrenergic system has been implicated in mental effort and pupil dilation. However, the precise mechanism related to PFC activity and fluctuations in pupil size is still not fully understood (Bradley et al., 2008; Mathot, 2018).

There is growing recognition that fNIRS is moderately tolerant to motion artifacts and is a portable, cost-effective

device (Ferrari and Quresima, 2012; Curtin and Ayaz, 2018). All these together make fNIRS a suitable tool to use in conjunction with an eye-tracker device for monitoring the brain activity of children in a naturalistic setting such as realistic educational situations. Together, pupillometry and neuroimaging techniques are valuable for integrating mental effort signs and shedding light on the neural mechanisms that underlie cognitive load (Hosseini et al., 2017; İşbilir et al., 2019). Although previous studies have investigated the cognitive effort's neural mechanism in adults, little is known about children's brain activity under visuospatial cognitive demands. Considering the spatial cognition relevance to school achievement (Giofrè et al., 2013), physiological measurements of students' mental effort during visuospatial tasks will make it possible to advance the understanding of cognitive aspects related to academic performance, specially geometry problem-solving development. We believe that traditional cognitive load measures could benefit from ET-fNIRS-based mental workload detection by combining additional information sources related to the mental effort of children during challenging tasks. Such a multimodal approach has the potential to contribute to a better understanding of young students' cognitive resources that underlie the learning process.

Multimodal applications can enhance physiological response detection from different modalities, which enable more accurate and robust measurements of cognitive states (Siegle et al., 2003; Shin et al., 2018), including mental effort (Laeng et al., 2012; Rozado and Andreas, 2015). Therefore, we propose the ET-fNIRS fusion as a promising framework for future investigations in educational research. Here, we sought to shed light on physiological responses related to children's cognitive resources during visuospatial tasks. More specifically, we wanted to improve our understanding of the relationship between pupil dilatation and hemodynamic PFC activity during a visuospatial task in children. Therefore, we use a multimodal approach by applying fNIRS and Eye-tracking simultaneously to investigate primary school children's mental effort expressed by brain hemodynamics changes and pupil size during the MRT. Based on the research cited above, we hypothesized that the higher bilateral PFC activity and pupil dilatation might be associated with higher mental effort during the spatial task and, consequently, lower children's performance (i.e., visuospatial scores).

Materials and methods

Participants

The local Ethics Committee of UFABC approved all aspects of our experiment. The experiment was performed following all relevant guidelines and federal regulations. All

subjects participated voluntarily and without any financial compensation. Thirty-two healthy students (fifteen boys and sixteen girls, 9–11 years old) from the 5th grade of a Brazilian elementary public school participated in this study. Parents/legal guardians of the young children provided written informed consent and were informed verbally of the purpose of the study and the safety of the experiments. All participants had normal or corrected to normal vision, and did not report present or previous neurological or psychiatric conditions.

Eye-tracking data acquisition

The pupil diameter was recorded using an ASL Mobile Eye-5 (Applied Sciences Laboratory, Bedford, MA, United States) with a capture rate of 60 Hz. The glasses were mounted with one small camera over the right eye. The ASL Mobil Eye-5 auto-calibrates to lighting conditions in the environment and compensates for head movements.

Functional near-infrared spectroscopy data acquisition

The hemodynamic signals were recorded using the continuous-wave NIRx Sport System (NIRx Medical Technologies, Glen Head, NY, United States). The cap montage was an array of eight pairs of optodes (sources and detectors), resulting in 28 channels (twenty long-range channels and eight short-distance channels) placed bilaterally over the prefrontal cortex (Figure 1), including the brain regions of the premotor cortex (BA 6), and dlPFC (BAs 9, 10, and 46; Koessler et al., 2009) (see **Supplementary material 1** and **Supplementary Figure 1** for more details). Source and detector distance were 30 mm with optodes positioned on the measuring cap regarding the 10–10 international system (Jurcak et al., 2007). The illumination sources emitted two wavelengths of near-infrared light (760 and 850 nm), and the signals were recorded at a sampling rate of 7.91 Hz. A dark cloth was placed over the measuring cap to reduce noise from the external light. We used the software NIRxStar 15.2 (NIRx Medical Technologies, Glen Head, NY, United States) to record the brain activity data and to evaluate the signal quality of the channels.

Geometry problem-solving test

To assess the geometry problem-solving performance, we selected the questions from the Brazilian Basic Education Assessment System (SAEB). It is a national test that assesses the mathematical and language skill levels of Brazilian students from the first and last year of elementary school,

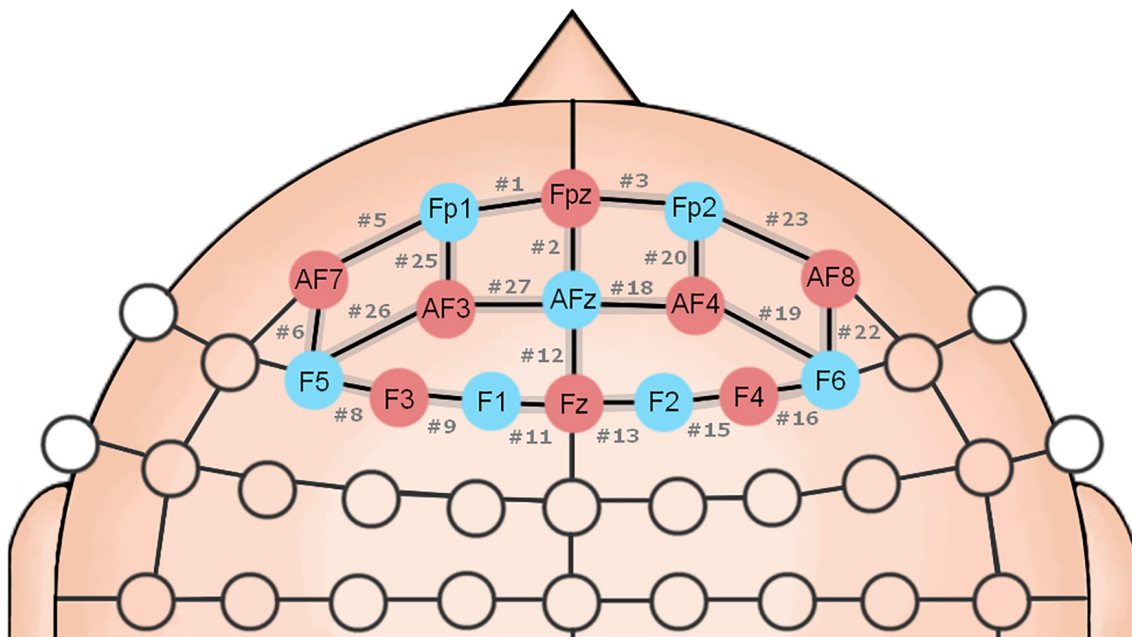


FIGURE 1

Schematic representation of the fNIRS optodes positioned over the scalp. The position of the optodes follows the universal configuration of the 10-10 EEG international system. Channels are composed of neighboring optodes represented by the black edges, sources in red (S1–S8) corresponding to F3, AF7, AF3, Fz, Fpz, AF4, AF8, F4, and detectors in blue (D1–D7) matching F5, F1, Fp1, AFz, F2, Fp2, F6. # channel.

5th and 9th grades, respectively (Carnoy et al., 2015). We selected and printed eleven geometry multiple-choice problems from the SAEB test public database (SAEB, 2011). The test was designed to measure the general geometry problem-solving skills of 5th graders, and we adapted the questions in such a way that the students did not know the results beforehand.

Mental rotations test

The MRT was adapted for the 9–11 years old children. The task consisted of 70 slides with images in two dimensions following the geometrical patterns (abstract figures). There were two kinds of abstract figures (Figure 2) in the stimuli session: (1) one gray abstract figure placed on the center of the slide with a gray arrow indicating the left or right direction of rotation to the next slide, and (2) a pair of gray abstract figures identified as “A” and “B” with a similar form, but different orientations. Only one of the options was a 90° rotation of the figure presented before. The task requires children to imagine the movement of rotation in the first slide to decide, in the next slide, which of the choices, “A” or “B,” match the stimuli presented before. There were ten trials with three slides of stimuli, three slides of alternatives, and one slide of the control condition in which subjects were instructed to fixate on the cross until the onset of the stimuli. The

child had 30 s to rest, 7 s to imagine the rotated geometric figure, and 3 s to choose the alternative, three times per block, yielding 30 s of stimuli. All the sessions had the same level of difficulty to counterbalance the different phases of the task. We used the software NIRxStim (NIRx Medical Technologies, Glen Head, NY, United States) to present the visual stimuli and collect the fNIRS data.

Experimental design

The MRT was performed in the students’ usual place at school. Each child participated in the task individually in a session that lasted approximately 30 min. First, students were randomly called and asked to answer eleven geometry questions from the Brazilian basic education assessment system (SAEB). After solving all the eleven mathematical problems with multiple choices printed on paper, participants were assigned to another room fitted with the fNIRS and Eye-tracker. Two experimenters controlled the equipment and observed the child during the task. Participants were instructed about the experiment and asked to answer a practice trial of MRT to make sure that they understood the task. If the child made a mistake, the experimenter would repeat the trial. After the instruction trials, experimenters set up the fNIRS and eye-tracker. Students were then told that they would next play on the computer screen the same kind of game as in the

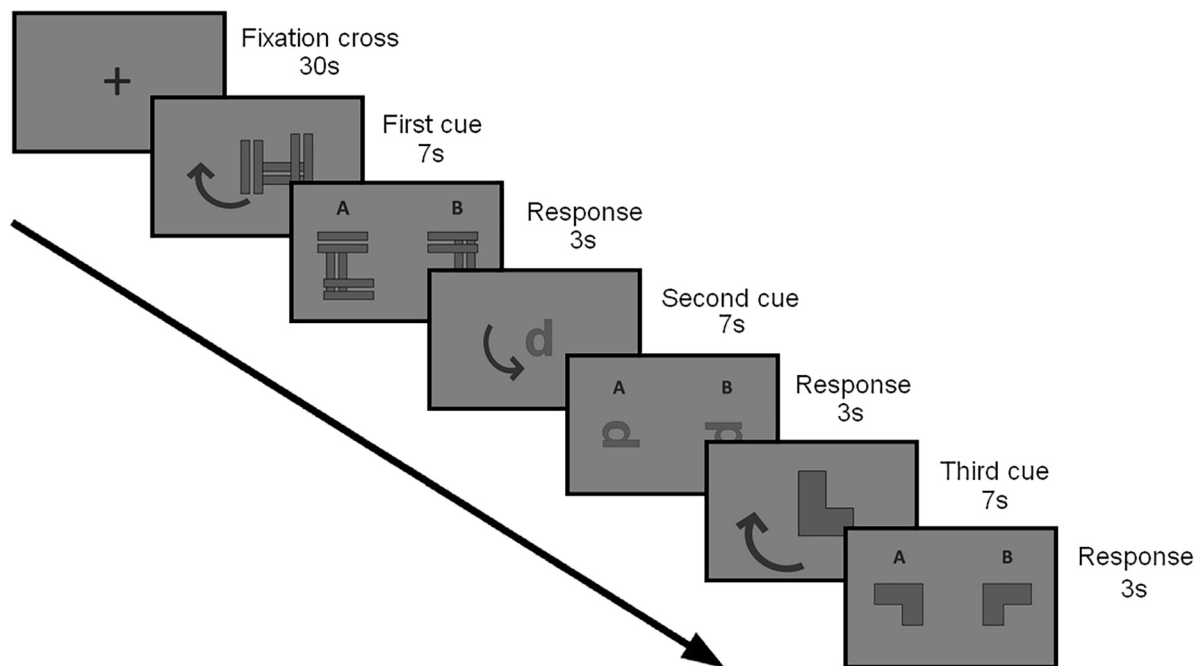


FIGURE 2

Schematic diagram of the experimental procedure. Fixation cross (the 30 s) was followed by the first cue (7 s) as one gray abstract figure placed on the center of the slide with a gray arrow indicating the left or right direction of rotation. A stimulus “response” was then presented (3 s) as a pair of gray abstract figures identified as “A” and “B” with a similar form, but different orientations at 90° according to the cue (to the right or left visual field) so the participants could answer verbally. The child had 30 s to rest, 7 s to imagine the rotated geometric figure, and 3 s to choose the alternative, three times per block, yielding 30 s of stimuli.

instruction trials of MRT. The stimuli were presented, followed by the choice “A” and “B,” in which students answered verbally. Participants’ answers were collected during the experiment for further analysis.

Eye-tracking data analysis

Five children were excluded from the sample due to excessive missing pupil data on the MRT (bad frames), yielding data from twenty-six children. The data were inspected for artifacts such as eye blinks or undetected signals during the task. The high-frequency noise also was also considered an eye tracker artifact. Therefore, we removed the noise from the data by smoothing it using a moving median (Gao et al., 2007). Then we implemented a computer algorithm to automate the artifact removal process for all subjects. We considered recordings insufficient when less than 40% of data remained after cleaning the data to remove artifacts. Children’s median pupil dilation variation (%) was calculated over each second of the task compared to the previous control condition. Then, we determined the median across the experimental trials. The 10 s interval preceding each stimuli block was treated as the baseline. Finally, a *t*-test was applied to test significant

differences in the pupil dilation mean between control and task conditions.

Functional near-infrared spectroscopy data analysis

The hemodynamic signals were processed and analyzed by the software nirsLAB v2019.04 (NIRx Medical Technologies, Glen Head, NY, United States). The signals were truncated before the first block and after the end of the last block. Then, the modified Beer-Lambert law was applied to convert the optical signals of each wavelength (760 and 850 nm) in concentration changes of oxy-Hb and deoxy-Hb. The general linear model (GLM) method (Huppert, 2016) was applied to analyze changes in both oxy-Hb and deoxy-Hb amplitudes during the MRT. In the individual-level analysis, we obtained the activation beta coefficients for each channel of the participants. Then, we used a group GLM to combine all participants’ beta coefficients yielding the group statistical activation map (*t*-test). The rate of accepted false-positive results was set at 5%. Bonferroni correction was applied for multiple comparisons for 20 long-distance channels (resulting *p*-value threshold: <0.0025). The hemodynamic signals were extracted from the channels that

presented significant activation in the group analysis. Finally, we calculated the signal block average among the participants to evaluate the channel's hemodynamic response. The spatial representation of the brain regions related to each channel was rendered from the BrainNet Viewer toolbox (Xia et al., 2013)¹.

Behavior data analysis

We analyzed the children's performance on the geometry problem-solving test and the MRT by counting the correct answers on tasks. The scores on both visuospatial tasks were correlated by Spearman's correlation coefficient. We also performed Spearman's correlation analysis to quantify the potential association between behavioral data with pupillary responses and cortical activation. In all cases, $p < 0.05$ (uncorrected) was considered statistically significant. When the short-distance channels were statistically significant to behavioral data, we considered the long-distance correspondent channel insufficient quality data for correlation analysis.

Results

Pupillary response

The timeline analysis of the pupil size data indicated three distinct periods of pupillary variation (%) from the beginning (0 s) until the end of the MRT ($p = 0.0006$), as shown in **Figure 3**. First, there was a period without any significant stimuli (10 s of resting state as the baseline), followed by an increase in the children's pupil size with the stimulus display. The results show a peak of pupil dilation around 10 s after the children imagined the figure. It suggests a pupil dilatation follows a visuospatial task. The responses to the mental rotation stimuli appeared to elicit more pupil dilation than the neutral stimuli. It indicates that it is possible to identify two distinct mental states of rest or activation during the MRT. There is no significant correlation between the mean amplitude of pupil dilation and the number of correct answers, neither in geometric problem-solving nor in the MRT ($p > 0.05$ in all cases).

Task performance

Spearman's correlation analysis revealed that task performance achieved during the MRT was statistically significantly correlated with the geometry problem-solving performance ($r = 0.38$; $p = 0.03$) (**Figure 4**).

Brain-behavior correlation

The Spearman's correlation analysis revealed correlation between MRT scores and oxy-Hb variation at channel 25 ($r = 0.43$; $p = 0.01$; **Figure 5B**) and channel 5 ($r = 0.41$; $p = 0.02$), but we considered the channel 5 as insufficient quality data because the corresponding short-distance was also significantly correlated to the task performance (Channel 7; $r = 0.38$; $p = 0.03$; see **Supplementary material 3**). There is also significant correlation between geometry performance and deoxy-Hb variation at channel 8 ($r = 0.40$; $p = 0.02$; **Figure 5C**), at short-distance channels 4 ($r = 0.33$; $p = 0.007$), 7 ($r = 0.41$; $p = 0.02$), 24 ($r = 0.39$; $p = 0.02$) and oxy-Hb variation at short-channel 7 ($r = 0.39$; $p = 0.03$; see **Supplementary material 3**). All other comparisons were not significant. The spatial representation of the brain regions related to each channel is shown in **Figure 5A**, for the mental rotation tasks scores and the geometry test scores (channels respectively, highlighted in purple and orange). Channel 25 is on the left prefrontal cortex (PFC), and channel 8 is on the left dorsolateral prefrontal cortex (dlPFC).

Cortical activation

The results revealed a significant deactivation (Deoxy-Hb) of the left dlPFC (Ch.6; p -value corrected = 0.002), and the right dlPFC (Ch.22; p -corrected = 0.001; and Ch.23; p -value corrected <0.001, see all results in **Supplementary material 2**). Brain regions represented in **Figure 6A** show that the channels associated with the correlation between Oxy-Hb variation and MRT are placed on dlPFC. Mean changes in oxy-Hb levels in the other prefrontal regions that we evaluated were not significant. **Figure 6B** shows the mean group oxy-Hb and deoxy-Hb signal change during the MRT for the channels 6, 22, and 23 (Ch.6, Ch.22, and Ch.23, respectively) indicating that the oxy-Hb concentration increase while the changes in deoxy-Hb were in the opposite direction, as expected.

Discussion

This study aimed to investigate the 9- to 11-year-old schoolchildren's cognitive effort expressed by changes in pupil diameter and PFC neural activity under spatial cognitive demand. Our findings suggest that the visuospatial task was able to demand students' mental effort, which was detectable by the multimodal approach using pupillometry and optical neuroimage. To our knowledge, this is the first study to apply fNIRS and Eye-tracking to assess the cognitive resources required under mental demands through cortical hemodynamic activity and the pupil responses of children during the MRT.

¹ <https://www.nitrc.org/projects/bnv>

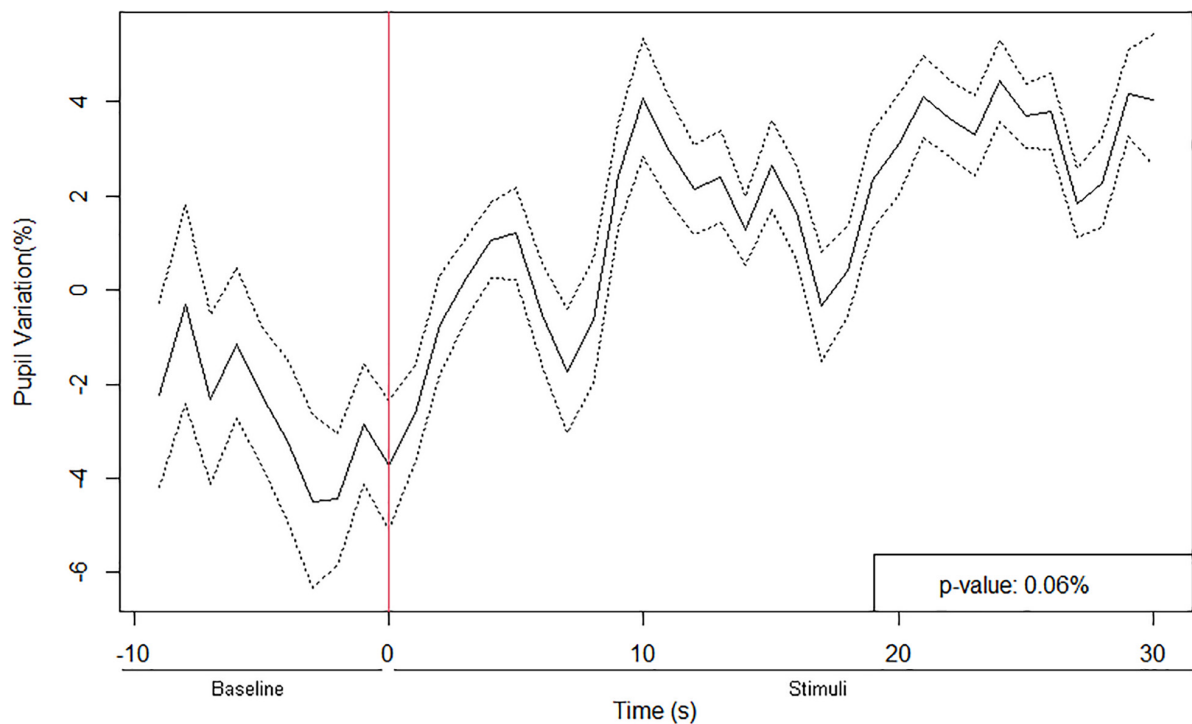


FIGURE 3

Pupil diameter variation timeline during the MRT. The red line indicates the beginning of the task (stimuli), which had three different figures to rotate per block, yielding 30 s of stimuli ($p = 0.0006$) and 10 s of baseline. The solid line indicates the mean (μ) and the dashed line indicates ± 1 standard error (σ).

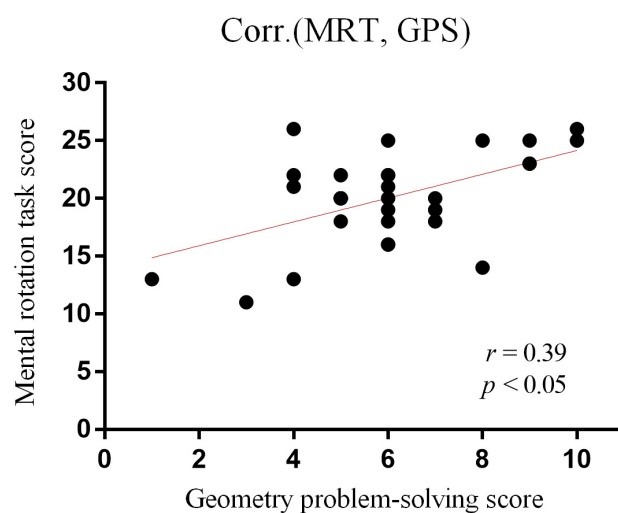


FIGURE 4

Scatter plot between the number of correct responses on MRT and geometry problem-solving test (scores).

