frontiers | Research Topics

New insights in the cognitive neuroscience of attention

Edited by
Tetsuo Kida and Hidehiko Okamoto

Published in
Frontiers in Human Neuroscience

ISSN 1664-8714
ISBN 978-2-8325-4109-8
DOI 10.3389/978-2-8325-4109-8

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

New insights in the cognitive neuroscience of attention

Topic editors

Tetsuo Kida — Aichi Developmental Disability Center, Institute for Developmental Research, Japan

Hidehiko Okamoto — International University of Health and Welfare (IUHW), Japan

Citation

Kida, T., Okamoto, H., eds. (2024). *New insights in the cognitive neuroscience of attention*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-4109-8

Table of contents

frontiers | Frontiers in Human Neuroscience

TYPE Editorial
PUBLISHED 29 November 2023
DOI 10.3389/fnhum.2023.1338738

OPEN ACCESS

EDITED AND REVIEWED BY
Lutz Jäncke,
University of Zurich, Switzerland

*CORRESPONDENCE
Tetsuo Kida
✉ tkida@inst-hsc.jp

RECEIVED 15 November 2023
ACCEPTED 15 November 2023
PUBLISHED 29 November 2023

CITATION
Kida T and Okamoto H (2023) Editorial: New insights into the cognitive neuroscience of attention. *Front. Hum. Neurosci.* 17:1338738. doi: 10.3389/fnhum.2023.1338738

Editorial: New insights into the cognitive neuroscience of attention

Tetsuo Kida[1]* and Hidehiko Okamoto[2]

[1]Higher Brain Function Unit, Department of Functioning and Disability, Aichi Developmental Disability Center, Institute for Developmental Research, Kasugai, Japan, [2]Department of Physiology, International University of Health and Welfare Graduate School of Medicine, Narita, Japan

KEYWORDS

attention, neuroimaging, electroencephalography, magnetoencephalography, executive functions, working memory, oscillatory activity, near infrared spectroscopy (NIRS)

Editorial on the Research Topic
New insights into the cognitive neuroscience of attention

Attention is associated with the selective directedness of our mental lives. It has been extensively examined since the nineteenth century, and recent studies have elucidated the underlying neural mechanisms and functionality in living humans using various human neuroscientific techniques. This Research Topic is a collection of studies relevant to attention in health and disease using human neuroscientific techniques, including event-related brain potentials (ERPs) and magnetic fields (ERFs), oscillatory activity on magnetoencephalograms (MEG) or electroencephalograms (EEG), sheet-type wearable EEG devices, functional magnetic resonance imaging (fMRI), functional near-infrared spectroscopy (fNIRS), and novel methods for the measurement and modeling of attention.

EEG is a classical technique to measure human neurophysiological activity, and recent advances in technology and experimental paradigms have rendered EEG a unique and powerful tool (Michel and Murray, 2012; Kida et al., 2015). This Research Topic consists of a number of EEG-based studies; five are based on the ERP technique, which has long been used to elucidate the neural mechanisms underlying attention (Hillyard et al., 1998; Martinez et al., 1999; Eimer and Driver, 2001; Muller et al., 2003; Eimer et al., 2009; Jones and Forster, 2013; Forster et al., 2016; Gherri et al., 2016, 2023; Qin et al., 2022). Using ERPs, Rutkowska et al. investigated a working memory (WM)-based guidance effect for naturalistic, complex stimuli. WM-based guidance is the automatic capture of visual attention generated when a sensory stimulus matches a stimulus representation voluntarily stored in WM. Memorized items did not evoke a lateralized N2pc ERP component, which was considered to indicate attention shifts (Luck and Hillyard, 1994; Eimer, 1996; Eimer and Kiss, 2007). In contrast, ERP modulation for very early prioritization specific to memorized faces was noted, which was consistent with the sensory recruitment theory of WM. Therefore, complex stimuli were prioritized by attention when maintained in WM, and the mechanism underlying this prioritization was considered to be based on a prolonged hold of spatial attention. Gherri et al. identified a reliable N140cc component (Forster et al., 2016), a somatosensory homolog of N2pc, for all set sizes in a tactile search task, and the N140cc amplitude decreased as the number of distractors increased. The addition of distractors appeared to impede the preattentive processing of the search array, resulting in increased uncertainty about the target location. This, in turn, increased the variability of directing attention to the