The results indicate a significant positive correlation between the children's MRT and geometry test scores. There are also significant positive correlations between dlPFC hemodynamic signals and visuospatial task performances (MRT and geometry problem-solving scores). Our results

are consistent with dlPFC activity related to higher-level cognitive processing (Miller and Cohen, 2001; Vassena et al., 2014, 2019). Moreover, we found significant activation in the amplitude of deoxy-Hb variation on the dlPFC and that pupil diameter increased during the MRT, suggesting

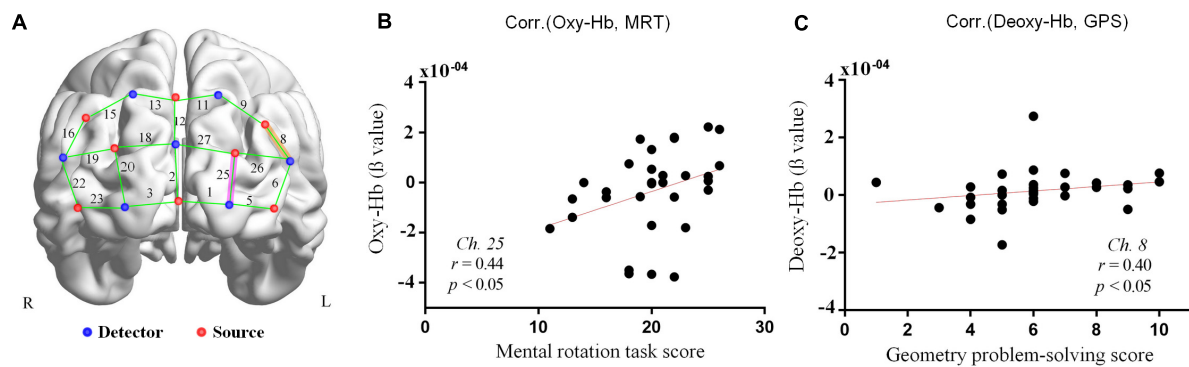


FIGURE 5

Correlation graphs and fNIRS channels positions. **(A)** 3D-rendered brain template (ICBM152) with left (L) and right (R) hemispheres from BrainNet Viewer toolbox (Xia et al., 2013). The fNIRS channels are positioned across the cortex. The red points represent sources, the blue points are detectors, and the green lines are channels. The purple highlight indicates the channel associated with the correlation between oxy-Hb variation and mental rotation task (MRT). The orange highlight indicates the channel associated with the correlation between Deoxy-Hb variation and geometry problem-solving (GPS). **(B)** Scatter plots between the number of correct responses on MRT and oxy-Hb variation at channel 25 ($r = 0.43$; $p = 0.01$). **(C)** Scatter plots between the number of correct responses on the geometry test and deoxy-hemoglobin (deoxy-Hb) variation at channel 8 ($r = 0.4$; $p = 0.02$).

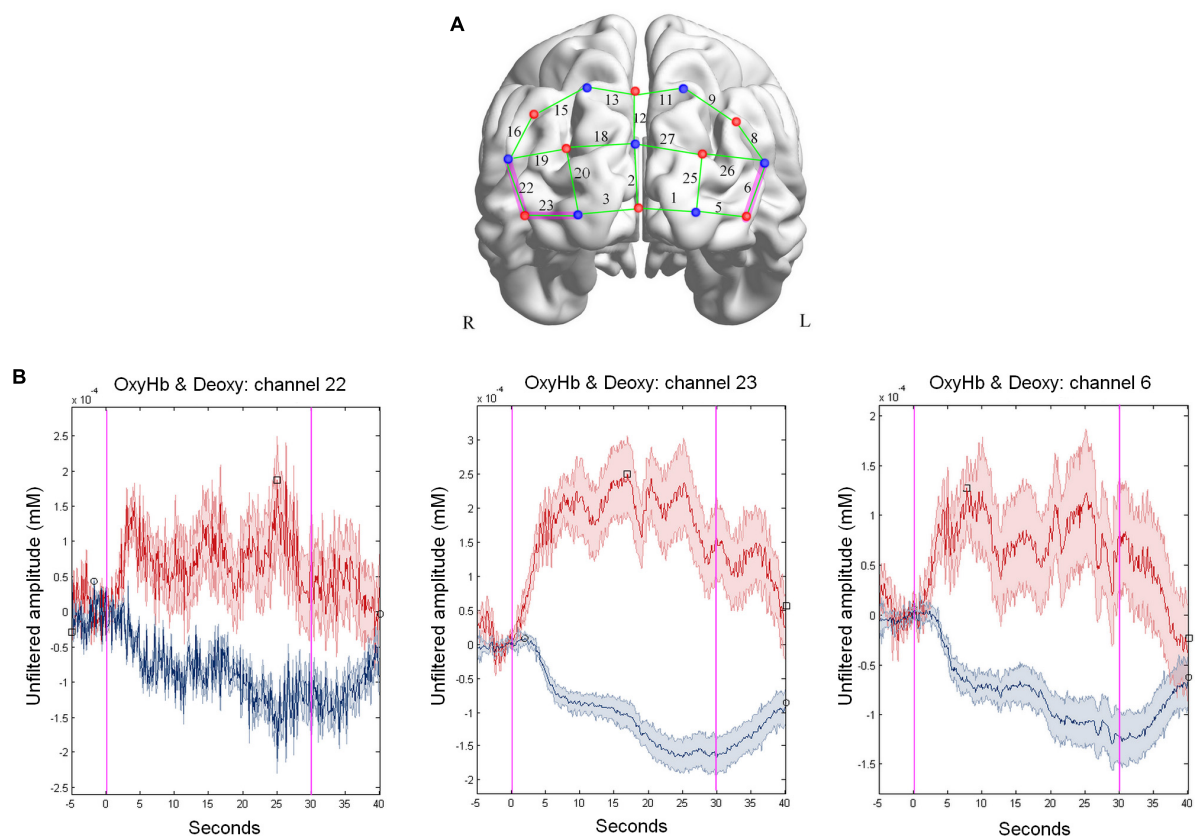


FIGURE 6

Group level HbO cortical activation map. **(A)** 3D-rendered brain template (ICBM152) with left (L) and right (R) hemispheres from BrainNet Viewer toolbox (Xia et al., 2013). The fNIRS channels are positioned across the cortex. The red points represent sources, the blue points are detectors, and the green lines are channels. The purple highlight indicates the channel associated with the correlation between deoxy-Hb variation and mental rotation task (MRT). Thresholded SPM t image, p -value = 0.0025 for deoxy-Hb. **(B)** Block-averaged hemodynamic response (mean \pm SD) computed by averaging the oxy-Hb (red) and deoxy-Hb (blue) signals during the MRT for channels 6, 22, and 23.

that both physiological responses are related to mental effort processes during the visuospatial task. Our findings indicate that children with more mental effort under the task performed better, which was unexpected. According to the general cognitive load theory (Sweller and Chandler, 1991; Sweller et al., 1998), better problem solvers tend to express less mental effort than novice learners' performances, as demonstrated in previous studies (Çakır et al., 2011; Shaw et al., 2018). A possible explanation is an age-related difference in our study. Unlike adults, children do not have well-developed visuospatial skills, so they need more activation to solve geometric problems and perform well. Despite the potential to investigate learning and development with eye-tracking, few studies used task-evoked pupil changes analysis on children (Eckstein et al., 2017). So, it is possible to apply pupillometry to elucidate students' cognitive processes and development.

Spatial cognition improves with age in childhood (Orde, 1996) and declines with age in adulthood (Pak, 2001). Furthermore, studies investigating cognitive efficiency in young and older adults show age-related differences in mental effort and performance (Baltes et al., 1999; Westbrook et al., 2013; Westbrook and Braver, 2015). Older adults tend to express a higher effortful reaction and lower performance on work memory tasks (McCabe et al., 2010; Westbrook et al., 2013) and math problem-solving (Hess and Ennis, 2012). On the other hand, studies with children demonstrated that young children primarily engage the frontal cortex when solving numerical tasks, decreasing by age and differences in brain development (Houdé et al., 2010; Artemenko et al., 2018; Soltanlou et al., 2018a,b). In this sense, we believe that adults would have different findings from children because of neural development. Such information has important implications for understanding students' spatial cognition and developing an intuitive understanding of fundamental geometric concepts.

A significant positive correlation was observed between MRT and geometry problem-solving performance. It indicates that children with good performance in figure rotation also present better performances in our geometry test based on the Brazilian mathematics assessment. It corroborates that our visuospatial task requires mental rotation skills associated with geometry problem-solving abilities. Geometry is a fundamental approach to interpreting the physical environment and everyday situations (Hwang et al., 2009). It contains basic mathematical concepts and demands spatial cognition during the learning process of shapes, size, quantity, and length (Biber et al., 2013; Hertanti et al., 2019). Previous studies have shown that spatial cognition is a good predictor of mathematics achievement, specifically geometry (Delgado and Prieto, 2004). Similarly, our results are in line with the idea that

spatial cognition is relevant to geometry performance. Our analysis correlation between MRT and geometry problem-solving underscores the critical role that spatial cognition plays in mathematical achievement and has implications for educational practices.

There seems to be a lack of evidence for neural correlates of children's cognitive effort and MRT, maybe for neuroimaging technique limitations on movement artifacts. However, the last few years have shown an increasing technology development that allows studies in schools with children (Soltanlou et al., 2018a), which includes naturalistic experimental settings with fNIRS (Ferrari and Quaresima, 2012; Balardin et al., 2017; Barreto et al., 2021) and Eye-Tracking (Epelboim and Suppes, 2001; Bolden et al., 2015) outside the laboratory environment. Our study is in line with the idea that fNIRS and eye-tracking are suitable for measuring the cognitive processes of schoolchildren in naturalistic settings (Mücke et al., 2018), and eye-tracking is a valuable device to use to measure physiological activations during spatial cognition stimuli in a familiar environment.

In addition to measuring cortical oxygenation, pupil diameter is also an important physiological measure of cognitive effort. Previous studies have suggested that changes in pupil size could be an index of how intensely the processing system is operating (Just and Carpenter, 1993; Laeng et al., 2012). The locus coeruleus (LC) is a brainstem nucleus that represents the primary source of norepinephrine (NE) in the brain (Kardon, 2005). The LC-NE system modulates the autonomic nervous system that influences pupil diameter through parasympathetic constriction and sympathetic dilation pathway (McDougal and Gamlin, 2008). The seminal study of Aston-Jones and Cohen (2005) demonstrated that pupil size is straightly related to fluctuations of the LC-NE activity. The same noradrenergic system has been implicated in arousal (Waterhouse, 2003; Aston-Jones, 2005; Aston-Jones et al., 2005) and mental effort (Bradley et al., 2008; Mathot, 2018). In fact, the LC widely projects throughout the PFC regions (Aston-Jones and Cohen, 2005) and seems to module neuronal network activity in cognitively demanding tasks (Robbins and Arnsten, 2009; Suttikus et al., 2021).

Although the LC-NE system presumably promotes the relationship between pupil dilation and mental effort, the mechanism in which the PFC activity areas influence the changes in pupil size is still not elucidated. Uncertainty about the mechanisms through the relationship between pupil size and PFC activity during cognitively demanding tasks emphasizes the need for further investigation into this matter. Multimodal studies have the potential to investigate how different areas of the nervous system are associated, revealing new levels of organization and

functioning. For example, a recent multimodal ET-fNIRS study demonstrated a functional relationship between the lateral PFC and the LC that may vary at different difficulty levels during the work memory task (Yeung et al., 2021). It is already known that pupil dilates with cognitive load in adults (Van Der Meer et al., 2010; Van Der Wel and van Steenbergen, 2018) and children (McGarrigle et al., 2017). Our data relates schoolchildren's pupil dilation during the MRT, suggesting a task-evoked increasing effort by spatial cognition demand. Further research should investigate the neural network specialization in visuospatial demand tasks across different age stages and neural development.

Interestingly, there are pieces of evidence that the brain activation of children can change after training. Soltanlou et al. (2018a) demonstrated that 2-week multiplication training with children reduced brain activation on the frontoparietal network. Furthermore, neuroimaging studies have suggested that children show decreased brain activation of the frontal cortex with increasing age (Houdé et al., 2010; Artemenko et al., 2018), which indicates reduced reliance on cognitive effort and attentional resources. Indeed, the cognitive effort is related to the degree of engagement with demanding tasks, and high engagement tends to enhance performance and attention (Kaplan and Berman, 2010). Considering the challenges to measuring the cognitive load (Paas et al., 2003), there is an increasing interest in the educational research field to apply a combination of measurements to obtain reliable and objective physiological data of cognitive effort in real-time (Kruger and Doherty, 2016).

The multimodal methodology has been used recently as a framework for the physiological measurement of cognitive load (Hosseini et al., 2017; İsbilir et al., 2019; Larmuseau et al., 2020; Vanneste et al., 2021; Yeung et al., 2021). Studies applying pupillometry-fMRI concurrently measure have shown that pupil dilation was temporally correlated with activation of adults' brain regions implicated in workload and cognitive control during working memory and decision-making tasks (Siegle et al., 2003; Satterthwaite et al., 2007). Similarly, we measured cortical activity and pupillary response concurrently using fNIRS and eye-tracking glasses to examine the neural systems linked to pupil dilation under mental rotation tasks. We also explored which PFC regions activate more during the visuospatial task. We believe that pupil dilations could be an index of mental effort differences indicating the mental effort linked to the allocation of cognitive resources on cognitively demanding tasks. We hypothesize that the change in pupil diameter could suggest how hard the student is trying to solve the problem, indicating the performance slightly. However, we did not apply different level difficulties in the MRT trials, which is a relevant limitation of our study. We found some short-channels signals significantly correlated to the task performance,

which indicate extracerebral perfusion (e.g., forehead skin vessels) may possibly be an interference in our hemodynamic signals. Although such short channels do not correspond to long-distance channels that correlated significantly with the task performance, we considered this limitation worthy of mention.

One point of remark is that two fNIRS statistical analyses were conducted. The first focused on the significance of the contrast between task and rest. The second had as concern whether the amplitude of activation was correlated with behavioral data. We could use the results of the first analysis to select the channels for the second one. However, we opted to consider all channels in the latter and to be conservative by applying Bonferroni correction for multiple comparisons. We justify this choice since the two analyses are not only complementary, but they are testing different null hypotheses. The second point of remark is that although both oxy- and deoxy-Hb changes are consequences of the same neurophysiological process (the neurohemodynamic coupling process), the signal-to-noise ratio and susceptibility to systemic artifacts are not the same and may also depend on brain location. Thus, the statistical findings are not always reciprocal since the effect sizes are different. Moreover, for the same reasons, the lack of statistical power could explain why the correlation with task performance was only found at the right dlPFC.

Conclusion

Our findings suggest that children with more mental effort during MRT performed better in the visuospatial and geometry problem-solving tests. The multimodal approach to monitoring students' mental effort can be of great interest in providing objective feedback on cognitive resource conditions and advancing our comprehension of the neural mechanisms that underlie cognitive effort. Hence, the ability to detect two distinct mental states of rest or activation of children in real-time during the MRT could eventually lead to an application for investigating visuospatial skills of young students using naturalistic educational paradigms with a multimodal approach by applying the fNIRS and Eye-Tracker devices.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the UFABC. Written informed consent to

participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

RS and JR designed the study, collected and analyzed the data, and wrote the manuscript. AA and CB wrote and reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding

JR and AA were supported by the São Paulo Research Foundation (FAPESP, Grants Nos. 2018/21934-5 and 2018/04654-9). AA was also supported by the São Paulo Research Foundation (FAPESP, Grant No. 2019/17907-5).

Acknowledgments

We are thankful to the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) – Finance Code 001 and Universidade Federal do ABC (UFABC) and this study was developed based on a doctoral research project at the Graduate Program in Neuroscience

and Cognition at the Universidade Federal do ABC.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnhum.2022.889806/full#supplementary-material>

References

- Ahern, S., and Beatty, J. (1979). Pupillary responses during information processing vary with Scholastic Aptitude Test scores. *Science* 205, 1289–1292. doi: 10.1126/science.472746
- Artemenko, C., Soltanlou, M., Ehli, A. C., Nuerk, H. C., and Dresler, T. (2018). The neural correlates of mental arithmetic in adolescents: a longitudinal fNIRS study. *Behav. Brain Funct.* 14, 1–13. doi: 10.1186/s12993-018-0137-8
- Aston-Jones, G. (2005). Brain structures and receptors involved in alertness. *Sleep medicine* 6, S3–S7.
- Aston-Jones, G., Gonzalez, M. M., and Doran, S. M. (2005). "Role of the locus coeruleus-norepinephrine system in arousal and circadian regulation of the sleep-waking cycle," in *Norepinephrine: Neurobiology and Therapeutics for the 21st Century*, eds G. A. Ordway, M. Schwartz, and A. Frazer (Cambridge, MA: Cambridge Univ. Press. in press).
- Aston-Jones, G., and Cohen, J. D. (2005). Adaptive gain and the role of the locus coeruleus-norepinephrine system in optimal performance. *J. Comp. Neurol.* 493, 99–110.
- Ayaz, H., Shewokis, P. A., Bunce, S., Izzetoglu, K., Willems, B., and Onaral, B. (2012a). Optical brain monitoring for operator training and mental workload assessment. *Neuroimage* 59, 36–47. doi: 10.1016/j.neuroimage.2011.06.023
- Ayaz, H., Shewokis, P. A., İzzetoglu, M., Çakır, M. P., and Onaral, B. (2012b). "Tangram solved? Prefrontal cortex activation analysis during geometric problem solving," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, (IEEE)*, 4724–4727. doi: 10.1109/EMBC.2012.6347022
- Ayaz, H., Willems, B., Bunce, B., Shewokis, P. A., Izzetoglu, K., Hah, S., et al. (2010). "Cognitive workload assessment of air traffic controllers using optical brain imaging sensors," in *Advances in Understanding Human Performance: Neuroergonomics, Human Factors Design, and Special Populations*, eds T. Marek, W. Karwowski, and V. Rice (Boca Raton: CRC Press Taylor & Francis Group), 21–31.
- Balardin, J. B., Zimeo Morais, G. A., Furucho, R. A., Trambaiolli, L., Vanzella, P., Biazoli, Jr., et al. (2017). Imaging brain function with functional near-infrared spectroscopy in unconstrained environments. *Front. Hum. Neurosci.* 11:258.
- Baltes, P. B., Staudinger, U. M., and Lindenberger, U. (1999). Lifespan psychology: theory and application to intellectual functioning. *Annu. Rev. Psychol.* 50, 471–507. doi: 10.1146/annurev.psych.50.1.471
- Barreto, C., de Albuquerque Bruneri, G., Brockington, G., Ayaz, H., and Sato, J. R. (2021). A new statistical approach for fNIRS hyperscanning to predict brain activity of preschoolers' using teacher's. *Front. Hum. Neurosci.* 15:181. doi: 10.3389/fnhum.2021.622146
- Bauer, R., Jost, L., Günther, B., and Jansen, P. (2021a). Pupillometry as a measure of cognitive load in mental rotation tasks with abstract and embodied figures. *Psycholog. Res.* 2021, 1–15. doi: 10.1007/s00426-021-01568-5
- Bauer, R., Jost, L., and Jansen, P. (2021b). The effect of mindfulness and stereotype threat in mental rotation: a pupillometry study. *J. Cogn. Psychol.* 33, 861–876.
- Beatty, J., and Kahneman, D. (1966). Pupillary changes in two memory tasks. *Psychon. Sci.* 5, 371–372. doi: 10.3758/BF03328444
- Biber, C., Tuna, A., and Korkmaz, S. (2013). The mistakes and the misconceptions of the eighth grade students on the subject of angles. *Eur. J. Sci. Mathemat. Educ.* 1, 50–59.
- Boersma, F., Wilton, K., Barham, R., and Muir, W. (1970). Effects of arithmetic problem difficulty on pupillary dilation in normals and educable retardates. *J. Exp. Child Psychol.* 9, 142–155. doi: 10.1016/0022-0965(70)90079-2

- Bolden, D., Barmby, P., Raine, S., and Gardner, M. (2015). How young children view mathematical representations: a study using eye-tracking technology. *Educ. Res.* 57, 59–79.
- Bradley, M. M., Miccoli, L., Escrig, M. A., and Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology* 45, 602–607.
- Bruce, C. D., and Hawes, Z. (2015). The role of 2D and 3D mental rotation in mathematics for young children: what is it? Why does it matter? And what can we do about it? *ZDM Math. Educ.* 47, 331–343. doi: 10.1007/s11858-014-0637-4
- Buckley, J., Canty, D., White, D., Seery, N., and Campbell, M. (2018). Spatial Working Memory and Neural Efficiency in Mental Rotations: an Insight from Pupillometry. *Eng. Design Graph. J.* 82:3.
- Çakır, M. P., Ayaz, H., İzzetoğlu, M., Shewokis, P. A., İzzetoğlu, K., and Onaral, B. (2011). Bridging brain and educational sciences: an optical brain imaging study of visuospatial reasoning. *Procedia-Soc. Behav. Sci.* 29, 300–309.
- Campbell, M. J., Toth, A. J., and Brady, N. (2018). Illuminating sex differences in mental rotation using pupillometry. *Biol. Psychol.* 138, 19–26. doi: 10.1016/j.biopsycho.2018.08.003
- Carnoy, M., Khavenson, T., Fonseca, I., Costa, L., and Marotta, L. (2015). Is Brazilian education improving? Evidence from PISA and SAEB. *Cadernos de Pesquisa* 45, 450–485.
- Causse, M., Chua, Z., Peysakhovich, V., Del Campo, N., and Matton, N. (2017). Mental workload and neural efficiency quantified in the prefrontal cortex using fNIRS. *Sci. Rep.* 7, 1–15. doi: 10.1038/s41598-017-05378-x
- Chatham, C. H., Frank, M. J., and Munakata, Y. (2009). Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control. *Proc. Natl. Acad. Sci. U.S.A.* 106, 5529–5533. doi: 10.1073/pnas.0810002106
- Chevalier, N. (2018). Willing to think hard? The subjective value of cognitive effort in children. *Child Dev.* 89, 1283–1295. doi: 10.1111/cdev.12805
- Cohen, M. S., Kosslyn, S. M., Breiter, H. C., DiGirolamo, G. J., Thompson, W. L., Anderson, A. K., et al. (1996). Changes in cortical activity during mental rotation: a mapping study using functional MRI. *Brain* 119, 89–100.
- Curtin, A., and Ayaz, H. (2018). The age of neuroergonomics: towards ubiquitous and continuous measurement of brain function with fNIRS. *Japan. Psychol. Res.* 60, 374–386.
- Delgado, A. R., and Prieto, G. (2004). Cognitive mediators and sex-related differences in mathematics. *Intelligence* 32, 25–32. doi: 10.1016/j.cognition.2018.10.005
- Eckstein, M. K., Guerra-Carrillo, B., Singley, A. T. M., and Bunge, S. A. (2017). Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development? *Developmental cognitive neuroscience* 25, 69–91. doi: 10.1016/j.dcn.2016.11.001
- Epelboim, J., and Suppes, P. (2001). A model of eye movements and visual working memory during problem solving in geometry. *Vision Res.* 41, 1561–1574. doi: 10.1016/S0042-6989(00)00256-x
- Ferrari, M., and Quaresima, V. (2012). A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *Neuroimage* 63, 921–935. doi: 10.1016/j.neuroimage.2012.03.049
- Fishburn, F. A., Norr, M. E., Medvedev, A. V., and Vaidya, C. J. (2014). Sensitivity of fNIRS to cognitive state and load. *Front. Hum. Neurosci.* 8, 1–11. doi: 10.3389/fnhum.2014.00076
- Gao, Y., Barreto, A., and Zhai, J. (2007). “Digital filtering of pupil diameter variations for the detection of stress in computer users,” in *Proc 11th world multi-conference on systemics, cybernetics and informatics*, (Orlando).
- Giofrè, D., Mammarella, I. C., Ronconi, L., and Cornoldi, C. (2013). Visuospatial working memory in intuitive geometry, and in academic achievement in geometry. *Learn. Individ. Diff.* 23, 114–122.
- Granholm, E., Asarnow, R. F., Sarkin, A. J., and Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology* 33, 457–461. doi: 10.1111/j.1469-8986.1996.tb01071.x
- Guillot, A., Hoyek, N., and Collet, C. (2012). Mental Rotation and Functional Learning. *Encyclopedia of the Sciences of Learning* 2012, 2222–2223. doi: 10.1007/978-1-4419-1428-6_493
- Harris, I. M., and Miniussi, C. (2003). Parietal lobe contribution to mental rotation demonstrated with rTMS. *J. Cogn. Neurosci.* 15, 315–323. doi: 10.1162/089892903321593054
- Hertanti, A., Retnawati, H., and Wutsqa, D. U. (2019). The role of spatial experience in mental rotation. *J. Phys.* 1320:012043.
- Hess, E. H., and Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science* 143, 1190–1192. doi: 10.1126/science.143.3611.1190
- Hess, T. M., and Ennis, G. E. (2012). Age differences in the effort and costs associated with cognitive activity. *J. Gerontol. Series* 67, 447–455.
- Hosseini, S. M., Bruno, J. L., Baker, J. M., Gundran, A., Harbott, L. K., Gerdes, J. C., et al. (2017). Neural, physiological, and behavioral correlates of visuomotor cognitive load. *Sci. Rep.* 7, 1–9. doi: 10.1038/s41598-017-07897-z
- Houdé, O., Rossi, S., Lubin, A., and Joliot, M. (2010). Mapping numerical processing, reading, and executive functions in the developing brain: an fMRI meta-analysis of 52 studies including 842 children. *Dev. Sci.* 13, 876–885. doi: 10.1111/j.1467-7687.2009.00938.x
- Hugdahl, K., Thomsen, T., and Ersland, L. (2006). Sex differences in visuo-spatial processing: an fMRI study of mental rotation. *Neuropsychologia* 44, 1575–1583. doi: 10.1016/j.neuropsychologia.2006.01.026
- Huppert, T. J. (2016). Commentary on the statistical properties of noise and its implication on general linear models in functional near-infrared spectroscopy. *Neurophotonics* 3:010401. doi: 10.1117/1.NPh.3.1.010401
- Hwang, W. Y., Su, J. H., Huang, Y. M., and Dong, J. J. (2009). A study of multi-representation of geometry problem solving with virtual manipulatives and whiteboard system. *J. Educ. Technol. Soc.* 12, 229–247.
- İşbilir, E., Çakır, M. P., Acartürk, C., and Tekerek, A. S. (2019). Towards a multimodal model of cognitive workload through synchronous optical brain imaging and eye tracking measures. *Front. Hum. Neurosci.* 2019:375. doi: 10.3389/fnhum.2019.00375
- Jainta, S., and Baccino, T. (2010). Analyzing the pupil response due to increased cognitive demand: an independent component analysis study. *Internat. J. Psychophysiol.* 77, 1–7. doi: 10.1016/j.ijpsycho.2010.03.008
- Johnson, E. L., Miller Singley, A. T., Peckham, A. D., Johnson, S. L., and Bunge, S. A. (2014). Task-evoked pupillometry provides a window into the development of short-term memory capacity. *Front. Psychol.* 5:218. doi: 10.3389/fpsyg.2014.00218
- Jordan, K., Heinze, H. J., Lutz, K., Kanowski, M., and Jäncke, L. (2001). Cortical activations during the mental rotation of different visual objects. *Neuroimage* 13, 143–152. doi: 10.1006/nimg.2000.0677
- Jurcak, V., Tsuzuki, D., and Dan, I. (2007). 10/20, 10/10, and 10/5 systems revisited: their validity as relative head-surface-based positioning systems. *Neuroimage* 34, 1600–1611. doi: 10.1016/j.neuroimage.2006.09.024
- Just, M. A., and Carpenter, P. A. (1993). The intensity dimension of thought: pupillometric indices of sentence processing. *Can. J. Exp. Psychol.* 47, 310–339. doi: 10.1037/h0078820
- Kaplan, S., and Berman, M. G. (2010). Directed Attention as a Common Resource for Executive Functioning and Self-Regulation. *Perspect. Psychol. Sci.* 5, 43–57. doi: 10.1177/1745691609356784
- Karatekin, C. (2004). Development of attentional allocation in the dual task paradigm. *Int. J. Psychophysiol.* 52, 7–21. doi: 10.1016/j.ijpsycho.2003.12.002
- Karatekin, C., Marcus, D. J., and Couperus, J. W. (2007). Regulation of cognitive resources during sustained attention and working memory in 10-year-olds and adults. *Psychophysiology* 44, 128–144. doi: 10.1111/j.1469-8986.2006.00477.x
- Kardon, R. H. (2005). “Anatomy and physiology of the autonomic nervous system,” in *Wash and Hoyt's Clinical Neuro-Ophthalmology*, 6th Edn, eds N. R. Miller, N. J. Newman, V. Biousse, and J. B. Kerrison (Philadelphia: Lippincott Williams & Wilkins), 649–714.
- Khine, M. S. (2016). “Spatial cognition: Key to STEM success,” in *Visual-spatial Ability in STEM Education: Transforming Research into Practice*, (New York, NY: Springer International Publishing), 3–8. doi: 10.1007/978-3-319-44385-0_1
- Klingner, J., Tversky, B., and Hanrahan, P. (2011). Effects of visual and verbal presentation on cognitive load in vigilance, memory, and arithmetic tasks. *Psychophysiology* 48, 323–332. doi: 10.1111/j.1469-8986.2010.01069.x
- Koessler, L., Maillard, L., Benhadid, A., Vignal, J. P., Felblinger, J., Vespignani, H., et al. (2009). Automated cortical projection of EEG sensors: anatomical correlation via the international 10–10 system. *Neuroimage* 46, 64–72. doi: 10.1016/j.neuroimage.2009.02.006

- Kruger, J. L., and Doherty, S. (2016). Measuring cognitive load in the presence of educational video: Towards a multimodal methodology. *Austral. J. Educ. Technol.* 32:6.
- Kyttälä, M., and Lehto, J. E. (2008). Some factors underlying mathematical performance: The role of visuospatial working memory and non-verbal intelligence. *Eur. J. Psychol. Educ.* 23, 77–94. doi: 10.1007/BF03173141
- Laeng, B., Sirois, S., and Gredebäck, G. (2012). Pupillometry: a window to the preconscious? *Perspect. Psychol. Sci.* 7, 18–27. doi: 10.1177/1745691611427305
- Larmuseau, C., Cornelis, J., Lancieri, L., Desmet, P., and Depaepe, F. (2020). Multimodal learning analytics to investigate cognitive load during online problem solving. *Br. J. Educ. Technol.* 51, 1548–1562.
- Lohman, D. F. (1994). "Spatial ability," in *Encyclopedia of intelligence*, ed. R. J. Sternberg (New York, NY: Macmillan), 1000–1007.
- Lum, J. A., Youssef, G. J., and Clark, G. M. (2017). Using pupillometry to investigate sentence comprehension in children with and without specific language impairment. *J. Speech, Lang. Hear. Res.* 60, 1648–1660. doi: 10.1044/2017_JSLHR-L-16-0158
- Mathot, S. (2018). Pupillometry: psychology, physiology, and function. *J. Cogn.* 1:1.
- McCabe, D. P., Roediger, H. L. III, McDaniel, M. A., Balota, D. A., and Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: evidence for a common executive attention construct. *Neuropsychology* 24:222. doi: 10.1037/a0017619
- McDougal, D. H., and Gamlin, P. D. R. (2008). "Pupillary control pathways," in *The Senses: A Comprehensive Reference*, Vol. 1, eds R. H. Masland and T. Albright (San Diego, CA: Academic Press), 521–536. doi: 10.1016/B978-012370880-9.00282-6
- McGarrigle, R., Dawes, P., Stewart, A. J., Kuchinsky, S. E., and Munro, K. J. (2017). Pupillometry reveals changes in physiological arousal during a sustained listening task. *Psychophysiology* 54, 193–203. doi: 10.1111/psyp.12772
- Miller, E. K., and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202.
- Mücke, M., Andrä, C., Gerber, M., Pühse, U., and Ludyga, S. (2018). Moderate-to-vigorous physical activity, executive functions and prefrontal brain oxygenation in children: a functional near-infrared spectroscopy study. *J. Sports Sci.* 36, 630–636. doi: 10.1080/02640414.2017.1326619
- Orde, B. J. (1996). *A correlational analysis of drawing ability and spatial ability*. Laramie, WY: University of Wyoming.
- Paas, F. G., Tuovinen, J. E., Tabbers, H., and Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educ. Psychol.* 38, 63–71. doi: 10.1207/S15326985EP3801_8
- Pak, R. (2001). "A further examination of the influence of spatial abilities on computer task performance in younger and older adults," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 45, (Los Angeles, CA: SAGE Publications), 1551–1555.
- Prat, S. C., Keller, A. T., and Just, M. (2007). Individual differences in sentence comprehension: a functional magnetic resonance imaging investigation of syntactic and lexical processing demands. *J. Cogn. Neurosci.* 2007:66155197. doi: 10.1184/R1/66155197.V1
- Robbins, T. W., and Arnsten, A. F. T. (2009). The neuropsychopharmacology of fronto-executive function: monoaminergic modulation. *Annu. Rev. Neurosci.* 32, 267–287. doi: 10.1146/annurev.neuro.051508.135535
- Roberts, J. E., and Bell, M. A. (2000). Sex differences on a mental rotation task: variations in electroencephalogram hemispheric activation between children and college students. *Dev. Neuropsychol.* 17, 199–223. doi: 10.1207/S15326942DN1702_04
- Rojas-Libano, D., Wainstein, G., Carrasco, X., Aboitiz, F., Crossley, N., and Ossandón, T. (2019). A pupil size, eye-tracking and neuropsychological dataset from ADHD children during a cognitive task. *Scientific Data* 6, 1–6. doi: 10.1038/s41597-019-0037-2
- Rozado, D., and Andreas, D. (2015). Combining EEG with pupillometry to improve cognitive workload detection. *Computer* 48, 18–25.
- SAEB (2011). *Geometry Problems from SAEB 2011 – INEP*. Available online at: <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/saeb> (accessed February 29, 2020).
- Satterthwaite, T. D., Green, L., Myerson, J., Parker, J., Ramaratnam, M., and Buckner, R. L. (2007). Dissociable but inter-related systems of cognitive control and reward during decision making: evidence from pupillometry and event-related fMRI. *NeuroImage* 37, 1017–1031. doi: 10.1016/j.neuroimage.2007.04.066
- Shaw, E. P., Rietschel, J. C., Hendershot, B. D., Pruziner, A. L., Miller, M. W., Hatfield, B. D., et al. (2018). Measurement of attentional reserve and mental effort for cognitive workload assessment under various task demands during dual-task walking. *Biolog. Psychol.* 134, 39–51.
- Shepard, R. N., and Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science* 171, 701–703. doi: 10.1126/science.171.3972.701
- Shepard, S., and Metzler, D. (1988). Mental Rotation: effects of Dimensionality of Objects and Type of Task. *J. Exp. Psychol. Hum. Percept. Perform.* 14, 3–11. doi: 10.1037/0096-1523.14.1.3
- Shin, J., Von Lühmann, A., Kim, D. W., Mehnert, J., Hwang, H. J., and Müller, K. R. (2018). Simultaneous acquisition of EEG and NIRS during cognitive tasks for an open access dataset. *Scient. Data* 5, 1–16. doi: 10.1038/sdata.2018.3
- Siegle, G. J., Steinhauer, S. R., Stenger, V. A., Konecky, R., and Carter, C. S. (2003). Use of concurrent pupil dilation assessment to inform interpretation and analysis of fMRI data. *NeuroImage* 20, 114–124. doi: 10.1016/S1053-8119(03)00298-2
- Sladky, R., Stepniczka, I., Boland, E., Tik, M., Lamm, C., Hoffmann, A., et al. (2016). Neurobiological differences in mental rotation and instrument interpretation in airline pilots. *Sci. Rep.* 6, 1–6. doi: 10.1038/srep28104
- Soltanlou, M., Artemenko, C., Ehlis, A. C., Huber, S., Fallgatter, A. J., Dresler, T., et al. (2018a). Reduction but no shift in brain activation after arithmetic learning in children: a simultaneous fNIRS-EEG study. *Sci. Rep.* 8, 1–15. doi: 10.1038/s41598-018-20007-x
- Soltanlou, M., Sitnikova, M. A., Nuerk, H. C., and Dresler, T. (2018b). Applications of functional near-infrared spectroscopy (fNIRS) in studying cognitive development: the case of mathematics and language. *Front. Psychol.* 9:277. doi: 10.3389/fpsyg.2018.00277
- Suttkus, S., Schumann, A., de la Cruz, F., and Bär, K. J. (2021). Working memory in schizophrenia: The role of the locus coeruleus and its relation to functional brain networks. *Brain Behav.* 11:e02130. doi: 10.1002/brb3.2130
- Sweller, J., and Chandler, P. (1991). Evidence for cognitive load theory. *Cogn. Instruct.* 8, 351–362.
- Sweller, J., van Merriënboer, J. J. G., and Paas, F. G. W. C. (1998). Cognitive Architecture and Instructional Design. *Educ. Psychol. Rev.* 10, 251–296. doi: 10.1023/A:1022193728205
- Tolar, T. D., Lederberg, A. R., and Fletcher, J. M. (2009). A structural model of algebra achievement: computational fluency and spatial visualisation as mediators of the effect of working memory on algebra achievement. *Educ. Psychol.* 29, 239–266. doi: 10.1080/01443410802708903
- Toth, A. J., and Campbell, M. J. (2019). Investigating sex differences, cognitive effort, strategy, and performance on a computerised version of the mental rotations test via eye tracking. *Sci. Rep.* 9, 1–11. doi: 10.1038/s41598-019-56041-6
- Tzuriel, D., and Egozi, G. (2010). Gender differences in spatial ability of young children: the effects of training and processing strategies. *Child Dev.* 81, 1417–1430. doi: 10.1111/j.1467-8624.2010.01482.x
- Van Der Meer, E., Beyer, R., Horn, J., Foth, M., Bornemann, B., Ries, J., et al. (2010). Resource allocation and fluid intelligence: Insights from pupillometry. *Psychophysiology* 47, 158–169. doi: 10.1111/j.1469-8986.2009.00884.x
- Van Der Wel, P., and van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: a review. *Psychon. Bull. Rev.* 25, 2005–2015. doi: 10.3758/s13423-018-1432-y
- Vanneste, P., Raes, A., Morton, J., Bombeke, K., Van Acker, B. B., Larmuseau, C., et al. (2021). Towards measuring cognitive load through multimodal physiological data. *Cogn. Technol. Work* 23, 567–585.
- Vassena, E., Gerrits, R., Demanet, J., Verguts, T., and Siugzdaitė, R. (2019). Anticipation of a mentally effortful task recruits Dorsolateral Prefrontal Cortex: an fNIRS validation study. *Neuropsychologia* 123, 106–115. doi: 10.1016/j.neuropsychologia.2018.04.033
- Vassena, E., Silveti, M., Boehler, C. N., Achten, E., Fias, W., and Verguts, T. (2014). Overlapping neural systems represent cognitive effort and reward anticipation. *PLoS One* 9:e91008. doi: 10.1371/journal.pone.0091008
- Wai, J., Lubinski, D., and Benbow, C. P. (2009). Spatial ability for STEM domains: aligning over 50 years of cumulative psychological knowledge solidifies its importance. *J. Educ. Psychol.* 101, 817–835. doi: 10.1037/a0016127
- Waterhouse, B. D. (2003). The locus coeruleus-noradrenergic system: modulation of behavioral state and state-dependent cognitive processes. *Brain Res. Brain Res. Rev.* 42, 33–84. doi: 10.1016/s0165-0173(03)00143-7
- Webb, R. M., Lubinski, D., and Benbow, C. P. (2007). Spatial ability: a neglected dimension in talent searches for intellectually precocious youth. *J. Educ. Psychol.* 99, 397–420. doi: 10.1037/0022-0663.99.2.397
- Wei, W., Yuan, H., Chen, C., and Zhou, X. (2012). Cognitive correlates of performance in advanced mathematics. *Br. J. Educ. Psychol.* 82, 157–181. doi: 10.1111/j.2044-8279.2011.02049.x

- Westbrook, A., and Braver, T. S. (2015). Cognitive effort: a neuroeconomic approach. *Cogn. Affect. Behav. Neurosci.* 15, 395–415.
- Westbrook, A., Kester, D., and Braver, T. S. (2013). What is the subjective cost of cognitive effort? load, trait, and aging effects revealed by economic preference. *PLoS One* 8:e68210. doi: 10.1371/journal.pone.0068210
- Wu, D. D., Yang, J. F., Xie, S., Luo, J. T., Chang, C. Q., and Li, H. (2020). An fNIRS examination of the neural correlates of mental rotation in preschoolers. *Hum. Behav. Brain* 1, 37–42.
- Xia, M., Wang, J., and He, Y. (2013). BrainNet viewer: a network visualization tool for human brain connectomics. *PLoS One* 8:e68910. doi: 10.1371/journal.pone.0068910
- Yeung, M. K., Lee, T. L., Han, Y. M., and Chan, A. S. (2021). Prefrontal activation and pupil dilation during n-back task performance: a combined fNIRS and pupillometry study. *Neuropsychologia* 159:107954. doi: 10.1016/j.neuropsychologia.2021.107954
- Zacks, J. M. (2008). Neuroimaging studies of mental rotation: a meta-analysis and review. *J. Cogn. Neurosci.* 20, 1–19. doi: 10.1162/jocn.2008.20013
- Zacks, J. M., Gilliam, F., and Ojemann, J. G. (2003). Selective disturbance of mental rotation by cortical stimulation. *Neuropsychologia* 41, 1659–1667. doi: 10.1016/S0028-3932(03)00099-X



OPEN ACCESS

EDITED BY
Takahiko Koike,
National Institute for Physiological
Sciences (NIPS), Japan

REVIEWED BY
Xian Zhang,
Yale University, United States
Dongchuan Yu,
Southeast University, China

*CORRESPONDENCE
Yi Liu
liuy930@nenu.edu.cn

†These authors have contributed
equally to this work

SPECIALTY SECTION
This article was submitted to
Interacting Minds and Brains,
a section of the journal
Frontiers in Human Neuroscience

RECEIVED 22 July 2022
ACCEPTED 19 October 2022
PUBLISHED 09 November 2022

CITATION
Liu Y, Li J, Wang Q and Li Y (2022) The
specificity, situational modulations,
and behavioral correlates
of parent-child neural synchrony.
Front. Hum. Neurosci. 16:1000826.
doi: 10.3389/fnhum.2022.1000826

COPYRIGHT
© 2022 Liu, Li, Wang and Li. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

The specificity, situational modulations, and behavioral correlates of parent-child neural synchrony

Yi Liu^{*†}, Jiaxin Li[†], Qi Wang and Yarong Li

School of Psychology, Northeast Normal University, Changchun, China

In recent years, aiming to uncover the neural mechanism of parent-child interaction and link it to the children's social development, a newly developed index, namely, parent-child inter-brain neural synchronization (INS) has attracted growing interest. Existing studies have mainly focused on three aspects of the INS; these are the specificity of the INS (i.e., stronger INS for parent-child dyads than stranger-child dyads), the situational modulations of the INS (i.e., how the valence of the situation or the types of interaction modulate INS), and the associations between the INS and the state-like behavioral tendencies or trait-like individual features of the parents and children. This review summarizes the existing findings in line with these three topics and provides preliminary suggestions to promote parent-child INS. In the meanwhile, the inconsistent findings and unstudied questions were discussed, opening new avenues for future studies.