target, resulting in reduced N140cc amplitudes. These findings indicated distinct patterns of attention between touch and vision. A study by Kimura and Katayama used ERP as an index of attention and prediction error and event-related spectral perturbation as an index of prediction to elucidate the predictive function of the visuo–tactile interaction in the peri-personal space. Their findings showed that approaching visual stimuli facilitated the prediction of subsequent tactile events regardless of whether they were relevant to the task. An ERP study by Kida et al. demonstrated the effects of focused and divided attention on N140 and P300 in somatosensory selective attention tasks with different response modes (motor response and silent counting), and also attentional modulation under different hand and arm postures. They also revealed attentional alterations in stimulus-response coupling using a single-trial analysis of the P300 latency and reaction time, which showed different patterns from data for dual-task vs. single-task conditions (Kida et al., 2012). An auditory ERP study by Sasaki et al. showed that the amplitude of phonological mapping negativity and N400, which are objective measures of phonological and semantic processing, respectively, increased when the picture of an event and speech mismatched. The authors suggested that established sound symbolic words and sound symbolic pseudowords undergo similar semantic processing. The ERP technique has long been used to elucidate the mechanisms underlying attention, and the findings published here indicate that this technique is still effective for that purpose.

Other studies published here using EEG proved its diverse usefulness for examining and estimating attentional states. Wan et al. measured EEG during the performance of a sustained attention task and in a resting state. A new classification model, Complexity-XGBoost, was proposed using EEG-based dynamical complexity features and Extreme Gradient Boosting (XGBoost) to discriminate multi-level attention states with greater accuracy. The proposed model outperformed the other classification methods, leading the authors to conclude the value of the proposed approach for classifying the attention status in order to improve safety and efficiency, and also its applicability to the brain-computer interaction. Dini et al. measured EEG while participants watched videos with high and low levels of narrativity, and calculated the inter-subject correlation of EEG and attentional engagement scores. High-level narrativity was associated with higher inter-subject correlation and attention scores, suggesting that the findings obtained are a step toward elucidating the viewers' way of processing. Callan et al. measured EEG gamma and alpha oscillations in association with inattentional deafness using a wireless EEG device. In an auditory detection task, task performance (hits and misses) was associated with alpha- and gamma-band activities, respectively, in different hemispheres of the auditory cortex. Additional activities were detected in the frontal and parietal regions that reflected attentional monitoring, selection, and switching. Therefore, these findings showed the involvement of gamma and alpha activities in frontal and modality-specific regions in selective attention. Ueno et al. used a sheet-type wearable EEG device to measure the frontal midline theta rhythm (fmθ), which emerges during the attentional focus state, in a resting state and mental calculation task. Participants with fmθ showed higher gamma and theta power and lower alpha power, leading the authors to conclude that gamma activity in the frontal region was associated with WM.