KEYWORDS

parent-child, neural synchrony, social interaction, social development, fNIRS

Introduction

In the past three decades, parent-child synchrony has been used to describe the temporal correlations of the behavior (e.g., gaze), affective states (e.g., positive affect), and biological rhythms (e.g., heart rhythms) between the parent and the child (see [Feldman, 2007a](#) for a review). The parent-child synchrony has also been shown to contribute to children's cognitive, social, and emotional growth ([Feldman, 2007a,b,c](#)). More recently, with the fast development of modern neuroimaging techniques (e.g., functional near-infrared spectroscopy, fNIRS and magnetoencephalography, MEG), it is possible to simultaneously record the neural activities of two or more individuals, i.e., hyperscanning, allowing the identification of synchronous features at the neural level when two or more persons are engaged in real-time reciprocal interactions. As a result, a growing number of researchers shift their attention from the single-brain activities toward the study of inter-brain neural synchronization (INS) to uncover the

neural mechanisms of social interaction. The INS describes similar patterns of the brain activities across participants at the same time, and it is associated with behavioral synchronization and successful communication (Cui et al., 2012; Jiang et al., 2012). Under the interactive context, the INS reflects the emergent dynamics such as mutual understanding or shared psychological state between interactors, while under non-interactive situations (e.g., watching videos) the INS reflects the shared representations of the dyads during social stimuli processing (Cui et al., 2012; Hasson et al., 2012; Redcay and Schilbach, 2019). In the field of developmental psychology, researchers employ the technique to uncover the neurobiological underpinnings of parent-child interaction. More specifically, researchers use the INS indices to describe the neural similarity of the parent-child dyads during social interaction or social stimuli processing and attempt to associate the parent-child INS with the interactive features of the interaction or the individual's features of the dyads.

Existing studies have investigated the parent-child INS using fNIRS (Reindl et al., 2018; Azhari et al., 2019, 2020, 2021; Miller et al., 2019; Nguyen et al., 2020a,b, 2021b; Quiñones-Camacho et al., 2020; Wang et al., 2020; Zhao et al., 2021), MEG (Levy et al., 2017), Electroencephalograph (EEG) (Endevelt-Shapira et al., 2021; Deng et al., 2022a,b; Zivan et al., 2022), and Optical topography (OT) (Bembich et al., 2022) to record the brain activities of parent-child dyads simultaneously. The simplest way to calculate the INS is to compute the cross-correlation of the time-series signals between each dyad (Azhari et al., 2020, 2021; Quiñones-Camacho et al., 2020; Zivan et al., 2022). The most used technique is the wavelet transform coherence (WTC) analysis, which estimates a coherence coefficient between two time series of each dyad as a function of frequency and time reflecting both time-related and frequency-related properties of the two time-series (Reindl et al., 2018, 2022; Miller et al., 2019; Nguyen et al., 2020a,b, 2021b; Wang et al., 2020; Kruppa et al., 2021; Zhao et al., 2021; Bembich et al., 2022). Another technique (yet seldomly) used to calculate INS is the dynamic time warping (DTW) time-series analysis, which arranges all sequence points to optimize the alignment of two sequences (Azhari et al., 2019). For MEG or EEG indices, the INS can be calculated as the inter-brain weighted phase lag index (wPLI) reflecting the phase coupling of inter-brain activities from the dyad (Levy et al., 2017; Endevelt-Shapira et al., 2021). In addition, the phase-locking-value (PLV) index measuring whether the EEG signals from the two interacting individuals are phase locked across time was also used as the INS index (Deng et al., 2022a,b). Further, the INS can be defined as time-aligned or time-lagged neural synchronization. For time-aligned INS, the synchronization is calculated using the temporally aligned brain activities of two individuals. Whereas for the time-lagged INS, the brain activity of one individual temporally lags behind that of the other individual, which reflects interpersonal predictive coding and delayed processing during social interaction (Jiang et al., 2021).

To our knowledge, only one study used time-lagged WTC as the index of parent-child INS (Zhao et al., 2021), while the remaining studies used time-aligned parent-child INS indices (Levy et al., 2017; Reindl et al., 2018; Azhari et al., 2019, 2020, 2021; Miller et al., 2019; Nguyen et al., 2020a,b, 2021b; Quiñones-Camacho et al., 2020; Wang et al., 2020).

By using these INS indices, existing research focused on three main aspects of the parent-child INS as follows: (i) Specificity of the parent-child INS. This line of research compares the parent-child INS and stranger-child INS in either interactive or non-interactive situations and points out how long-term parent-child attachment shapes the interbrain function between parents and the child. (ii) Modulations of situational factors on the parent-child INS. This line of investigation compares the parent-child INS in different situations (e.g., cooperation or competition) to investigate how the parent-child INS is modulated by the ongoing social cognitive processes. Last but not least, (iii) the behavioral correlates of the parent-child INS. This line of research associates the parent-child INS with the state-like behavioral tendencies or trait-like features of the parent and the child, with the aim to formulate interpretations and identify implications concerning the parent-child INS. The current work will review the parent-infant/child/adolescent hyperscanning research published before October 2022 based on these three lines of investigation (see Table 1). Moreover, we will highlight the progress achieved by these studies and most importantly, we will bring to light critical limitations that can nevertheless be addressed in future studies.

Specificity of the parent-child inter-brain neural synchronization

The specificity of the parent-child INS, namely, the stronger INS of parent-child dyads than stranger-child dyads, was found in both non-interactive (e.g., watching videos without communication) and interactive situations (e.g., cooperative problem solving).

In non-interactive situations, parent-child dyads passively watched videos together without explicit cognitive tasks. Since the contents of the videos were consistent across all participants, the INS of parent-child dyads was compared with that of randomly paired stranger-child dyads (i.e., in data analysis, each child was randomly paired with a parent of another child to calculate the stranger-child INS as control). We regard the stronger parent-child INS (compared to INS of randomly paired stranger-child) as the specificity of the parent-child INS. For example, Azhari et al. used fNIRS to record the INS of mother-child dyads (Azhari et al., 2019) and father-child dyads (Azhari et al., 2021) while watching animated video clips (including scenes of positive family interactions and emotionally arousing family conflict). Compared with randomly paired stranger-child

TABLE 1 Summary of the literatures about parent-child neural synchrony.

References	Subjects	Techniques	Task	Variables manipulation	INS calculation	Brain regions	Results		
							Specificity	Situational modulations	Behavioral correlates
Azhari et al., 2019	31 Mother-child dyads; Children's gender: 18 boys and 13 girls; Mother's age: 34.9 ± 4.16 years; Children's age: 3.47 ± 0.51 years	fNIRS	Non-interaction (video watching)	Mother-child dyads vs. randomly paired stranger-child dyads	Time aligned; DTW	Left inferior frontal gyrus; frontal eye field; dorsolateral PFC	Mother-child dyads > randomly paired dyads (uncorrected)	Video positivity (null)	Mother-child dyads: ● Parenting stress (-)
Azhari et al., 2020	31 Mother-child dyads; Children's gender: 18 boys and 13 girls; Mother's age: 34.9 ± 4.16 years; Children's age: 3.47 ± 0.51 years	fNIRS	Non-interaction (video watching)	Mother-child dyads vs. randomly paired stranger-child dyads	Time aligned; Cross-correlation	Medial PFC	Mother-child dyads = randomly paired stranger-child dyads		Mother-child dyads: ● Maternal attachment anxiety (-)
Azhari et al., 2021	29 Father-child dyads; Children's gender: 18 boys and 11 girls; Father's age: 38.1 ± 3.67 years; Children's age: 3.52 ± 0.44 years	fNIRS	Non-interaction (video watching)	Positive scene vs. conflict scene; Father-child dyads vs. randomly paired stranger-child dyads	Time aligned; Cross correlation	Medial left PFC	Conflict scene: father-child dyads > randomly paired dyads	Conflict scene: father-child dyads > randomly paired dyads; Positive scenes: father-child dyads = randomly paired dyads	Father-child dyads: ● Father's age (-)
Bembich et al., 2022	16 Mother-infant dyads; Children's gender: 8 boys and 8 girls; Mother's age: 37.84 ± 3.09 years; infants' age: 2 days	Optical topography	Non-interaction (Watching painful stimulation on infants)	Baseline phase vs. Disinfection phase vs. Painful stimulation	Time aligned; WTC	Mother's parietal cortex; infants' motor/somato sensory cortex		Significant INS during painful stimulation	
Deng et al., 2022a	12 Parent-LSA dyads (5 mothers and 7 fathers): Children's gender: 5 boys and 7 girls; Mother's age: 40.20 ± 3.49 years; Father's age: 47.43 ± 4.54 years; Children's age: 11.75 ± 1.48 years; 13 Parent-HSA dyads (7 mothers and 6 fathers): Children's gender: 9 boys and 4 girls; Mother's age: 42.49 ± 5.78 years; Father's age: 42.67 ± 3.50 years; Children's age: 12.08 ± 1.12 years	EEG	Non-interaction (Watching emotional pictures)	HAS vs. LSA; Positive vs. negative vs. neutral	Time aligned; PLV	Fz, Pz, Cz		Positive: HSA > LSA; Negative: LSA > HSA; LSA: Negative > Neutral; HSA: Positive > Negative	

(Continued)

TABLE 1 (Continued)

References	Subjects	Techniques	Task	Variables manipulation	INS calculation	Brain regions	Results		
							Specificity	Situational modulations	Behavioral correlates
Deng et al., 2022b	15 LFC parent-adolescent dyads (6 mothers and 9 fathers): Children's gender: 11 boys and 4 girls; Mother's age: 43.56 ± 5.43 years; Father's age: 44.22 ± 4.63 years; Children's age: 12.00 ± 1.25 years; 14 HFC parent-adolescent dyads (7 mothers and 7 fathers): Children's gender: 11 boys and 3 girls; Mother's age: 42.57 ± 1.90 years; Father's age: 42.76 ± 3.21 years; Children's age: 12.36 ± 1.08 years	EEG	Non-interaction (video watching)	Positive vs. negative vs. neutral; LFC vs. HFC	Time aligned; PLV	Fz, Pz, Cz		Positive: HFC > LFC; HFC: Positive > Negative	
Endevelt-Shapira et al., 2021	Face-to-face: 62 Mother-infant dyads; Children's gender: 34 boys and 28 girls; Mothers' age: 33.3 ± 4.0 years; Children's age: 7.0 ± 1.49 months; Back-to-back: 39 Mother-infant dyads and 51 stranger-infant dyads; Mother-infant dyads: Children's gender: 23 boys and 16 girls; Mother's age: 32.7 ± 4.3 years; Children's age: 6.8 ± 1.1 months; Stranger-infant dyads: Children's gender: 26 boys and 25 girls; Stranger age: 34–35 years; Children's age: 6.8 ± 1.3 months	EEG	Interaction (free interaction)	Mother-child dyads vs. stranger-child dyads; face-to-face interaction vs. back-to-back interaction; With vs. without maternal body odor	Time aligned; wPLI	Central, parietal, and temporal areas	Face-to-face: mother-child dyads > stranger-child dyads; With maternal body odor: Mother-child = stranger-child; Stranger-child dyads: with > without maternal body odor	Mother-child dyads: Face-to-face > back-to-back	Stranger-child dyads with maternal body odor: • Child's visual attention (+); • Child's positive arousal (+); • Child's safety and engagement (null)
Kruppa et al., 2021	Control sample (TD group): 41 Parent-child dyads (40 mother and 1 father) and 32 stranger-child dyads; Children's gender: 18 boys and 23 girls; Parents' age: 44.98 ± 5.14 years; Strangers' age: 23.40 ± 3.72 years; Children's age: 12.66 ± 2.79 years; Matched sample: 18 Parent-ASD child dyads, 18 Parent-TD child TD dyads and 32 stranger-child dyads; Parent-ASD dyads: Children's gender: 18 boys; Parents' age: 46.49 ± 5.57 years; Strangers' age: 23.40 ± 3.72 years; Children's age: 13.54 ± 2.96 years; Parent-TD dyads: Children's gender: 18 boys; Parents' age: 47.15 ± 4.68 years; Strangers' age: 23.53 ± 4.40 years; Children's age: 13.53 ± 2.99 years	fNIRS	Interaction (button press task)	Parent-child dyads vs. Stranger-child dyads vs. randomly paired stranger-child dyads; Cooperation vs. Competition; ASD vs. TD	Time aligned; WTC	DLPFC, FPC	Control sample: Parent-child dyads > Stranger-child dyads (marginally); Parent-child dyads > randomly paired stranger-child dyads; Matched sample: Parent-ASD child dyads = randomly paired stranger-child dyads;	Control sample: Competition > Cooperation	Control sample: • Child's age (+);

(Continued)

TABLE 1 (Continued)

References	Subjects	Techniques	Task	Variables manipulation	INS calculation	Brain regions	Results		
							Specificity	Situational modulations	Behavioral correlates
Levy et al., 2017	Mother-child dyads Children's gender: not reported Mothers' age: 41.58 ± 4.69 years; Children's age: 11.67 ± 0.89 years	MEG	Non-interaction (video watching)	Own videos vs. other's videos; Behavioral synchrony episodes vs. non-synchrony episodes	Time aligned; wPLI (gamma-band phase coupling); correlation coefficients (gamma-band power coupling)	STS		Own videos: phase coupling: behavioral synchronized episodes > behavioral non-synchronized episodes; power coupling: behavioral synchronized episodes = behavioral non-synchronized episodes	
Miller et al., 2019	28 Mother-child dyads; Children's gender: 13 boys and 15 girls; Mother's age: 45.93 ± 3.76 years; Children's age: 11.27 ± 1.27 years	fNIRS	Interaction (button press task)	Cooperation vs. independent task	Time aligned; WTC	Right dorsolateral and frontopolar PFC		Cooperation > independent task	Cooperation: • Children's anxious attachment (null) • Children's avoidant attachment (null)
Nguyen et al., 2020a	40 Mother-child dyads; Children's gender: 20 boys and 20 girls; Mother's age: 36.37 ± 4.51 years; Children's age: 5.07 ± 0.04 years	fNIRS	Interaction (free conversation)		Time aligned; WTC	TPJ; dorsolateral PFC			• Amounts of turn-taking (+) • High turn-taking conversation: time course (+)
Nguyen et al., 2020b	42 Mother-child dyads; Children's gender: 19 boys and 23 girls; Mother's age: 36.42 ± 4.81 years; Children's age: 5.08 ± 0.04 years	fNIRS	Interaction (puzzle-solving task)	Cooperation vs. independent task	Time aligned; WTC	TPJ; dorsolateral PFC		Cooperation > independent and rest	Cooperation: • Behavioral reciprocity (+) • Cooperative performance (+) • Maternal stress (-) • Child agency (+) • Children's temperamental negative affectivity (null) • Maternal sensitivity (null)
Nguyen et al., 2021a	72 Mother-infant dyads; Children's gender: 39 boys and 33 girls; Mothers' age: 33.97 ± 4.94 years; Children's age: 4.7 months ± 16 days	fNIRS	Non-interaction (video watching); Interaction (free play)	Proximal joint watching vs. distal joint watching vs. free play (face-to-face interaction)	Time aligned; WTC	Lateral PFC; Medial PFC		Free play > proximal joint watching > distal joint watching	• Affectionate touch (+) • Stimulating touch (-) • Passive touch (null) • Functional touch (null) • Interpersonal physiological synchronization (null)
Nguyen et al., 2021b	66 Father-child dyads; Children's gender: 35 boys and 31 girls; Father's age: 339.2 ± 5.17 years; Children's age: 5.32 ± 0.31 years	fNIRS	Interaction (puzzle-solving task)	Cooperation vs. independent task	Time aligned; WTC	Bilateral dorsolateral PFC; left TPJ		Cooperation > independent and rest	Cooperation: • Father's role attitude (+) • Behavioral reciprocity (null) • Cooperative performance (null) • Paternal sensitivity (null) • Paternal stress (-, marginal) • Child agency (null)

(Continued)

TABLE 1 (Continued)

References	Subjects	Techniques	Task	Variables manipulation	INS calculation	Brain regions	Results		
							Specificity	Situational modulations	Behavioral correlates
Quiñones-Camacho et al., 2020	116 Mother-child dyads; Children's gender: 63 boys and 53 girls; Mother's age: not reported; Children's age: 4.86 ± 0.60 years	fNIRS	Interaction (puzzle-solving task)	Frustration phase vs. recovery phase	Time aligned; Robust correlation coefficient	Lateral PFC		Frustration = recovery	<ul style="list-style-type: none"> • Frustration: Behavioral synchrony (+) • Recovery: Behavioral synchrony (null) • Child irritability (-) • Maternal irritability (null)
Reindl et al., 2018	30 Mother-child dyads and 3 father-child dyads; 21 stranger-child dyads; Children's gender: 18 boys and 15 girls; Parents' age: 41.24 ± 4.32 years; Strangers' age: 24.33 ± 4.70 years; Children's age: 7.52 ± 0.87 years	fNIRS	Interaction (button press task)	Parent-child dyads vs. stranger-child dyads; Cooperation vs. competition	Time aligned; WTC	left dorsolateral and frontopolar PFC	Cooperation: parent-child > stranger-child	Parent-child: cooperation > competition	<ul style="list-style-type: none"> • Parent-child cooperation: • Parent's and child's emotion regulation ability (+) • Cooperative performance (+)
Reindl et al., 2022	34 mother-child dyads and 29 stranger-child dyads; Children's gender: 9 boys and 4 girls; Mother's age: 45.32 ± 4.95 years; Stranger's age: 23.07 ± 2.09 years; Children's age: 14.26 ± 2.21 years	fNIRS	Interaction (button press task)	Parent-child dyads vs. Stranger-child dyads; Cooperation vs. Competition	Time aligned; WTC	Prefrontal cortex	Mother-child dyads > Stranger-child dyads		<ul style="list-style-type: none"> • Competition: • Autonomic nervous system synchrony (+); • Behavioral synchronous responses (-)
Wang et al., 2020	12 Mother-child dyads and 4 father-child dyads; Children's gender: 15 boys and 1 girls; Parents' age: not reported; Children's age: 8.2 ± 1.7 years	fNIRS	Interaction (button press task)	Cooperation vs. independent task	Time aligned; WTC	Superior frontal cortex		Cooperation > independent task	<ul style="list-style-type: none"> • Cooperation-independent: • Autism symptoms (-) • Cooperative performance (+)
Zhao et al., 2021	25 Mother-child dyads and stranger-child dyads; Children's gender: 13 boys and 12 girls; Mother's age: 37.84 ± 3.09 years; Stranger's age: 37.12 ± 4.48 years; Children's age: 7.40 ± 0.28 years	fNIRS	Interaction (cooperative drawing)	Mother-child dyads vs. stranger-child dyads	Time lagged (4s); WTC	Left TPJ			<ul style="list-style-type: none"> • Mother-child dyads: • Children's committed compliance (+) • Children's responsiveness (+) • Mutual responsiveness (-) • Mother's responsiveness (null)
Zivan et al., 2022	24 Mother-child dyads; Children's gender: 16 boys and 8 girls; Mother's age: 35 ± 5 years; Children's age: 33.5 ± 5.8 months	EEG	Interaction (Dialogic Reading)	Uninterrupted vs. Interrupted; Mother-child dyads vs. randomly paired dyads	Time aligned; Circular Correlation	Mother's left and child's right hemispheres	Mother-child dyads > randomly paired dyads	Uninterrupted > Interrupted	<ul style="list-style-type: none"> • Uninterrupted: • Child's age (+); • Mother reading skills (+); • Mother's passive screen exposure (-); • Uninterrupted vs. interrupted: • Joint attention (+)

fNIRS, functional near-infrared spectroscopy; MEG, magnetoencephalography; WTC (), wavelet transform coherence (frequency band); DTW, dynamic time warping; wPLI, weighted phase lag index; PLV, phase-locking-value; STS, superior temporal sulcus; PFC, prefrontal cortex; TPJ, temporo-parietal junction; TD, typically developing; ASD, autism spectrum disorder; LSA, low social anxiety; HSA, high social anxiety; LFC, low family cohesion; HFC, high family cohesion; (+)/(-)/(null), positive/negative/no correlation between the behavioral indices with INS; >, the INS in the former condition was stronger than that in the later condition. = : The INS in the former condition was comparable with that in the later condition. The behavioral correlates results only listed the behavioral indices that associated with the INS indices, while that associated with brain activations was not listed.

dyads, father-child dyads showed stronger INS in left frontal regions while watching family conflict scenes, moreover, the specificity of father-child INS was more obvious for younger fathers (Azhari et al., 2021). Somewhat similar, the specificity of mother-child INS in the left frontal region while video watching was also significant (under $p < 0.05$ level but did not survive multiple comparisons correction), but only when analyzing the data of both positive and conflict scenes together (Azhari et al., 2019). An exception was observed in Azhari et al. (2020), in which the INS of mother-child dyads was comparable with randomly paired stranger-child dyads. The authors concluded that concurrent experiences in the laboratory (e.g., watching the same video) induced high levels of INS even for randomly paired dyads which in turn overruled the specificity of parent-child INS normally derived from specific relationship attachment. In general, these results demonstrate a tendency of parent-child INS to be stronger than randomly paired stranger-child INS under non-interactive situations. Nevertheless, there are other important factors that could interact with effect and should be taken into consideration, such as the gender and age of the parent and valence of the scenes.

In interactive situations, the control condition was set as stranger-child dyads under the same interactive tasks as parent-child dyads. The specificity of parent-child INS was indexed by the stronger INS of parent-child dyads compared to that of interactive stranger-child dyads. Endevelt-Shapira et al. (2021) showed that the parent-infant (about 7-month-old) INS was stronger than stranger-infant INS during face-to-face interaction, and the exposure to maternal chemosignals during stranger-infant interaction attenuated this difference. For children about 8-years-old, Reindl et al. (2018) asked parent-child dyads and stranger-child dyads to play a computer game cooperatively (pressing a button as simultaneously as possible to win the game) or competitively (pressing a button as soon as possible to beat the other player). Stronger INS was found in the left dorsolateral prefrontal and frontopolar cortex for the parent-child dyads in comparison to the stranger-child dyads during cooperation. With the same task, Reindl et al. (2022) replicated the stronger INS for parent-child dyads than stranger-child dyads during cooperation as well as competition. While for typically developing adolescents (8–18 years) in Kruppa et al. (2021), the difference between parent-adolescent and stranger-adolescent (actual dyads) INS was marginally significant for HbR signals, and the difference between the parent-adolescent dyads and randomly paired dyads was significant during competition as well as cooperation. However, the increased parent-adolescent INS was not found for adolescents with autism spectrum disorder. These results demonstrated a trend of the specificity of parent-child INS in interactive situations, especially for typically developing children.

Taken together, these results demonstrate that the specificity of the parent-child INS exists both in the interactive and non-interactive situations. The main explanation is based on

the specific parent-child attachment. That is, the parent-child attachment that formed through long-term parenting results in stronger affective bonding and more joint experiences than that of non-parent relationships. Hence, the non-interactive situations such as watching videos together, would induce more joint processing or shared representations of the social scenes for parent-child dyads, thus leading to stronger INS compared to randomly paired stranger-child dyads (Azhari et al., 2021). Similarly, in the interactive situations parent-child dyads could better understand each other (compared to stranger-child dyads) during cooperation (Reindl et al., 2018) and could engaged in strong social comparison process during competition (Reindl et al., 2018; Kruppa et al., 2021) which resulted in stronger parent-child INS. Such conclusions, however, must be interpreted with caution. For instance, the explanations regarding the effect of attachment on the parent-child INS are rather speculative and lack empirical evidence. What is more, the potential of genetic effect on the specific parent-child INS cannot not be ignored, although this line of investigation remains largely untapped. Finally, the present review demonstrates that there are critical limitations on the specificity of the parent-child INS, especially regarding cooperative or competition situations (Reindl et al., 2018; Kruppa et al., 2021), processing conflict narrative scenes (Azhari et al., 2021), and the gender and age of the parent (Azhari et al., 2020, 2021; Kruppa et al., 2021). However, these limitations have been interpreted from scattered findings and were not tested directly.

Situational modulations on the parent-child inter-brain neural synchronization

For infants as early as the second postnatal day, the mother's left parietal cortex activity synchronized to her newborn's brain activity in superior motor/somatosensory cortex when the newborn was in potential danger situation (painful stimulation) (Bembich et al., 2022). Later, namely, 4–8 months, the mother-infant INS during face-to-face interaction was stronger than that during non-interactive situations (Endevelt-Shapira et al., 2021; Nguyen et al., 2021a), and in the non-interactive situations, the mother-infant INS was stronger during proximal than distal joint video watching (Nguyen et al., 2021a). These findings suggest that the parent-child INS occurs early in life and highly depends on the interactive situations.

For children (>2 years) participants, the situational modulations on the parent-child INS were tested in interactive and non-interactive situations.

In non-interactive situations, the main modulator is the valence of the videos/pictures watched by the parent-child dyads. Levy et al. (2017) videotaped a positive dialog (i.e., planning a fun day) and a conflict dialog (a discussion regarding

a conflict that happened before the dialog took place) between mother and child at home. The videos of the dialogs were encoded as behavioral synchronized episodes (moments when mother and child expressed simultaneous positive affect) and non-synchronized episodes (neither mother nor child expressed positive affect). The mother-child dyads' brain activities were recorded using MEG while they watched videos of themselves and of other mother-child dyads. The results showed that the parent-child gamma-band phase coupling in superior temporal sulcus (STS) was significant while watching their own synchronized episodes. Given that the behavioral synchrony was defined by the mother and child simultaneous expression of positive affect, while non-synchrony was defined by the lack of positive affect from both parts, from the perspective of situational valence, these results indicated that positive situations during social interactions may promote the mother-child INS. However, [Azhari et al. \(2021\)](#) found that the stronger INS of father-child dyads (vs. randomly paired stranger-child dyads) in the frontal regions was only observed while watching videos of family conflict scenes but not during the positive scenes, indicating that the father-child INS was more likely to be induced by situations of negative valence. The authors explained this result as that negative scenes induce higher emotional arousal compared to positive scenes, thus requiring individuals to make greater effort to mentalize the other's emotional states, which finally produces higher INS in the social brain regions. In addition, [Azhari et al. \(2019\)](#) reported null correlations between the mother-child INS and video positivity. Whereas in a later study, [Azhari et al. \(2020\)](#) distinguished the valence of the video stimuli but did not report the difference of the mother-child INS between positive and negative situations. With adolescents as participants (10–14 years), [Deng et al., 2022a,b](#) asked parent-adolescent dyads to watch emotional pictures or videos during EEG recording. The parent-adolescent INS (PLV in gamma band) was stronger during watching positive stimuli than negative stimuli for high social anxiety adolescents and their parents ([Deng et al., 2022a](#)) and for the dyads from high family cohesion (i.e., emotion, support, helpfulness, and caring among family members) ([Deng et al., 2022b](#)). In contrast, the INS in negative condition was stronger than that in positive condition for low social anxiety adolescents ([Deng et al., 2022a](#)). The dyads from low family cohesion showed comparable INS for positive and negative conditions ([Deng et al., 2022b](#)). Overall, researchers have already realized that in non-interactive situations, the valence of the situation affects parent-child INS. Nonetheless, the results were inconsistent and future research on this question will strengthen this idea by considering the individual and family features of the dyads such as gender of the parent, age of the child, social traits and family functioning.

In interactive situations, [Zivan et al. \(2022\)](#) showed that continuous storytelling interaction induced stronger parent-child INS than interrupted situations, in which the parent answered the messages on the mobile phone during the

storytelling, suggesting that the continuity of the interaction was important for INS. Besides this, the main modulators that have been taken into account are the ways in which the parent-child dyads interact, i.e., cooperation, competition, or perform the task independently. [Reindl et al. \(2018\)](#) showed that when compared with competition (pressing a button as soon as possible), cooperation (pressing a button as simultaneously as possible) induced stronger parent-child INS in the dorsolateral prefrontal and frontopolar cortex. Similarly, [Miller et al. \(2019\)](#) used an independent task in which the mother and the child completed a (computer) reaction time task independently as the control condition and found that cooperation induced stronger mother-child INS in the right dorsolateral and frontopolar prefrontal cortex. A stronger parent-child INS in the frontal cortex during cooperation (vs. independent condition) in the reaction time task was also found for children with autism and their parents ([Wang et al., 2020](#)). In addition, in a tangram puzzle-solving task (arranging seven geometric shapes to recreate different templates), similar results were found for mother-child ([Nguyen et al., 2020b](#)) and father-child ([Nguyen et al., 2021b](#)) dyads. That is, compared with independent problem solving or rest (no task), cooperative problem solving induced stronger parent-child INS in the bilateral prefrontal cortex and temporo-parietal areas. Further, using the same cooperative puzzle-solving task, [Quiñones-Camacho et al. \(2020\)](#) distinguished the frustration phase (solving difficult puzzles that induced negative affect on the child) and the following recovery phase (freely playing with toys to recover from the negative affect) and showed comparable mother-child INS in these two phases. One exception was [Kruppa et al. \(2021\)](#), in which stronger parent-adolescent INS was found in competition than cooperation for typically developing groups. But for adolescents with autism spectrum disorder, the INS during competition and cooperation were comparable. Of note, they used adolescents as participants (8–18 years). The development of parent-adolescent relationships and conflict interactions were raised as potential explanations for the high INS during competition. Overall, exclusive of adolescents, the results presented herein consistently showed that cooperative interactions of the parent and the child induced stronger INS in the social brain regions (e.g., frontal cortex, temporo-parietal areas, and STS) compared to competitive interactions, no interactions and rest. The cooperation advantage for parent-child INS was consistently explained in terms of the situational features of the cooperation. That is, compared with competition and independent tasks, cooperation induces more joint attention and similar emotional reactions, and needs continuous adjustment of the behaviors to achieve the goals of cooperation, which requires more effort to understand each other's intentions and mental states. These similar social cognitive processes induced the high levels of parent-child INS during cooperative interactions ([Reindl et al., 2018](#); [Miller et al., 2019](#); [Nguyen et al., 2020b, 2021b](#); [Wang et al., 2020](#)).

Taken together, there is no conclusive evidence that the valence of a situation modulates parent-child INS in non-interactive situations. Whereas in interactive situations, findings frequently show that cooperation (compared to competition or independent tasks) induces stronger parent-child INS possibly because of joint attention, shared emotion, mentalizing and behavioral adjustment during cooperation.

Behavioral correlates of the parent-child inter-brain neural synchronization

In the field of developmental psychology, how the parent-child INS correlates with the parents' and children's social behaviors or traits are of great significance to understand the development of the children's social cognition and social brain. Previous studies have focused on two types of behavioral indices that correlate with parent-child INS, namely, state-like, and trait-like indices. The former is highly dependent on current situations and is often coded during ongoing social interactions, such as communicative reciprocity and cooperative performance. The latter relates to inherent traits or behavioral tendencies that are consistent across time and independent of the ongoing situations, such as the parents' age and children's irritability.

State-like indices

In previous studies, the most mentioned state-like indices that correlate with parent-child INS are the interactive features during interaction and the consequence of the interaction (e.g., cooperative performance). Interactive features such as affective touch during parent-infant free playing was positively correlated with INS, while stimulating touch was negatively correlated with INS (Nguyen et al., 2021a). Communicative reciprocity (i.e., verbal turn-taking) during free conversation between the mother and child positively predicted mother-child INS in frontal and temporoparietal areas, while the INS also increased over the course of the conversation (Nguyen et al., 2020a). The behavioral reciprocity (i.e., contingent responses) during the puzzle-solving cooperation was positively correlated with mother-child INS in the bilateral prefrontal cortex and temporo-parietal areas (Nguyen et al., 2020b) but not with father-child INS (Nguyen et al., 2021b). During computer game competition, less behavioral synchronous responses were associated with higher INS in Reindl et al. (2022). However, Kruppa et al. (2021) did not find significant correlations between motor synchrony (differential response times of each dyad) and the INS. Using gaze features as the measurement of joint attention, Zivan et al. (2022) showed a positive association between the parent-child INS and joint attention during storytelling. Apart from these

reciprocal features, child agency (i.e., interest, vigor, enthusiasm, and eagerness to do the tasks) during cooperation tended to be positively correlated with mother-child INS (Nguyen et al., 2020b) but not with father-child INS (Nguyen et al., 2021b). Whereas parent's sensitivity (i.e., prompt, appropriate, and sensitive responses to the child's signals) during cooperation did not correlate with the father-child (Nguyen et al., 2021b) or the mother-child (Nguyen et al., 2020b) INS. The explanation offered for these positive associations between interactive features and INS was that, the high quality of interaction meant behavioral coordination or synchronization with the signs of joint attention, shared emotions or social adaptive motivation thus promoting INS (Nguyen et al., 2020a, 2021b, Zivan et al., 2022).

However, it should be noted that the positive association is not always observed for father-child dyads (Nguyen et al., 2021b), or on the contrary, it is even shown to be negative (Zhao et al., 2021). Zhao et al. (2021) used a "mother-guided" cooperative drawing task and coded the child's, mother's and mutual responsiveness and the child's committed compliance during drawing. The results showed that the mother-child time-lagged INS in the left temporoparietal junction was positively correlated with the child's responsiveness and committed compliance behavior but not with the responsiveness of the mother. In contrast, the mother-child time-lagged INS was negatively correlated with the mutual responsiveness and the child's committed compliance behavior, which is inconsistent with the results of Nguyen et al., 2020a,b, 2021b. The inconsistency may be a result of the different INS indices. Zhao et al. (2021) used time-lagged (i.e., the child's brain activity lagged behind that of the mother by 4 s) INS, while other studies used time-aligned INS (Nguyen et al., 2020a,b, 2021b). The drawing task in Zhao et al. (2021) was easy for the mother but relatively difficult for the child, and the mother had to teach and guide the child in the drawing task. Thus, the 4 s time lag might reflect a delayed processing of the mother's beliefs in the child's mind during interactions. The inconsistency observed in these results might suggest that mutual communication benefits parent-child INS only for cooperation between equally contributed dyads but not for the leader-follower dyads. Apart from the different INS indices (time-aligned or time lagged), it is also worth noting that some of these studies tested mother-child dyads and some tested father-child dyads and these studies used different cooperative tasks (free conversation, puzzle-solving, drawing, etc.). These different experiment settings restrict the comparability of results and may host potential modulators on the parent-child INS, such as task difficulty, and parent gender.

Apart from the interactive features during interaction, the final cooperative performance also received great attention from investigators. For example, Nguyen et al. (2020b) found that stronger mother-child INS in the bilateral prefrontal cortex and temporo-parietal areas during puzzle-solving cooperation was positively correlated with the cooperative performance

(number of templates solved). Similar results were also found for children with autism spectrum disorder, where the stronger the parent-child INS in the frontal cortex was, the better was their cooperative performance (simultaneous button press) (Wang et al., 2020). What is more, Reindl et al. (2018) tested the causal relation between the parent-child INS and cooperative performance. They divided the cooperation processes (button press task) into two blocks and found that, the parent-child INS in block 1 positively predicted the cooperative performance in block 2, while the cooperation performance in block 1 did not predict the parent-child INS in block 2. This result suggested that the INS could yield better cooperative performance while the cooperative performance did not affect the INS. Note that in the above-mentioned studies, all the participants (Nguyen et al., 2020b) or most of them (Reindl et al., 2018; Wang et al., 2020) were mother-child dyads (Nguyen et al., 2020b). On the contrary, when using father-child dyads as participants, null correlation between the cooperative performance and the INS during puzzle-solving cooperation (vs. independent condition) was found, suggesting that the neural dynamics in father-child problem solving can diverge from mother-child problem solving (Nguyen et al., 2021b). To sum up, most of the existing evidence supported a positive association between parent-child INS (especially for mother-child INS) and cooperative performance (Reindl et al., 2018; Nguyen et al., 2020b; Wang et al., 2020). The explanation is that the INS during cooperation facilitates exchanges of social information between the dyad which helps to achieve the cooperation goal. These findings indicate that the INS may serve as a neural mechanism of successful social cooperation (Reindl et al., 2018; Nguyen et al., 2020b), while the low level of INS may also underlie the social deficits of children with autism disorder (Wang et al., 2020).

Taken together, the association between interactive features and parent-child INS was not consistently observed. In contrast, findings that higher levels of parent-child INS led to better cooperative performance were relatively robust, suggesting a brain-to-brain neural mechanism of successful parent-child cooperation.

Trait-like indices

The trait-like indices of the parent and the child have both been shown to be correlated with parent-child INS. The most mentioned parents' trait-like indices are the emotion-related indices. For example, mother's greater parenting stress was associated with weaker mother-child INS in the medial prefrontal cortex during non-interactive video watching (Azhari et al., 2019). The higher the general stress (including but not limited to economic stress, and interpersonal stress) the mother perceived, the weaker the mother-child INS in the frontal cortex and temporal-parietal areas was during cooperation (Nguyen et al., 2020b). In the same vein, mother-child INS in the medial

prefrontal cortex while watching a video reduced as the maternal attachment anxiety increased (Azhari et al., 2020). Findings concerning the negative associations between stress or anxiety and mother-child INS were explained in terms of the effect of stress or anxiety on the social cognitive process. That is, strong feelings of stress or anxiety may interrupt emotion sharing with the child and prevent the mother from taking the child's perspective, which result in a low level of mother-child INS (Azhari et al., 2019, 2020; Nguyen et al., 2020b). Nonetheless, the negative association between the father's perceived parenting stress and father-child INS was relatively weak (marginally significant effects) (Nguyen et al., 2021b). Furthermore, a study by Reindl et al. (2018) focused on the association of parents' and child's emotion regulation and the parent-child INS. A mediation model was built and demonstrated that the parents used reappraisal as a strategy to regulate their emotions thus promoting the emotion regulation abilities of the child through increasing the parent-child INS in the dorsolateral prefrontal and frontopolar cortex during cooperation. This finding suggests that the parent-child INS might potentially be one mechanism through which the parent's emotion regulation influences the child's emotional development. In addition, the father-child INS in the frontal cortex while watching a video decreased with the father's age. One possible reason is older fathers processed narrative scenes in the presence of their children without feeling the need to excessively attune to the child's emotional state, which in turn decreased the father-child INS (Azhari et al., 2021). Nguyen et al. (2021b) found that when the father's attitude toward their role in child-rearing was more involved, sensitive, and positive, the father-child INS was stronger in the left TPJ and bilateral dorsolateral prefrontal areas during cooperative problem solving. The explanation was that the identification with being a warm and supporting father promoted a higher quality of interaction with their child and resulted in a high level of father-child INS.

Child related trait-like indices include temperament and attachment related indices. However, no evidence was found for the significant associations between the mother-child INS and the negative affectivity of temperament (Nguyen et al., 2020b), or the child's anxious or avoidant attachment (Miller et al., 2019) during cooperation. Distinguishing the frustration (difficult puzzle-solving) and recovery (free playing) phases, the children's irritability was not associated with the mother-child INS in frustration phase, while in the following recovery phase the children's irritability was negatively correlated with the mother-child INS (Quiñones-Camacho et al., 2020). The parent-child INS during storytelling increased with the child's age (Zivan et al., 2022). Additionally, the children with severe autism symptoms (measured by the communication and imagination subscales of the Autism Spectrum Quotient) showed lower levels of parent-child INS in the frontal regions during cooperation (Wang et al., 2020).

Generally, the trait-like indices includes parents' stress (Azhari et al., 2019; Nguyen et al., 2020b, 2021b), the father's age (Azhari et al., 2021) and attitude toward the parenting role (Nguyen et al., 2021b), the mother's and child's attachment styles (Miller et al., 2019; Azhari et al., 2020), the temperament related features (Nguyen et al., 2020b; Quiñones-Camacho et al., 2020), emotion regulation of the parents and child (Reindl et al., 2018), and the children's autism symptom (Wang et al., 2020). Among which, the mother's stress (Azhari et al., 2019; Nguyen et al., 2020b), father's age (Azhari et al., 2021) and parenting role attitude (Nguyen et al., 2021b), child's age (Zivan et al., 2022) and temperamental irritability (Quiñones-Camacho et al., 2020), parents' and child's emotion regulation (Reindl et al., 2018), and the children's autism symptom (Wang et al., 2020) showed associations with the parent-child INS in the frontal regions and temporal-parietal areas, while the remaining indices did not show clear associations with parent-child INS. It is worth noting that each trait-like index listed above was mentioned in only one or two studies, and were tested under different tasks (e.g., video watching, cooperation, etc.). Thus, the associations between trait-like indices and parent-child INS should be considered with caution and verified with future studies.

To conclude, the results regarding associations between behavioral indices and parent-child INS remain somewhat unclear. In spite of this, the explanations about the association results were relatively consistent and focused on the association between the behavioral features of the dyad and the brain functions during social interaction. Specifically, the associations between parent-child INS and the state-like indices (e.g., turn-taking and cooperative performance) were explained in terms of the high quality of parent-child interaction reflecting the coordination or synchronization of their behaviors, which facilitated INS (Nguyen et al., 2020a, 2021b). As a result, the INS could also promote the social exchanges of the dyad and increase their behavioral performance (Reindl et al., 2018; Nguyen et al., 2020b). In other words, effective social interaction provided the possibility of INS, whereas the high level of INS served as the neural bases for the following effective social interaction (e.g., cooperation). Notably, this explanation emphasized that the INS highly relied on the ongoing social situations, thus it could also explain why the INS showed relatively weak associations with trait-like indices which are independent of the ongoing situations. For example, the attachment style is an important index highly related with parent-child relationships but was not related with parent-child INS during specific social interactive situations (Miller et al., 2019; Azhari et al., 2020). One possible explanation for the null association may be that, the measurement of the attachment reflected long-term parent-child behavioral tendencies, therefore it is difficult to show the effects of attachment styles on parent-child INS in temporary social interactive situations manipulated in the laboratory setting (Azhari et al., 2020; Nguyen et al., 2020b).

Discussion

Psychological perspectives

As reviewed above, up to now, the relatively consistent findings regarding parent-child INS depart from two main aspects, namely, the specificity of the parent-child INS (i.e., stronger INS for parent-child dyads than stranger-child dyads) and the situational dependency of the parent-child INS (i.e., the stronger parent-child INS during cooperation than competition or independent situations, and the positive association between parent-child INS and cooperative performance). In contrast, the associations between the parents'/child's trait-like behavioral indices and INS were complex and did not allow for clear conclusions. Nonetheless, two main aspects influencing the INS could be observed, namely, interpersonal relationship (i.e., parent-child vs. stranger-child) and interactive features (i.e., cooperation vs. competition). Other factors, however, such as individual features (e.g., attachment styles, etc.) of the parents or the children did not show clear influence on the parent-child INS. To recap, interpersonal relationship was the main psychological explanation regarding the specificity of the parent-child INS. That is, the long-term attachment formed between parents and their children marks a specific affective connection with shared experiences and memories, which induce similar neural responses when experiencing the same social events, that ultimately result in high level of parent-child INS (Reindl et al., 2018; Azhari et al., 2021). In contrast, the situational dependency of the parent-child INS was explained in term of the interactive features of the ongoing social interactions such as the shared attention, similar emotional responses, strong intention to understand the other and the mutual adaptability to achieve the cooperative goals (Reindl et al., 2018; Miller et al., 2019; Nguyen et al., 2020b, 2021b; Quiñones-Camacho et al., 2020; Wang et al., 2020). Although the influence of the individual features on the parent-child INS were complex, the explanations were also mostly focused on the social cognition processes. For example, some traits effectively facilitated mentalizing or emotion sharing, which affected the INS (Reindl et al., 2018; Azhari et al., 2019, 2020, 2021; Miller et al., 2019; Nguyen et al., 2020b, 2021b; Quiñones-Camacho et al., 2020; Wang et al., 2020).

To sum up, two main explanations for the parent-child INS were observed in this review, i.e., long-term formed attachment relationship and the effective social cognitive processes based on the interactive features. These two explanations are complementary to each other. Nevertheless, these explanations were all *post-hoc* inferences and, therefore they lack empirical evidence. For the attachment account, previous studies did not find clear associations between the attachment styles and parent-child INS (Miller et al., 2019; Azhari et al., 2020), while associations between parent-child INS and other relevant indices, such as parent-child relationship, companionship, and

co-living time, were not mentioned so far. In addition, the hereditary account for the specificity of the parent-child INS was surprisingly ignored. For the ongoing social cognition explanation, the frequently mentioned emotion sharing and mentalizing processes that depend on the situations were not directly measured in previous studies to provide objective evidence.