This Research Topic was also contributed to by other neuroscientific techniques, such as MEG, fNIRS, genetics neuroimaging, and behavioral measures. EEG is a record of extracellular volume currents (secondary current) that originate from the current source and flow in the brain, whereas MEG is a record of the in/outflux of magnetic fields generated following intracellular currents (primary current; Hamalainen et al., 1993; Kida et al., 2015). The magnetic field is highly permeable in the skull, scalp, meninges, and cerebrospinal fluid and, thus, MEG has advantages for accuracy and the computational demand of source localization. MEG studies reported attentional modulations using event-related (evoked) activity in audition (Okamoto et al., 2007), vision (Kida et al., 2011), and touch (Kida et al., 2018) as well as oscillatory activity (Siegel et al., 2008). In this Research Topic, Onishi and Yokosawa used MEG to examine theta-band oscillatory activity during a WM task. Theta activity exhibited a dynamic pattern from attentional control and inhibition control in the prefrontal cortex to the inhibition of task-irrelevant information in the occipito-parietal cortex. An fNIRS study by Ohshima et al. examined the neural correlates of executive functions (EFs) based on an emotional vocal cue in an N-back task. The findings obtained showed that the right precentral and inferior frontal gyri were activated during the emotional N-back task, reflecting the function of an attentional network with auditory top-down processing. Significant activation was also detected in the ventrolateral prefrontal cortex, an important area for WM. Therefore, this study demonstrated the neural correlate of emotional judgments based on vocal cues in comparison with that for gender judgments. Huang et al. reported genetics neuroimaging in attention deficit/hyperactivity disorder (ADHD), one of the most common neurodevelopmental disorders. The study performed a functional connectivity (FC) analysis of resting-state fMRI data in 60 ADHD patients and 28 healthy controls, and genotyped single nucleotide polymorphisms in synaptosomal-associated protein-25 (SNAP-25) to separate patients into two groups: TT homozygotes and G-allele carriers. In comparisons with the TT group, the TG group showed significant changes in FC between the hippocampus and other regions. The authors proposed that hippocampal FC may serve as an imaging biomarker for ADHD. Younger et al. developed a novel behavioral measure and modeling of EFs closely associated with attentional functions, and used a large-scale diverse dataset in middle childhood to reveal the evolving diversity of EFs, including WM, context monitoring, and interference resolution, while simultaneously accounting for their unity. The findings obtained showed that the eight tasks were organized into three stable components by the age of 10 years, whereas refinements in the composition of these components continued through to at least the age of 14 years.

The novel analytical approaches, measurement devices, ERP components, and oscillatory activities discussed in this Research Topic will advance research on the neuroscience of attention. The variety of methodologies and themes represents the expansion of this research field, suggesting that the cognitive neuroscience of attention will continue to progress in the future.

Author contributions

TK: Conceptualization, Writing – original draft. HO: Conceptualization, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

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.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

References

Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalogr. Clin. Neurophysiol.* 99, 225–234. doi: 10.1016/0013-4694(96)95711-9

Eimer, M., and Driver, J. (2001). Crossmodal links in endogenous and exogenous spatial attention: evidence from event-related brain potential studies. *Neurosci. Biobehav. Rev.* 25, 497–511. doi: 10.1016/S0149-7634(01)00029-X

Eimer, M., and Kiss, M. (2007). Attentional capture by task-irrelevant fearful faces is revealed by the N2pc component. *Biol. Psychol.* 74, 108–112. doi: 10.1016/j.biopsycho.2006.06.008

Eimer, M., Kiss, M., Press, C., and Sauter, D. (2009). The roles of feature-specific task set and bottom-up salience in attentional capture: an ERP study. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 1316–1328. doi: 10.1037/a0015872

Forster, B., Tziraki, M., and Jones, A. (2016). The attentive homunculus: ERP evidence for somatotopic allocation of attention in tactile search. *Neuropsychologia* 84, 158–166. doi: 10.1016/j.neuropsychologia.2016.02.009

Gherri, E., Gooray, E., and Forster, B. (2016). Cue-locked lateralized components in a tactile spatial attention task: evidence for a functional dissociation between ADAN and LSN. *Psychophysiology* 53, 507–517. doi: 10.1111/psyp.12596

Gherri, E., White, F., and Venables, E. (2023). On the spread of spatial attention in touch: evidence from event-related brain potentials. *Biol. Psychol.* 178:108544. doi: 10.1016/j.biopsycho.2023.108544

Hamalainen, M., Hari, R., Ilmoniemi, R. J., Kunuutila, J., and Lounasmaa, O. V. (1993). Magnetoencephalography-theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev. Mod. Phys.* 65, 413–497. doi: 10.1103/RevModPhys.65.413

Hillyard, S. A., Vogel, E. K., and Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 353, 1257–1270. doi: 10.1098/rstb.1998.0281

Jones, A., and Forster, B. (2013). Lost in vision: ERP correlates of exogenous tactile attention when engaging in a visual task. *Neuropsychologia* 51, 675–685. doi: 10.1016/j.neuropsychologia.2013.01.010

Kida, T., Inui, K., Tanaka, E., and Kakigi, R. (2011). Dynamics of within-, inter-, and cross-modal attentional modulation. *J. Neurophysiol.* 105, 674–686. doi: 10.1152/jn.00807.2009