Neuroscience perspectives

The above mentioned psychological explanations of INS are based on the psychological perspective of social interaction. On the other hand, neuroscience findings of the brain functions also support these explanations. Existing studies about the specificity, situational modulation and the behavioral correlates of the parent-child INS typically showed on the frontal cortex (Reindl et al., 2018; Azhari et al., 2019, 2020, 2021; Miller et al., 2019; Nguyen et al., 2020a,b, 2021b; Quiñones-Camacho et al., 2020; Wang et al., 2020), temporal-parietal junction (Nguyen et al., 2020a,b, 2021b; Zhao et al., 2021), and superior temporal sulcus (Levy et al., 2017). These are typical regions of the social brain network, which responsible for mentalizing, emotional control, and attention to social information (Brothers, 1990; Frith and Frith, 2007; Blakemore, 2008). Based on the functions of these brain regions, the parent-child INS in the dorsolateral prefrontal cortex was suggested to represent similar cognitive control processes of the dyad, especially the emotional control process (Reindl et al., 2018). The INS in the medial prefrontal cortex, frontopolar and the temporal-parietal regions were suggested to represent the similar processing of the social information or mentalizing of the dyad (Reindl et al., 2018; Miller et al., 2019; Azhari et al., 2020; Nguyen et al., 2020a,b, 2021b). The parent-child INS in the superior temporal sulcus was suggested to represent the integrations of the bottom-up and top-down processes to create online bio-behavioral coordination between the dyad (Levy et al., 2017).

Concerning these explanations, three points need to be noted. (i) The explanations were drawn from *post-hoc* inferences based on the general function of these social brain regions rather than empirical evidence through strict manipulation or measurement of the social cognition process (e.g., mentalizing or emotional control). Since one brain region would be engaged in multiple cognitive functions, the *post-hoc* functional inferences based on the brain functions should be regarded with caution. (ii) Although the existing studies consistently focused on the social brain regions such as frontal and temporal-parietal regions, the parent-child INS was found in different regions in different studies. The specificity of parent-child INS was consistently found in the frontal regions (Reindl et al., 2018; Azhari et al., 2021). The situational modulations were widely found in the frontal, temporal-parietal junction or superior temporal sulcus (Reindl et al., 2018; Miller et al., 2019; Nguyen

et al., 2020b, 2021b; Wang et al., 2020). The INS showed significant associations with behavioral indices in the frontal regions or temporal-parietal regions (Reindl et al., 2018; Nguyen et al., 2020a,b, 2021b; Quiñones-Camacho et al., 2020; Wang et al., 2020; Zhao et al., 2021). In addition, these regions showed different hemisphere asymmetry in different studies (Reindl et al., 2018; Azhari et al., 2019, 2020, 2021; Miller et al., 2019; Wang et al., 2020; Zhao et al., 2021), however, these differences on the locations of the parent-child INS were not discussed. (iii) Most of the existing studies used the fNIRS to record the brain activities (Reindl et al., 2018; Azhari et al., 2019, 2020, 2021; Miller et al., 2019; Nguyen et al., 2020a,b, 2021b; Quiñones-Camacho et al., 2020; Wang et al., 2020; Zhao et al., 2021). Because of the limitations posed by fNIRS, only the brain activities of the cerebral cortex regions could be recorded. The parent-child INS in other key nodes in the social brain network such as insular, medial prefrontal cortex, cingulate cortex and precuneus (Brothers, 1990; Frith and Frith, 2007; Blakemore, 2008) have not been investigated yet.

Inspirations

Generally, the current findings reviewed in the current article indicate that parent-child INS have positive meanings on the children's social and affective development. For example, high level of parent-child INS reflected high quality of parent-child interaction, facilitated cooperation (Reindl et al., 2018; Nguyen et al., 2020b; Wang et al., 2020) and correlates with child's committed compliance (Zhao et al., 2021). The parents' emotional regulation strategies could also promote the development of the child's emotion regulation abilities through parent-child INS (Reindl et al., 2018). Thus, it can be seen that creating favorable conditions to promote parent-child INS may benefit children's development.

The existing findings may provide valuable information for the parents as to how to promote parent-child INS. First, existing studies consistently showed that cooperative tasks (relative to competition or independent task) induced higher parent-child INS in the social brain regions possibly through motivating joint attention, inducing similar emotional responses, and promoting understanding and inferring the other's mental states (Reindl et al., 2018; Miller et al., 2019; Nguyen et al., 2020b, 2021b; Wang et al., 2020). Thus, creating cooperative situations in family life may help to produce parent-child INS that benefit the social development of the children. Second, high levels of parent-child INS can be achieved when parents provide appropriate, and in-time responses during parent-child interaction because high quality of social communication and information exchange, such as joint attention, similar emotional responses and motivation to adapt to each other, produces high levels of INS (Nguyen et al., 2020b, 2021b). Notably, this conclusion builds on the

equal cooperation of the parent and the child. When the cooperation task was too difficult for the child and the parent's guidance was needed, mutual responsiveness might decrease the parent-child INS (Zhao et al., 2021). Third, parent-child INS could potentially be improved when parents maintain positive psychological states. There is evidence that high levels of stress or anxious attachment prevents the mother from taking the child's perspective and sharing their emotion, which makes INS hard to achieve (Azhari et al., 2019, 2020; Nguyen et al., 2020b). What is more, parents' effective emotion regulation could promote the development of the child's emotion regulation ability through parent-child INS (Reindl et al., 2018). Thus, by keeping a positive psychological state (e.g., good stress coping and emotion regulation) parents can promote the parent-child INS and generate a positive effect on the children's social and emotional development. Last but not least, the INS can serve as a potential index for evaluating social dysfunctions such as those observed in autistic population. As found in Wang et al. (2020), for children with autism spectrum disorder, the weaker the parent-child INS, the severer the autism symptoms, and the worse the cooperative performance. These findings suggest that the abnormality of INS might underlie the social dysfunctions of autism. Thus, using the INS as a potential biological index of social deficits may be a promising means for the diagnosis and intervention of autism.

Although these benefits from parent-child INS require targeted investigation, the findings aforementioned provided meaningful inspirations and primary guidance for parent-child relationship and family education, which will benefit the child's social development.

Limitations and future directions

Overall, the existing research on parent-child INS has made great progress and set the path for the investigation of children's social development from the perspective of social neuroscience. In spite of this, there are some limitations and unanswered questions that need to be investigated in future studies.

First, the existing findings were preliminary and the explanations of these findings need further empirical evidence. It can be seen that there were some inconsistent findings, such as whether the valence of a situation modulates parent-child INS in non-interactive situations, or which interactive or individual features were associated with the parent-child INS, need to be clarified in future studies. Additionally, some of the findings were only evident in one or two studies (e.g., the associations between trait-like indices and parent-child INS) and would benefit from further replications. Moreover, as mentioned above, the explanations about the parent-child INS (i.e., attachment relationship and social cognitive processes) are mostly speculative and lack empirical evidence, and alternative explanations, such as heredity of the brain functions, need to be

further studied. In addition, given that the associations between the parent-child INS and children's behavioral indices were inconclusive, the real meaning of the parent-child INS to the children's social development still needs more investigation.

Second, comparisons between mother, father or other caregivers (e.g., grandparents) are scarce. To date, most studies used mother-child dyads (Levy et al., 2017; Azhari et al., 2019, 2020; Miller et al., 2019; Nguyen et al., 2020a, 2021b,a; Quiñones-Camacho et al., 2020; Endevelt-Shapira et al., 2021; Zhao et al., 2021; Bembich et al., 2022; Reindl et al., 2022; Zivan et al., 2022), five studies used both mother-child and father-child dyads (Reindl et al., 2018; Wang et al., 2020; Kruppa et al., 2021; Deng et al., 2022a,b), two studies used father-child dyads (Azhari et al., 2021; Nguyen et al., 2021b), and no studies tested other caregiver-child dyads. The fact that most studies tested mother-child dyads makes sense as mothers are primary caregiver, but it has been shown, that the father plays important role in children's development (Feldman, 2003; Liong, 2017), besides other caregivers such as grandparents also provide considerable help during the children's growth (Attar-Schwartz et al., 2009; Mahne and Huxhold, 2015). The current explanations (i.e., attachment and social cognitive processes) regarding INS are not limited to mother-child relationships. For instance, father and other caregivers can also form a specific attachment relationship with the child which could result in specific INS with the child. Of note, the mother may induce very specific features compared with the father or other caregivers. For example, mothers and fathers co-create distinct types of behavioral synchrony with the infant; maternal synchrony is more cyclic and social oriented while paternal synchrony orients toward the environment and encourages exploration (Feldman, 2003). Recent findings have already shown a few different results between mother- and father-child dyads. For example, the mother-child INS decreased with the mother's parenting stress (Azhari et al., 2019), but it is not the case for the father (Nguyen et al., 2021b). Nonetheless, there are also similar results for the mother-child and father-child dyads. For instance, the INS during cooperative puzzle-solving tasks was stronger than the independent task for both the mother-child (Nguyen et al., 2020b) and father-child (Nguyen et al., 2021b) dyads. Hence, future research would benefit from the use of systematic comparisons between the INS of the mother-, father-, and even other caregiver-child dyads as these would help to understand the importance of the different parenting roles on children's social development.

Third, whether and how the parent-child INS changes with the children's age were not tested. The children tested in existing studies were within four age groups, i.e., infants younger than 1 year old (Endevelt-Shapira et al., 2021; Nguyen et al., 2021a; Bembich et al., 2022), children about 2–5 years (Azhari et al., 2019, 2020, 2021; Nguyen et al., 2020a,b, 2021b; Quiñones-Camacho et al., 2020), children about 7–12 years

(Levy et al., 2017; Reindl et al., 2018; Miller et al., 2019; Wang et al., 2020; Zhao et al., 2021; Zivan et al., 2022), and adolescents about 10–18 years (Deng et al., 2022a,b; Kruppa et al., 2021; Reindl et al., 2022). It is known that from infants to 5 years are preschool ages while 7–12 years are school ages and 10–18 years are in puberty. School education could also affect the social brain development of the children (see Lieberman, 2012 for a review) and the parent-child relationship changes a lot from childhood to puberty (Branje, 2018). Although parent-child behavioral synchrony affected the development of children's social skills across childhood and up to adolescence (see Feldman, 2012 for a review), it is worth to test whether the parent-child INS would be different before and after the school age or before and after the puberty. In other words, whether the parent-child INS changes with the growth of the children needs to be answered in future research, which would provide a full view as to what extent the parent-child relationship affects the social development of the children.

Forth, two aspects of methodology pitfalls are worth noting. One is that the recording of brain activities using fNIRs only targets cerebral cortex regions (e.g., TPJ) of the social brain network. Whether other key nodes in the social brain network such as insular, medial prefrontal cortex, cingulate cortex and precuneus would show similar features of the parent-child INS are still unknown. The other pitfall is the inconsistency of the INS measurement. Currently, the INS measurement is far from established. For example, some studies used cross-correlation analysis (Azhari et al., 2020, 2021; Quiñones-Camacho et al., 2020), while other studies used coherence analysis (i.e., WTC) (Reindl et al., 2018; Miller et al., 2019; Nguyen et al., 2020a,b, 2021b; Wang et al., 2020; Zhao et al., 2021). The discrepancy in INS calculation has to be taken into account when comparing the results of different studies. In addition, most studies used time-aligned indices of parent-child INS (e.g., Levy et al., 2017; Azhari et al., 2019; Miller et al., 2019; Nguyen et al., 2020b, 2021b) and only one study used the time-lagged index (Zhao et al., 2021). The pattern of time-lagged INS has been found between the leader and the follower (i.e., the neural activity of the follower was a few seconds lags behind that of the leader) during the leader guided cooperation (Sänger et al., 2013; Konvalinka et al., 2014; Jiang et al., 2015). This kind of leader-follower relationship during interaction was similar to the parent-child interaction in daily life, for example, the parents always needed to teach the children life skills or provide guidance when the children got stuck in difficult tasks. In these types of situations, it is possible that the parent-child INS would also show the time-lagged pattern. However, up to now, only one study used the parent-guided drawing task and the time-lagged parent-child INS as indices (Zhao et al., 2021), while the remaining studies used time-aligned INS indices (e.g., Levy

et al., 2017; Azhari et al., 2019; Miller et al., 2019; Nguyen et al., 2020b, 2021b). Future investigation should focus on the parent-guided interactions, whereas testing the time-lagged INS would provide new insights to understand the neural correlates of the parent-child interaction.

To summarize, existing studies have demonstrated the specificity of the parent-child INS, investigated the situational factors that affect the parent-child INS, and correlated the parent-child INS with the parents' and child's state-like and trait-like behavioral indices. The existing findings also indicate that the parent-child INS may provide a very promising avenue for parent-child relationship and the social development of the children. In spite of this, it is worth noting that these studies used different tasks, different gender of participants, and different indices, which constrained the comparisons between findings, and raised a number of questions to be answered by future studies.

Author contributions

YiL and JL organized all the existing findings and wrote the manuscript. QW and YaL searched the literatures and revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the Humans and Social Sciences Youth Fund of the Ministry of Education of China (No. 18YJCZH109).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Attar-Schwartz, S., Tan, J. P., Buchanan, A., Flouri, E., and Griggs, J. (2009). Grandparenting and adolescent adjustment in two-parent biological, lone-parent, and step-families. *J. Fam. Psychol.* 23, 67–75. doi: 10.1037/a0014383
- Azhari, A., Bizzego, A., and Esposito, G. (2021). Father-child dyads exhibit unique inter-subject synchronization during co-viewing of animation video stimuli. *Soc. Neurosci.* 16, 522–533. doi: 10.1080/17470919.2021.1970016
- Azhari, A., Gabrieli, G., Bizzego, A., Bornstein, M. H., and Esposito, G. (2020). Probing the association between maternal anxious attachment style and mother-child brain-to-brain coupling during passive co-viewing of visual stimuli. *Attach. Hum. Dev.* 1, 1–16. doi: 10.1080/14616734.2020.1840790
- Azhari, A., Leck, W. Q., Gabrieli, G., Bizzego, A., Rigo, P., Setoh, P., et al. (2019). Parenting stress undermines mother-child brain-to-brain synchrony: A hyperscanning study. *Sci. Rep.* 9:11407. doi: 10.1038/s41598-019-47810-4
- Bembich, S., Saksida, A., Mastromarino, S., Travan, L., Di Risio, G., Cont, G., et al. (2022). Empathy at birth: Mother's cortex synchronizes with that of her newborn in pain. *Eur. J. Neurosci.* 55, 1519–1531. doi: 10.1111/ejn.15641
- Blakemore, S. J. (2008). The social brain in adolescence. *Nat. Rev. Neurosci.* 9, 267–277. doi: 10.1038/nrn2353
- Branje, S. (2018). Development of parent-adolescent relationships: Conflict interactions as a mechanism of change. *Child Dev. Perspect.* 12, 171–176. doi: 10.1073/pnas.0709815105
- Brothers, L. (1990). The social brain: A project for integrating primate behaviour and neurophysiology in a new domain. *Concepts Neurosci.* 1, 27–51.
- Cui, X., Bryant, D. M., and Reiss, A. L. (2012). NIRS-based hyperscanning reveals increased interpersonal coherence in superior frontal cortex during cooperation. *Neuroimage* 59, 2430–2437. doi: 10.1016/j.neuroimage.2011.09.003
- Deng, X., Chen, X., Zhang, L., Gao, Q., Li, X., and An, S. (2022a). Adolescent social anxiety undermines adolescent-parent interbrain synchrony during emotional processing: A hyperscanning study. *Int. J. Clin. Health Psychol.* 22:100329. doi: 10.1016/j.ijchp.2022.100329
- Deng, X., Lin, M., Zhang, L., Li, X., and Gao, Q. (2022b). Relations between family cohesion and adolescent-parent's neural synchrony in response to emotional stimulations. *Behav. Brain Funct.* 18:11. doi: 10.1186/s12993-022-00197-1
- Endevelt-Shapira, Y., Djalovski, A., Dumas, G., and Feldman, R. (2021). Maternal chemosignals enhance infant-adult brain-to-brain synchrony. *Sci. Adv.* 7:eabg6867. doi: 10.1126/sciadv.abg6867
- Feldman, R. (2003). Infant-mother and infant-father synchrony: The coregulation of positive arousal. *Infant Ment. Health J.* 24, 1–23. doi: 10.1002/imhj.10041
- Feldman, R. (2007a). Parent-infant synchrony and the construction of shared timing: physiological precursors, developmental outcomes, and risk conditions. *J. Child Psychol. Psychiatry* 48, 329–354. doi: 10.1111/j.1469-7610.2006.01701.x
- Feldman, R. (2007b). On the origins of background emotions: From affect synchrony to symbolic expression. *Emotion* 7, 601–611. doi: 10.1037/1528-3542.7.3.601
- Feldman, R. (2007c). Mother-infant synchrony and the development of moral orientation in childhood and adolescence: Direct and indirect mechanisms of developmental continuity. *Am. J. Orthopsychiatry* 77, 582–597. doi: 10.1037/0002-9432.77.4.582
- Feldman, R. (2012). Bio-behavioral synchrony: A model for integrating biological and microsocial behavioral processes in the study of parenting. *Parent. Sci. Pract.* 12, 154–164. doi: 10.1080/15295192.2012.683342
- Frith, C. D., and Frith, U. (2007). Social cognition in humans. *Curr. Biol.* 17, R724–R732.
- Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., and Keysers, C. (2012). Brain-to-brain coupling: A mechanism for creating and sharing a social world. *Trends Cogn. Sci.* 16, 114–121. doi: 10.1016/j.tics.2011.12.007
- Jiang, J., Chen, C., Dai, B., Shi, G., Ding, G., Liu, L., et al. (2015). Leader emergence through interpersonal neural synchronization. *Proc. Natl. Acad. Sci. U.S.A.* 112, 4274–4279. doi: 10.1073/pnas.1422930112
- Jiang, J., Dai, B., Peng, D., Zhu, C., Liu, L., and Lu, C. (2012). Neural synchronization during face-to-face communication. *J. Neurosci.* 32, 16064–16069. doi: 10.1523/JNEUROSCI.2926-12.2012
- Jiang, J., Zheng, L. F., and Lu, C. M. (2021). A hierarchical model for interpersonal verbal communication. *Soc. Cogn. Affect. Neurosci.* 16, 246–255. doi: 10.1093/scan/nsaa151
- Konvalinka, I., Bauer, M., Stahlhut, C., Hansen, L. K., Roepstorff, A., and Frith, C. D. (2014). Frontal alpha oscillations distinguish leaders from followers: Multivariate decoding of mutually interacting brains. *Neuroimage* 94, 79–88. doi: 10.1016/j.neuroimage.2014.03.003
- Kruppa, J. A., Reindl, V., Gerloff, C., Oberwilleand Weiss, E., Prinz, J., Herpertz-Dahlmann, B., et al. (2021). Brain and motor synchrony in children and adolescents with ASD—a fNIRS hyperscanning study. *Soc. Cogn. Affect. Neurosci.* 16, 103–116. doi: 10.1093/scan/nsaa092
- Levy, J., Goldstein, A., and Feldman, R. (2017). Perception of social synchrony induces mother-child gamma coupling in the social brain. *Soc. Cogn. Affect. Neurosci.* 12, 1036–1046. doi: 10.1093/scan/nsx032
- Lieberman, M. D. (2012). Education and the social brain. *Trends Neurosci. Educ.* 1, 3–9.
- Liong, M. (2017). *Chinese fatherhood, gender and family: Father mission*. Berlin: Springer. doi: 10.1057/978-1-137-44186-7
- Mahne, K., and Huxhold, O. (2015). Grandparenthood and subjective well-being: Moderating effects of educational level. *J. Gerontol. Ser. B Psychol. Sci. Soc. Sci.* 70, 782–792. doi: 10.1093/geronb/gbu147
- Miller, J. G., Vrticka, P., Cui, X., Shrestha, S., Hosseini, S. M. H., Baker, J. M., et al. (2019). Inter-brain synchrony in mother-child dyads during cooperation: An fNIRS hyperscanning study. *Neuropsychologia* 124, 117–124. doi: 10.1016/j.neuropsychologia.2018.12.021
- Nguyen, T., Schleihau, H., Kungl, M., Kayhan, E., Hoehl, S., and Vrticka, P. (2021b). Interpersonal neural synchrony during father-child problem solving: An fNIRS hyperscanning study. *Child Dev.* 92, E565–E580. doi: 10.1111/cdev.13510
- Nguyen, T., Abney, D. H., Salamander, D., Bertenthal, B. I., and Hoehl, S. (2021a). Proximity and touch are associated with neural but not physiological synchrony in naturalistic mother-infant interactions. *Neuroimage* 244:118599.
- Nguyen, T., Schleihau, H., Kayhan, E., Matthes, D., Vrticka, P., and Hoehl, S. (2020a). Neural synchrony in mother-child conversation: Exploring the role of conversation patterns. *Soc. Cogn. Affect. Neurosci.* 16, 93–102. doi: 10.1093/scan/nsaa079
- Nguyen, T., Schleihau, H., Kayhan, E., Matthes, D., Vrticka, P., and Hoehl, S. (2020b). The effects of interaction quality on neural synchrony during mother-child problem solving. *Cortex* 124, 235–249. doi: 10.1016/j.cortex.2019.11.020
- Quiñones-Camacho, L. E., Fishburn, F. A., Camacho, M. C., Hlutsowsky, C. O., Huppert, T. J., Wakschlag, L. S., et al. (2020). Parent-child neural synchrony: A novel approach to elucidating dyadic correlates of preschool irritability. *J. Child Psychol. Psychiatry* 61, 1213–1223. doi: 10.1111/jcpp.13165
- Redcay, E., and Schilbach, L. (2019). Using second-person neuroscience to elucidate the mechanisms of social interaction. *Nat. Rev. Neurosci.* 20, 495–505. doi: 10.1038/s41583-019-0179-4
- Reindl, V., Gerloff, C., Scharke, W., and Konrad, K. (2018). Brain-to-brain synchrony in parent-child dyads and the relationship with emotion regulation revealed by fNIRS-based hyperscanning. *Neuroimage* 178, 493–502. doi: 10.1016/j.neuroimage.2018.05.060
- Reindl, V., Wass, S., Leong, V., Scharke, W., Wistuba, S., Wirth, C. L., et al. (2022). Multimodal hyperscanning reveals that synchrony of body and mind are distinct in mother-child dyads. *Neuroimage* 251:118982. doi: 10.1016/j.neuroimage.2022.118982
- Sänger, J., Müller, V., and Lindenberger, U. (2013). Directionality in hyperbrain networks discriminates between leaders and followers in guitar duets. *Front. Hum. Neurosci.* 7:234. doi: 10.3389/fnhum.2013.00234
- Wang, Q., Han, Z., Hu, X., Feng, S., Wang, H., Liu, T., et al. (2020). Autism Symptoms modulate interpersonal neural synchronization in children with autism spectrum disorder in cooperative interactions. *Brain Topogr.* 33, 112–122. doi: 10.1007/s10548-019-00731-x
- Zhao, H., Cheng, T., Zhai, Y., Long, Y., Wang, Z., and Lu, C. (2021). How mother-child interactions are associated with a child's compliance. *Cereb. Cortex* 31, 4398–4410. doi: 10.1093/cercor/bha094
- Zivan, M., Gashri, C., Habuba, N., and Horowitz-Kraus, T. (2022). Reduced mother-child brain-to-brain synchrony during joint storytelling interaction interrupted by a media usage. *Child Neuropsychol.* 28, 918–937. doi: 10.1080/09297049.2022.2034774



OPEN ACCESS

EDITED BY
Pedro Gomez-Vilda,
Polytechnic University of Madrid, Spain

REVIEWED BY
Xin Wang,
Shenzhen Institutes of Advanced
Technology (CAS), China
Ludwig Smeele,
Antoni van Leeuwenhoek Hospital,
Netherlands

*CORRESPONDENCE
Orlando Guntinas-Lichius
orlando.guntinas@med.uni-jena.de

†These authors have contributed
equally to this work and share first
authorship



SPECIALTY SECTION
This article was submitted to
Motor Neuroscience,
a section of the journal
Frontiers in Human Neuroscience

RECEIVED 31 August 2022
ACCEPTED 22 November 2022
PUBLISHED 12 December 2022

CITATION
Mueller N, Trentzsch V, Grassme R,
Guntinas-Lichius O, Volk GF and
Anders C (2022) High-resolution
surface electromyographic activities
of facial muscles during mimic
movements in healthy adults:
A prospective observational study.
Front. Hum. Neurosci. 16:1029415.
doi: 10.3389/fnhum.2022.1029415

COPYRIGHT
© 2022 Mueller, Trentzsch, Grassme,
Guntinas-Lichius, Volk and Anders.
This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License](#)
(CC BY). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

High-resolution surface electromyographic activities of facial muscles during mimic movements in healthy adults: A prospective observational study

Nadiya Mueller^{1,2†}, Vanessa Trentzsch^{1,2†}, Roland Grassme^{1,3},
Orlando Guntinas-Lichius ^{2,4,5*}, Gerd Fabian Volk^{2,4,5} and
Christoph Anders ¹

¹Division Motor Research, Pathophysiology and Biomechanics, Department of Trauma, Hand and Reconstructive Surgery, Jena University Hospital, Friedrich-Schiller-University Jena, Jena, Germany, ²Department of Otolaryngology, Jena University Hospital, Friedrich-Schiller-University Jena, Jena, Germany, ³Department of Prevention, Biomechanics, German Social Accident Insurance Institution for the Foodstuffs and Catering Industry, Erfurt, Germany, ⁴Facial-Nerve-Center Jena, Jena University Hospital, Jena, Germany, ⁵Center for Rare Diseases, Jena University Hospital, Jena, Germany

Objectives: Surface electromyography (sEMG) is a standard tool in clinical routine and clinical or psychosocial experiments also including speech research and orthodontics to measure the activity of selected facial muscles to objectify facial movements during specific facial exercises or experiments with emotional expressions. Such muscle-specific approaches neglect that facial muscles act more as an interconnected network than as single facial muscles for specific movements. What is missing is an optimal sEMG setting allowing a synchronous measurement of the activity of all facial muscles as a whole.

Methods: A total of 36 healthy adult participants (53% women, 18–67 years) were included. Electromyograms were recorded from both sides of the face using an arrangement of electrodes oriented by the underlying topography of the facial muscles (Fridlund scheme) and simultaneously by a geometric and symmetrical arrangement on the face (Kuramoto scheme). The participants performed a standard set of different facial movement tasks. Linear mixed-effects models and adjustment for multiple comparisons were used to evaluate differences between the facial movement tasks, separately for both applied schemes. Data analysis utilized sEMG amplitudes and also their maximum-normalized values to account for amplitude differences between the different facial movements.

Results: Surface electromyography activation characteristics showed systematic regional distribution patterns of facial muscle activation for both schemes with very low interindividual variability. The statistical significance to discriminate between the different sEMG patterns was good for both

schemes (significant comparisons for sEMG amplitudes: 87.3%, both schemes, normalized values: 90.9%, Fridlund scheme, 94.5% Kuramoto scheme), but the Kuramoto scheme performed considerably superior.

Conclusion: Facial movement tasks evoke specific patterns in the complex network of facial muscles rather than activating single muscles. A geometric and symmetrical sEMG recording from the entire face seems to allow more specific detection of facial muscle activity patterns during facial movement tasks. Such sEMG patterns should be explored in more clinical and psychological experiments in the future.

KEYWORDS

facial muscle activations, facial movement analysis, facial nerve (FN), surface EMG (sEMG), high-density sEMG, face

Introduction

Facial muscle contractions form mimic expressions important for many functions in human behavior (Cattaneo and Pavesi, 2014). They enable facial motor control, for instance, for eye closure or mouth control during drinking and eating. Moreover, facial expressions reflect emotions and are essential for non-verbal communication. Therefore, facial electromyography (EMG) represents a standard in psychological experiments for psychophysiological measures to test the group activity of some facial muscles and its association with specific emotions (Hubert and de Jong-Meyer, 1991). Furthermore, EMG investigations of facial muscles are a standard for facial muscle diagnostics in patients with facial nerve and muscle diseases (Guntinas-Lichius et al., 2020). Facial EMG recordings in a psychological and clinical setting usually use only two to six pairs of needle or surface electrodes to record the resting state or voluntary movements of some facial muscles on the assumption that this reflects the respective mimic muscles as a whole. This approach and assumption neglect that the facial muscular system forms a complex interdependent system of the individual anatomically defined facial muscles that are connected to the skin (Cattaneo and Pavesi, 2014). They have no fixed bone-based insertion points, are interwoven, and are partly overlapping. The activation of one muscle with the aim to induce a movement in a certain direction involuntarily elicits the activation of other facial muscles. This activation of other facial muscles is needed to create a fixation point for the intended movement. Therefore, each facial exercise leads to a complex activation of almost all facial muscles (Schumann et al., 2021).

To approach a realistic setting, Fridlund and Cacioppo recommended surface EMG (sEMG) recordings of 1 masticatory and 10 facial muscles (Fridlund and Cacioppo, 1986) for psychophysiological investigations.

Kuramoto et al. (2019) recommended the use of 21 sEMG electrodes in an EEG-like arrangement. The idea behind this particular approach is that such a muscle site-independent recording over the face might better reflect the complex facial activity during exercises or emotional expressions (Kuramoto et al., 2019). Recently, experimental high-density sEMG (HD sEMG) applications with 90 electrodes even allowed the visualization of mimic muscle activation projected on the facial surface (Cui et al., 2020). Lately, we published a study and an atlas on facial muscle activation patterns based on multichannel sEMG (Schumann et al., 2010, 2021). The sEMG amplitude characteristics were transposed into a color-coded atlas of voluntary facial muscle activation tasks. We could show that facial muscles act more as a whole connected muscle network than as single facial muscles, responsible for certain movements.

As a next step and in a new clinical trial, we, therefore, wanted to use both the sEMG electrode schemes of Fridlund and Cacioppo (1986) and Kuramoto et al. (2019) synchronously to analyze which scheme allows optimal visualization of the facial muscle activity of the entire face during specific facial movements. Furthermore, we asked which of the two schemes is better suited to distinguish different facial expressions from each other.

Materials and methods

Healthy participants

The study included 36 healthy adult volunteers (52.8% women) with no neurological diseases or history of facial surgery (mean age: 33.3 ± 15.2 years; range: 18–67). All participants gave written informed consent to participate in this study. The ethics committee of the Jena University Hospital approved this study (No. 2019-1539). The individual shown

in the figures of this manuscript has given written informed consent to publish these case details.

Standardization of mimic exercises

The participants were instructed about the sequence of the examination and the facial exercises. The experimental workup is shown in **Figures 1A,B**. Room temperature and relative humidity were controlled to be approximately 24°C and 40–65%, respectively. Participants sat in a relaxed upright position in front of a computer screen and followed a self-explanatory video tutorial published recently (Schaede et al., 2017). This provides standardized and reliable instruction to perform facial movements guided by a human instructor (Volk et al., 2019). Participants followed the instructions on the computer screen. Details are presented elsewhere (Volk et al., 2019). The sequence of the facial movements is shown in **Supplementary Table 1**.

Facial sEMG registration

Surface electromyography measurement was applied using a monopolar montage utilizing reusable surface electrodes (Ag-AgCl discs, diameter of 4 mm, DESS052606, GVB-geliMED, Bad Segeberg Germany). The reference electrodes were positioned at both mastoid processes and connected. The sEMG recording was performed with a multichannel EMG system (gain: 100, frequency range: 10–1,861 Hz; sampling rate: 4,096/s; resolution: 5.96 nV/bit; DeMeTec, Langgöns, Germany). Electromyograms were recorded from both sides of the face simultaneously using both the arrangement of electrodes by Fridlund and Cacioppo (1986) and by Kuramoto et al. (2019) (cf. **Figures 1C–G**). The schemes were labeled as “Fridlund” and “Kuramoto.” The Fridlund scheme is following the topography of the facial muscles. Eleven pairs of electrodes (interelectrode distance for each pair: 1 cm) per hemiface were needed. The corrugator supercilii and depressor supercilii muscles shared one electrode simultaneously. Hence, 21 electrodes were placed on each side of the face. In total, 42 electrodes were needed for the Fridlund scheme. In contrast, the Kuramoto scheme does not take muscle topography into consideration. Similar to an electroencephalogram (EEG), the electrodes were placed in symmetrical standard positions with fixed distances between the electrodes (Beniczky and Schomer, 2020). Eight active electrodes were placed on each side of the face. In addition, three electrodes were placed in the midline of the face. Hence, 19 electrodes were needed in total for the Kuramoto scheme. Six electrodes were shared by the simultaneous application of both schemes (cf. **Figures 1C–G**). One ground electrode and one reference electrode on each mastoid were needed. Therefore, overall 58 electrodes were placed on the face for both the Fridlund and the Kuramoto



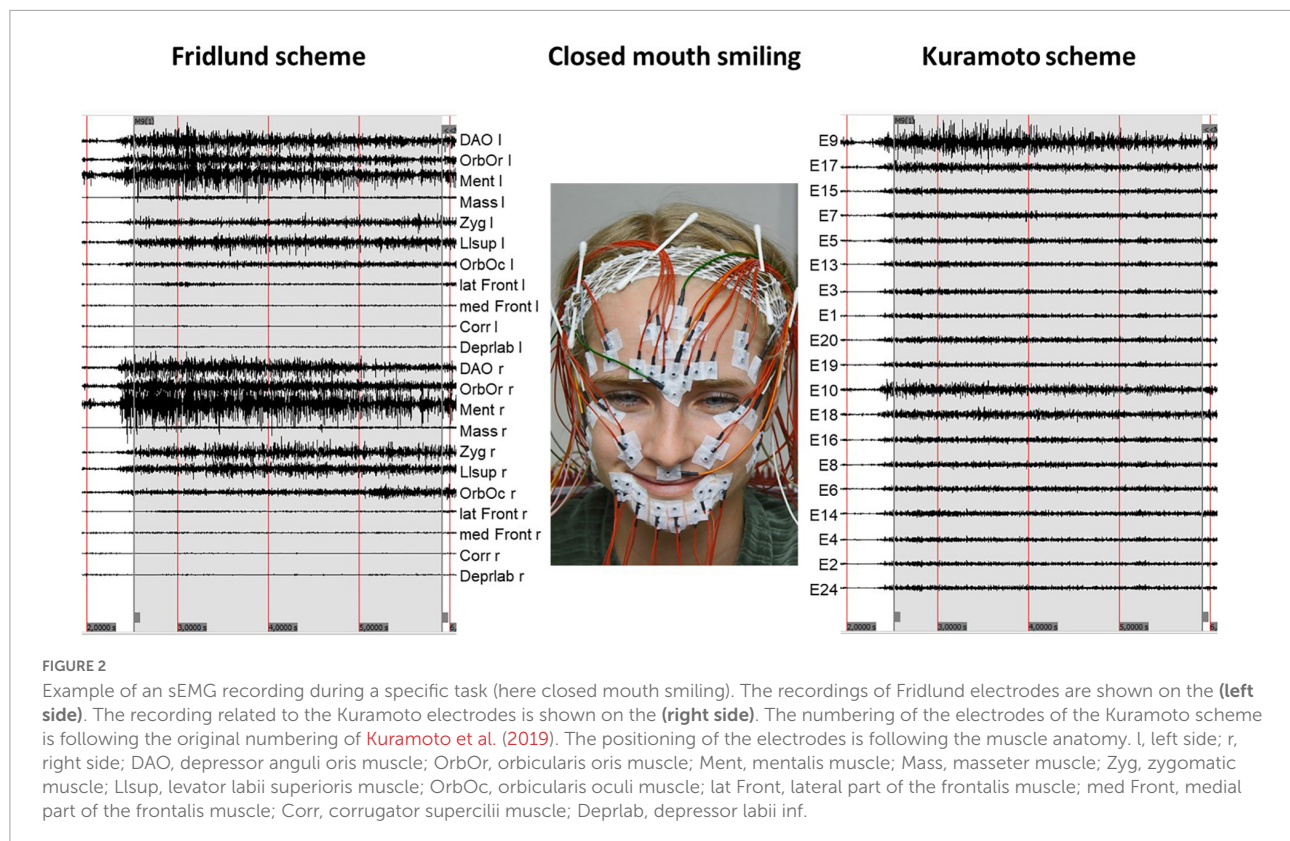
FIGURE 1

The experimental setup. **(A)** The participant is sitting at a standardized distance from the computer screen and is following the instructions. **(B)** The electromyography (EMG) amplifiers are placed in the back of the chair. **(C–E)** EMG electrode schemes after Fridlund and Cacioppo (dots) and after (Kuramoto et al., 2019) (crosses) are marked in the face of the participant with the help of a standardized template (circled crosses refer to electrodes shared by both schemes). **(F)** Frontal view on the fixed surface EMG (sEMG) electrodes. **(G)** Side view on the fixed sEMG electrodes.

scheme and the additional electrodes. An example of an EMG recording for one task is shown in **Figure 2**.

To account for possible technical and also physiological artifacts prior to any further analysis, signals were centered and bandpass filtered between 10 and 500 Hz. Also, a 50 Hz notch filter was applied to account for interferences from the electrical circuit.

For the Fridlund scheme, from the monopolar measured electrodes, bipolar channels were calculated by subtracting the signals from the respective electrode pairs



(Fridlund and Cacioppo, 1986). Data for the Kuramoto scheme were monopolar analyzed. sEMG amplitudes were quantified as mean RMS values (single interval duration for calculation: 125 ms) during the steady state contraction phases of every movement and sEMG channel. Also, maximum-normalized data were calculated to account for individual and movement-dependent amplitude variations and also general amplitude differences between the facial movements.

Heat maps for visualization of the sEMG activity

Heat maps were calculated to provide a topographic impression of the myogenic activity distribution over the face for both schemes. A modified 4-nearest neighbor interpolation of the EMG RMS values was used similar to the EEG mapping with the inverse square of the distance as weight (Jäger et al., 2016). Interpolation in EEG mapping is often performed on the basis of spherical functions because the conducting area in EEG is approximately a hemisphere (Lehmann, 1992). This precondition is not given in the face. Therefore, the 4-nearest neighbor interpolation was used. To avoid spatial discontinuity, the weight of the most distant (fourth) electrode was steadily pulled to zero in the vicinity of the change between two fourth neighbors (Soong et al., 1993).

Statistics

Separately for each scheme, we applied a linear mixed-effects model (LMM) to evaluate the main effects of the parameters “movement,” “side,” and “electrode position,” together with their interactions. “Movement,” “side,” and “electrode position” were modeled as fixed effects with a random intercept per subject. Initially, all main effects and interactions were calculated, but for the final analysis, only the significant main effects together with significant interactions remained in the calculation. Adjustment for multiple comparisons for differences between the tested facial movements was performed by the least significant difference. To differentiate between amplitude-driven and distribution-driven differences, these calculations were applied separately for the RMS and the normalized dataset. The significance level was set to 5%.

Results

Effects of the specific movement, electrode position, and side of the face

Independent of the analyzed scheme and dataset, the initial statistical calculation revealed significant main effects of the parameter “movement” (RMS dataset, Fridlund: $p < 0.001$,

F : 290.5; Kuramoto: $p < 0.001$, F : 364.9; normalized dataset, Fridlund: $p < 0.001$, F : 96.3; Kuramoto: $p < 0.001$, F : 384.3) and “electrode position” (RMS dataset, Fridlund: $p < 0.001$, F : 309.1; Kuramoto: $p < 0.001$, F : 220.7; normalized dataset, Fridlund: $p < 0.001$, F : 370.3; Kuramoto: $p < 0.001$, F : 230.8), but not for “side” in the Fridlund scheme. In the Kuramoto scheme, the parameter “side” had a significant effect in the RMS dataset ($p = 0.027$, F : 4.9). The pairwise comparisons revealed significant differences between both facial sides and the centrally located electrodes (left and right vs. central $p < 0.001$). The values were always lower for the centrally located electrodes. There was no difference between the left and right sides ($p = 0.984$). Hence, the left and right body sides themselves had no significant main effect. Therefore, we decided to further analyze the RMS dataset of the Kuramoto scheme independent of the body side.

Interactions were significant for the parameters “movement*electrode position” (RMS dataset, Fridlund: $p < 0.001$, F : 87.4; Kuramoto: $p < 0.001$, F : 77.9; normalized dataset, Fridlund: $p < 0.001$, F : 104.3; Kuramoto: $p < 0.001$, F : 78.8) and “side*electrode position” (RMS dataset, Fridlund: $p = 0.003$, F : 2.5; Kuramoto: $p < 0.001$, F : 17.2; normalized dataset, Fridlund: $p < 0.001$, F : 16.6; Kuramoto: $p < 0.001$, F : 17.2). In the Kuramoto scheme also, the interaction “movement*side” was a significant parameter ($p = 0.003$, F : 2.7). Based on these results, the different facial movements were then compared with each other.

sEMG recordings with the Fridlund scheme

The heat map presentation of the 11 facial movements is shown in **Figures 3A,B**. Activation characteristics showed systematic regional distribution patterns as could be expected by the facial movements. The variability of these values across the face is provided in an amplitude-normalized fashion, i.e., as variation coefficients (VCs), and are presented in **Figure 3C**. Some facial tasks showed a very low interindividual variability of the activation patterns (for instance, “Wrinkling the forehead”), whereas others showed a higher variability (for instance, “Smiling with open mouth”). Overall, the variability of the activation in μV was very low. In **Figure 4A**, the results of the LMM calculations of the pairwise comparisons between all facial movements are provided for the RMS and normalized datasets. The sEMG recording allowed highly significant discrimination between most of the exercises ($p < 0.0001$) but some failed. The comparisons of the tasks “depress lower lips” vs. “wrinkle nose” and “open mouth smile,” “close eyes forcefully” vs. “wrinkle nose” and “lip pucker,” “wrinkle forehead” vs. “closed mouth smile,” “lip pucker” vs. “close eyes forcefully,” and “rest” vs. “close eyes normally” showed no significant difference (all $p > 0.05$). After normalization, the number of non-significant comparisons dropped, but still the comparisons “closed mouth

smile” vs. “open mouth smile,” “wrinkle forehead,” and “wrinkle nose” and also “close eyes forcefully” vs. “lip pucker” remained non-significant (all $p > 0.05$).

sEMG recording with the Kuramoto scheme

In **Figures 5A,B**, the maps for the Kuramoto scheme of all 11 facial movements and in **Figure 5C**, the respective VC values are provided. As can be seen from **Figure 4B**, the number of highly significant differences between separate and specific facial movements was considerably larger as could be obtained by the application of the Fridlund scheme (cf. **Figure 4A**). The following comparisons of facial movement tasks could not reach significant results: “wrinkle nose” vs. “lip pucker” and “blow-out cheek,” “depress lower lip” vs. “wrinkle forehead” and “closed mouth smile,” “wrinkle forehead” vs. “closed mouth smile,” and finally “rest” vs. “close eyes normally.”

After normalization of the data, nearly all facial exercises revealed highly specific sEMG activity patterns ($p < 0.0001$). Only “wrinkle nose” vs. “wrinkle forehead,” “blow-out cheek” vs. “snarl,” and “depress lower lip” vs. “open mouth smile” failed to reach the required significance level (all $p > 0.05$).

Discussion

The present study was the first prospective clinical trial directly and synchronously comparing the facial sEMG electrode arrangements of Fridlund and Cacioppo vs. the new arrangement of Kuramoto et al. in the same group of healthy probands (Fridlund and Cacioppo, 1986; Kuramoto et al., 2019). Overall, both sEMG schemes allowed to record of characteristic and discriminable facial muscle activation patterns during different facial movement tasks. The geometric sEMG recording from the entire face introduced by Kuramoto et al. seemed to allow more specific detection of facial muscle activity patterns during facial movement tasks (Kuramoto et al., 2019). As shown by others, there was no relevant side difference (Schumann et al., 2010, 2021; Kuramoto et al., 2019; Cui et al., 2020). In contrast, HD sEMG recordings of the facial and neck muscles show mainly a symmetrical activation but also some side differences during phonation tasks (Zhu et al., 2022). Hence, it might be that the degree of left/right symmetry might be task-dependent.