Kida, T., Kaneda, T., and Nishihira, Y. (2012). Modulation of somatosensory processing in dual tasks: an event-related brain potential study. *Exp. Brain Res.* 216, 575–584. doi: 10.1007/s00221-011-2961-z

Kida, T., Tanaka, E., and Kakigi, R. (2015). Multi-dimensional dynamics of human electromagnetic brain activity. *Front. Hum. Neurosci.* 9:713. doi: 10.3389/fnhum.2015.00713

Kida, T., Tanaka, E., and Kakigi, R. (2018). Adaptive flexibility of the within-hand attentional gradient in touch: an MEG study. *Neuroimage* 179, 373–384. doi: 10.1016/j.neuroimage.2018.06.063

Luck, S. J., and Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology* 31, 291–308. doi: 10.1111/j.1469-8986.1994.tb02218.x

Martinez, A., Anllo-Vento, L., Sereno, M. I., Frank, L. R., Buxton, R. B., Dubowitz, D. J., et al. (1999). Involvement of striate and extrastriate visual cortical areas in spatial attention. *Nat. Neurosci.* 2, 364–369. doi: 10.1038/7274

Michel, C. M., and Murray, M. M. (2012). Towards the utilization of EEG as a brain imaging tool. *Neuroimage* 61, 371–385. doi: 10.1016/j.neuroimage.2011.12.039

Muller, M. M., Malinowski, P., Gruber, T., and Hillyard, S. A. (2003). Sustained division of the attentional spotlight. *Nature* 424, 309–312. doi: 10.1038/nature01812

Okamoto, H., Stracke, H., Wolters, C. H., Schmael, F., and Pantev, C. (2007). Attention improves population-level frequency tuning in human auditory cortex. *J. Neurosci.* 27, 10383–10390. doi: 10.1523/JNEUROSCI.2963-07.2007

Qin, N., Wiens, S., Rauss, K., and Pourtois, G. (2022). Effects of selective attention on the C1 ERP component: a systematic review and meta-analysis. *Psychophysiology* 59, e14123. doi: 10.1111/psyp.14123

Siegel, M., Donner, T. H., Oostenveld, R., Fries, P., and Engel, A. K. (2008). Neuronal synchronization along the dorsal visual pathway reflects the focus of spatial attention. *Neuron.* 60, 709–719. doi: 10.1016/j.neuron.2008.09.010

frontiers in Human Neuroscience

ORIGINAL RESEARCH
published: 01 June 2021
doi: 10.3389/fnhum.2021.673955

Frontal EEG-Based Multi-Level Attention States Recognition Using Dynamical Complexity and Extreme Gradient Boosting

Wang Wan[1†], Xingran Cui[2,3†], Zhilin Gao[1] and Zhongze Gu[1,3*]*

[1] State Key Laboratory of Bioelectronics, School of Biological Science & Medical Engineering, Southeast University, Nanjing, China, [2] Key Laboratory of Child Development and Learning Science, Ministry of Education, School of Biological Science & Medical Engineering, Southeast University, Nanjing, China, [3] Institute of Biomedical Devices (Suzhou), Southeast University, Suzhou, China

OPEN ACCESS

Edited by:
Jessica A. Turner,
Georgia State University,
United States

Reviewed by:
Aaron Kemp,
University of Arkansas for Medical Sciences, United States
Márk Molnár,
Hungarian Academy of Sciences (MTA), Hungary

***Correspondence:**
Xingran Cui
cuixr@seu.edu.cn
Zhongze Gu
gu@seu.edu.cn

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

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

Received: 28 February 2021
Accepted: 26 April 2021
Published: 01 June 2021

Citation:
Wan W, Cui X, Gao Z and Gu Z (2021) Frontal EEG-Based Multi-Level Attention States Recognition Using Dynamical Complexity and Extreme Gradient Boosting. Front. Hum. Neurosci. 15:673955. doi: 10.3389/fnhum.2021.673955