Since its publication in 1986, the recommendations of Fridlund and Cacioppo for sEMG electrode placements for a differential recording over the most important facial muscles [Figure 1 in Fridlund and Cacioppo (1986)] probably become the most popular facial recording scheme. More than 1,200 articles (Web of Science, August 2022, mainly psychological research) refer to this scheme when describing their facial sEMG approach. It is important to notice that the Fridlund scheme

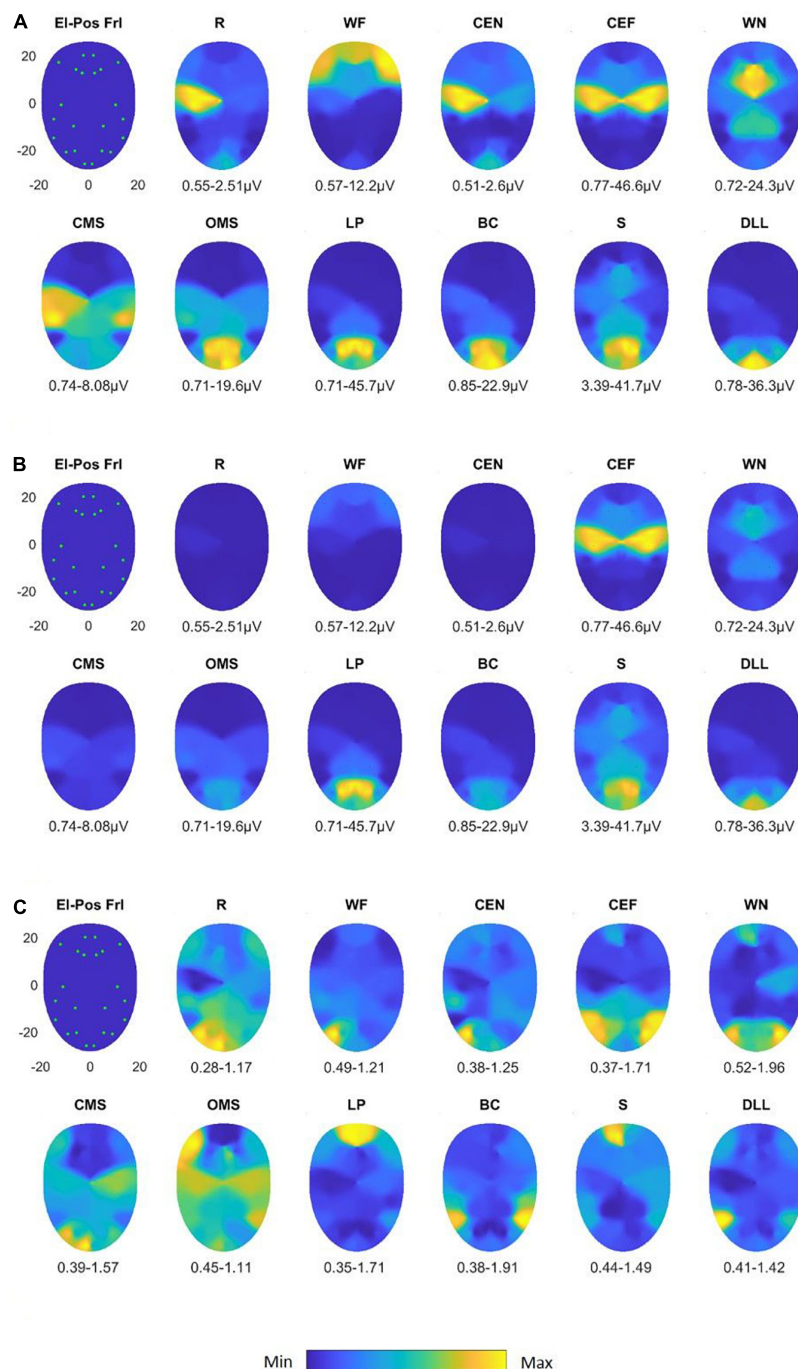


FIGURE 3

Activation heat maps of facial muscle activity during different mimic exercises using the Fridlund scheme (El-Pos Fri). According to the provided color bars, blue stands for low values and yellow for high values. Below each map, the respective minimal (Min) and maximal (Max) values are provided. **(A)** Heat maps with always individual Min to Max scaling; **(B)** heat maps scaled according to the Min and Max values across all facial movements. **(C)** Variability (variation coefficient) heat maps. R, at rest; WF, wrinkling of the forehead; CEF, closing the eyes normally; CEF, closing the eyes forcefully; WN, wrinkling of the nose; CMS, closed mouth smiling; OMS, open mouth smiling; LP, lip puckering; BC, blowing-out the cheeks; S, snarling; DLL, depressing lower lip.

covers nine facial muscles (called by Fridlund and Cacioppo as “major” facial muscles) and not all facial muscles. Second, many articles cite the Fridlund scheme, even when the recordings

are just performed from some of the nine facial muscles. As shown in the present study, the Fridlund scheme is a powerful tool when placing all 22 electrode pairs. We have doubts,

A											
Fridlund scheme											
rms	WF	CEN	CEF	WN	CMS	OMS	LP	BC	S	DLL	
R	<0.001	0.888	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
WF		<0.001	<0.001	<0.001	0.182	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
CEN			<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
CEF				0.051	<0.001	0.004	0.062	<0.001	<0.001	<0.001	<0.001
WN	n.s.	7	12.7%		<0.001	0.341	<0.001	<0.001	<0.001	<0.001	0.070
CMS	<0.05	1	1.8%			<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
OMS							<0.001	0.001	<0.001	0.389	
LP	<0.01	2	3.6%					0.000	<0.001	<0.001	<0.001
BC	<0.001	45	81.8%						<0.001	0.017	
S										<0.001	
normalized	WF	CEN	CEF	WN	CMS	OMS	LP	BC	S	DLL	
R	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
WF		<0.001	<0.001	0.045	0.624	0.045	<0.001	<0.001	0.001	<0.001	<0.001
CEN			<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.025	<0.001	<0.001
CEF				<0.001	<0.001	<0.001	0.987	<0.001	<0.001	0.029	<0.001
WN	n.s.	5	9.1%		0.130	1.000	<0.001	<0.001	<0.001	<0.001	<0.001
CMS	<0.05	5	9.1%			0.130	<0.001	<0.001	<0.001	<0.001	<0.001
OMS							<0.001	<0.001	<0.001	<0.001	<0.001
LP	<0.01	1	1.8%					<0.001	<0.001	0.030	<0.001
BC	<0.001	44	80.0%						<0.001	<0.001	<0.001
S										<0.001	

B											
Kuramoto scheme											
rms	WF	CEN	CEF	WN	CMS	OMS	LP	BC	S	DLL	
R	<0.001	0.536	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
WF		<0.001	<0.001	<0.001	0.079	<0.001	<0.001	<0.001	<0.001	<0.001	0.054
CEN			<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
CEF				<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
WN	n.s.	7	12.7%		<0.001	0.003	0.103	0.343	<0.001	<0.001	<0.001
CMS	<0.05	0	0.0%			<0.001	<0.001	<0.001	<0.001	<0.001	0.863
OMS							<0.001	<0.001	<0.001	<0.001	<0.001
LP	<0.01	1	1.8%					0.494	<0.001	<0.001	<0.001
BC	<0.001	47	85.5%						<0.001	<0.001	<0.001
S										<0.001	<0.001
normalized	WF	CEN	CEF	WN	CMS	OMS	LP	BC	S	DLL	
R	<0.001	0.007	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
WF		<0.001	<0.001	<0.001	<0.001	<0.001	0.000	<0.001	<0.001	<0.001	<0.001
CEN			<0.001	<0.001	<0.001	<0.001	0.000	<0.001	<0.001	<0.001	<0.001
CEF				<0.001	<0.001	<0.001	0.000	<0.001	<0.001	<0.001	<0.001
WN	n.s.	3	5.5%		<0.001	<0.001	0.000	<0.001	<0.001	<0.001	<0.001
CMS	<0.05	0	0.0%			<0.001	0.000	<0.001	<0.001	<0.001	<0.001
OMS							0.000	0.002	0.001	0.949	<0.001
LP	<0.01	5	9.1%					<0.001	<0.001	<0.001	<0.001
BC	<0.001	47	85.5%						0.867	0.002	<0.001
S										0.001	

FIGURE 4

Results of the between-movement comparisons by applying a linear mixed-effects model. (A) Fridlund scheme. (B) Kuramoto scheme. Upper row for both schemes: row data; lower row: normalized data. Significance levels are color-coded: $p < 0.05$: pink, $p < 0.001$: orange, $p < 0.001$: red. Also, the absolute numbers and relative numbers of non-significant and significant differences are provided. R, at rest; WF, wrinkling of the forehead; CEF, closing the eyes normally; CEF, closing the eyes forcefully; WN, wrinkling of the nose; CMS, closed mouth smiling; OMS, open mouth smiling; LP, lip puckering; BC, blowing-out the cheeks; S, snarling; DLL, depressing lower lip; n.s., not significant.

and that was not the aim of Fridlund and Cacioppo, that a reduced scheme does allow to control of specific facial muscle activations during mimetic investigations. The accuracy of an EMG activation pattern to reflect a specific mimetic task is not only important for clinical or psychosocial studies. There are first approaches to use specific human facial gestures based on EMG data for human-machine interface (HMI) technologies for steering of machines (Hamed et al., 2011). This will only work when reliable sEMG facial muscle patterns are used.

In this light, it is important to bring to mind that although facial muscles are conventionally classified as individual muscles, these muscles are interconnected, and movements are interdependent on each other (Cattaneo and Pavesi, 2014). Even more, the activation of one muscle may define the pulling direction of another facial muscle just in this actual moment. Therefore, most facial and emotional expressions

need the coordinated action of several facial muscles (cf. Figure 2: not only muscles around the mouth are activated for smiling). Furthermore, the correct identification of all respective facial muscles requires highly trained personnel and is time-consuming to produce reliable results. Also, the post-processing (monopolar recording has to be converted into bipolar channels) bears some risks. For this reason, the Kuramoto scheme is of such interest. For this approach, called by the authors “myogenic potential topography,” the electrodes are placed on the face with fixed inter-electrode distance covering the entire face without reference to the underlying facial muscles. In the original study, the probands had to perform two mimic expressions (smiling and a disgusting face) different from the present study. Therefore, a comparison of the results is not possible. Using much more movement tasks, we were able to show that the Kuramoto scheme allows

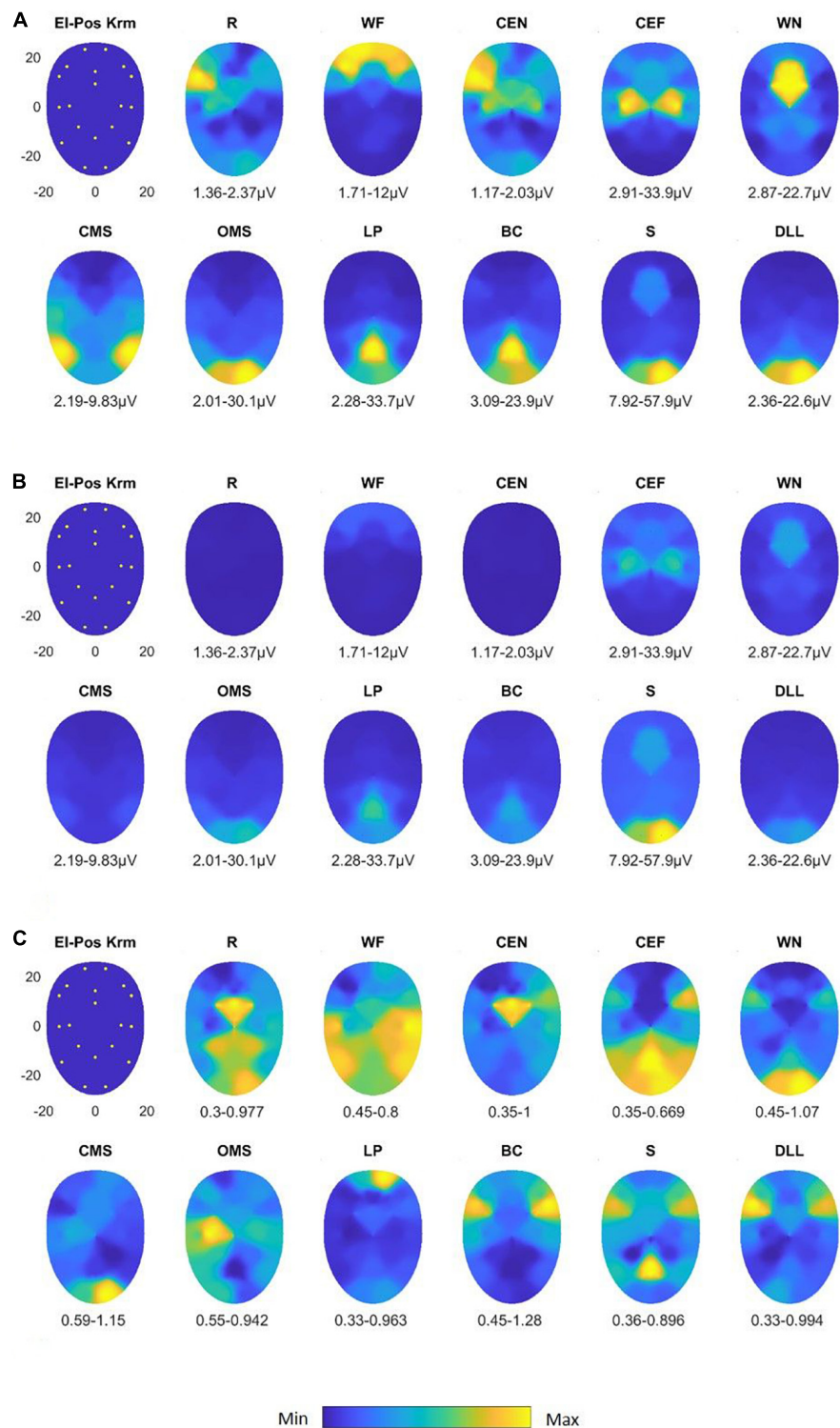


FIGURE 5

Activation heat maps of facial muscle activity during different mimic exercises using the Kuramoto scheme (El-Pos Krm). According to the provided color bars, blue stands for low values and yellow for high values. Below each map, the respective minimal (Min) and maximal (Max) are provided. **(A)** Heat maps with always individual Min to Max scaling; **(B)** heat maps scaled according to the Min and Max values across all facial movements. **(C)** Variability (variation coefficient) heat maps. R, at rest; WF, wrinkling of the forehead; CEF, closing the eyes normally; CEF, closing the eyes forcefully; WN, wrinkling of the nose; CMS, closed mouth smiling; OMS, open mouth smiling; LP, lip puckering; BC, blowing-out the cheeks; S, snarling; DLL, depressing lower lip.

discrimination of the most important (non-emotional) facial movements. Cui et al. used six facial movements (raising eyebrows, closing eyes, bulging cheek, grinning, pouting, and wrinkling the nose) in combination with a facial high-density sEMG (HD sEMG) to produce center of gravity color-coded activation maps but finally refer the activation to single facial muscles (Cui et al., 2020).

The present study has some limitations. We analyzed only facial movement tasks. In the next experiments, we would like to compare both schemes when the probands perform emotional expressions. Furthermore, the differences between the reliability of the individual activation patterns for both sEMG schemes are unknown. In the present study, there was some interindividual variability in the muscle activation patterns. Facial muscle size shows some interindividual variability (Volk et al., 2014). This, but also age and gender, might influence activity patterns. The influence of these factors should be studied in larger cohorts of healthy adults. Furthermore, the performance of the tasks might vary if performed repeatedly on different days (Hess et al., 2017). Therefore, the retest reliability has to be investigated in future studies. Cui et al. used it for their HD sEMG work-up of 90 electrodes (Cui et al., 2020). In contrast to the presented schemes, it remains doubtful if such an elaborate experimental setup is applicable to clinical routine. It might be more interesting to develop easy-to-apply wearable, multichannel sEMG electrodes (Inzelberg et al., 2018, 2020; Gat et al., 2022). This would simplify routine application and would further allow investigations with freely moving probands, hence, would allow a more realistic setting, especially for psychosocial experiments.

Conclusion

The high-resolution sEMG recording of healthy probands shows that specific facial movement tasks lead to complex facial muscle activation patterns. These patterns can only be detected if the entire facial muscles on both sides of the face are recorded simultaneously. Specific facial movement tasks lead to specific sEMG activation patterns. The discrimination seems to be more sensitive if an sEMG recording scheme with the geometrical arrangement (as typical for EEG) and independent of any facial muscle topography is used. Nevertheless, a classical sEMG recording scheme that is orientated to the topography of facial muscles produces also specific sEMG activation patterns for most facial movement tasks.

Data availability statement

The original contributions presented in this study are included in the article/**Supplementary material**, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of the Jena University Hospital. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual for the publication of any potentially identifiable images or data included in this article.

Author contributions

OG-L, GV, and CA: conceptualization and supervision. OG-L and CA: first draft preparation. NM and VT: data acquisition. NM, VT, RG, and CA: data analysis. All authors contributed to the article and approved the final version.

Funding

Both NM and VT received a doctoral scholarship from the Interdisziplinäres Zentrum für Klinische Forschung (IZKF) of the Jena University Hospital. OG-L acknowledges support by the Deutsche Forschungsgemeinschaft (DFG), grant no. GU-463/12-1.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnhum.2022.1029415/full#supplementary-material>

References

- Beniczky, S., and Schomer, D. (2020). Electroencephalography: Basic biophysical and technological aspects important for clinical applications. *Epileptic Disord.* 22, 697–715. doi: 10.1684/epd.2020.1217
- Cattaneo, L., and Pavesi, G. (2014). The facial motor system. *Neurosci. Biobehav. Rev.* 38, 135–159. doi: 10.1016/j.neubiorev.2013.11.002
- Cui, H., Zhong, W., Yang, Z., Cao, X., Dai, S., Huang, X., et al. (2020). Comparison of facial muscle activation patterns between healthy and bell's palsy subjects using high-density surface electromyography. *Front. Hum. Neurosci.* 14:618985. doi: 10.3389/fnhum.2020.618985
- Fridlund, A., and Cacioppo, J. (1986). Guidelines for human electromyographic research. *Psychophysiology* 23, 567–589. doi: 10.1111/j.1469-8986.1986.tb00676.x
- Gat, L., Gerston, A., Shikun, L., Inzelberg, L., and Hanein, Y. (2022). Similarities and disparities between visual analysis and high-resolution electromyography of facial expressions. *PLoS One* 17:e0262286. doi: 10.1371/journal.pone.0262286
- Guntinas-Lichius, O., Volk, G., Olsen, K., Makitie, A., Silver, C., Zafereo, M., et al. (2020). Facial nerve electrodiagnostics for patients with facial palsy: A clinical practice guideline. *Eur. Arch. Otorhinolaryngol.* 277, 1855–1874. doi: 10.1007/s00405-020-05949-1
- Hamed, M., Salleh, S. H., Tan, T. S., Ismail, K., Ali, J., Dee-Uam, C., et al. (2011). Human facial neural activities and gesture recognition for machine-interfacing applications. *Int. J. Nanomed.* 6, 3461–3472. doi: 10.2147/IJN.S26619
- Hess, U., Arslan, R., Mauersberger, H., Blaison, C., Dufner, M., Denissen, J., et al. (2017). Reliability of surface facial electromyography. *Psychophysiology* 54, 12–23. doi: 10.1111/psyp.12676
- Hubert, W., and de Jong-Meyer, R. (1991). Psychophysiological response patterns to positive and negative film stimuli. *Biol. Psychol.* 31, 73–93. doi: 10.1016/0301-0511(90)90079-c
- Inzelberg, L., David-Pur, M., Gur, E., and Hanein, Y. (2020). Multi-channel electromyography-based mapping of spontaneous smiles. *J. Neural Eng.* 17, 026025. doi: 10.1088/1741-2552/ab7c18
- Inzelberg, L., Rand, D., Steinberg, S., David-Pur, M., and Hanein, Y. A. (2018). Wearable high-resolution facial electromyography for long term recordings in freely behaving humans. *Sci. Rep.* 8:2058. doi: 10.1038/s41598-018-20567-y
- Jäger, J., Klein, A., Buhmann, M., and Skrandies, W. (2016). Reconstruction of electroencephalographic data using radial basis functions. *Clin. Neurophysiol.* 127, 1978–1983. doi: 10.1016/j.clinph.2016.01.003
- Kuramoto, E., Yoshinaga, S., Nakao, H., Nemoto, S., and Ishida, Y. (2019). Characteristics of facial muscle activity during voluntary facial expressions: Imaging analysis of facial expressions based on myogenic potential data. *Neuropsychopharmacol. Rep.* 39, 183–193. doi: 10.1002/npr2.12059
- Lehmann, D. (1992). [Evaluation of evoked potential or event-related potential mapping]. *EEG EMG Z. Elektroenzephalogr. Elektromyogr. Verwandte Geb* 23, 1–11.
- Schaefer, R., Volk, G., Modersohn, L., Barth, J., Denzler, J., and Guntinas-Lichius, O. (2017). [Video instruction for synchronous video recording of mimic movement of patients with facial palsy]. *Laryngorhinotologie* 96, 844–849. doi: 10.1055/s-0043-101699
- Schumann, N., Bongers, K., Guntinas-Lichius, O., and Scholle, H. (2010). Facial muscle activation patterns in healthy male humans: A multi-channel surface emg study. *J. Neurosci. Methods* 187, 120–128. doi: 10.1016/j.jneumeth.2009.12.019
- Schumann, N., Bongers, K., Scholle, H., and Guntinas-Lichius, O. (2021). Atlas of voluntary facial muscle activation: Visualization of surface electromyographic activities of facial muscles during mimic exercises. *PLoS One* 16:e0254932. doi: 10.1371/journal.pone.0254932
- Soong, A., Lind, J., Shaw, G., and Koles, Z. (1993). Systematic comparisons of interpolation techniques in topographic brain mapping. *Electroencephal. Clin. Neurophysiol.* 87, 185–195. doi: 10.1016/0013-4694(93)90018-q
- Volk, G., Sauer, M., Pohlmann, M., and Guntinas-Lichius, O. (2014). Reference values for dynamic facial muscle muscle ultrasonography in adults. *Muscle Nerve* 50, 348–357. doi: 10.1002/mus.24204
- Volk, G., Schaefer, R., Thielker, J., Modersohn, L., Mothes, O., Nduka, C., et al. (2019). Reliability of grading of facial palsy using a video tutorial with synchronous video recording. *Laryngoscope* 129, 2274–2279. doi: 10.1002/lary.27739
- Zhu, M., Wang, X., Deng, H., He, Y., Zhang, H., Liu, Z., et al. (2022). Towards evaluating pitch-related phonation function in speech communication using high-density surface electromyography. *Front. Neurosci.* 16:941594. doi: 10.3389/fnhum.2022.941594

Frontiers in Human Neuroscience

Bridges neuroscience and psychology to
understand the human brain

The second most-cited journal in the field of
psychology, that bridges research in psychology
and neuroscience to advance our understanding
of the human brain in both healthy and diseased
states.

Discover the latest Research Topics

[See more →](#)

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact



Frontiers in Human Neuroscience