Measuring and identifying the specific level of sustained attention during continuous tasks is essential in many applications, especially for avoiding the terrible consequences caused by reduced attention of people with special tasks. To this end, we recorded EEG signals from 42 subjects during the performance of a sustained attention task and obtained resting state and three levels of attentional states using the calibrated response time. EEG-based dynamical complexity features and Extreme Gradient Boosting (XGBoost) classifier were proposed as the classification model, Complexity-XGBoost, to distinguish multi-level attention states with improved accuracy. The maximum average accuracy of Complexity-XGBoost were 81.39 ± 1.47% for four attention levels, 80.42 ± 0.84% for three attention levels, and 95.36 ± 2.31% for two attention levels in 5-fold cross-validation. The proposed method is compared with other models of traditional EEG features and different classification algorithms, the results confirmed the effectiveness of the proposed method. We also found that the frontal EEG dynamical complexity measures were related to the changing process of response during sustained attention task. The proposed dynamical complexity approach could be helpful to recognize attention status during important tasks to improve safety and efficiency, and be useful for further brain-computer interaction research in clinical research or daily practice, such as the cognitive assessment or neural feedback treatment of individuals with attention deficit hyperactivity disorders, Alzheimer's disease, and other diseases which affect the sustained attention function.

Keywords: attention recognition, sustained attention task, electroencephalogram, dynamical complexity, extreme gradient boosting

INTRODUCTION

Sustained attention refers to the ability to focus on task-related information stimuli while consciously trying to ignore other stimuli over a relatively long period. As a foundational cognitive function, sustained attention underlies other cognitive domains, such as learning and memory (Fortenbaugh et al., 2017). However, due to the lack of a monitoring mechanism, assessing whether

people are sustainably focused and improving concentration in learning activities is a challenge. Many studies have applied electroencephalography (EEG) to explore neural mechanisms because it has both high time-resolution and applicability (Schu, 1999). In addition to being used to diagnose various brain-related diseases, EEG has shown great potential in studying related brain activities such as cognition, memory, and emotion. It is also an essential measurement for assessing attention status (Aoki et al., 1999; Müller et al., 2000; Palva and Palva, 2007; Srinivasan et al., 2009).

It has been shown that EEG activities in different frequency bands can be related to specific physiological states. The traditional EEG analysis method is to divide the brain activity into different frequency bands, including delta (1–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (30–60 Hz) waves (Teplan, 2002). A lot of studies were considering band power as an important parameter to characterize the state of attention. EEG power-based indices [P_β/P_α, $1/P_\alpha$, and $P_\beta/(P_\alpha + P_\theta)$] were used to assess the sustained attention level in healthy controls and diffused axonal injury patients, and they found significant negative correlations between P_β/P_α, $1/P_\alpha$ indices and the variations of mean reaction time during sustained attention test (Coelli et al., 2018). Hu et al. (2018) extracted the Hjorth parameters and power spectral features to distinguish three attention levels evaluated by a self-assessment model according to 10 subjects' self-reports during a learning process. They proposed a combined procedure with correlation-based feature selection and k-nearest neighbors classification algorithm to achieve the highest accuracy of 80.04%.

In addition to linear characteristics, nonlinear analysis methods have great potential in EEG analysis based on its nonlinear and nonstationary characteristics. Nonlinear dynamical analysis make it possible to study self-organization and pattern formation in the complex neuronal networks of the brain (Stam, 2005). Several studies have reported the association between attentional function and EEG single-scale complexity. Bob et al. (2011) using pointwise correlation dimension to analyze attentional processes related to dissociative states. Rezaeezadeh et al. (2020) focused on entropy measures, including univariate features from individual EEG channels and multivariate features from brain lobes, to diagnosis Attention Deficit Hyperactivity Disorder. Recent studies have shown that applying nonlinear multiscale information analysis to EEG can provide new information about the complex dynamics of brain cognitive function, such as emotion recognition (Gao et al., 2019). Ke et al. (2014) conducted two experiments instructing all subjects to perform tasks with three different levels of attention (i.e., attention, no attention, and rest). Nonlinear parameters including entropy and multiscale entropy were extracted, and a Support Vector Machine (SVM) model was performed for classification between each experiment state and resting states, with 76.19 and 85.24% accuracy, respectively, in the two experiments.

The relationship between EEG activity and attention state is not limited to EEG amplitude changes with experimental tasks, it also includes phase and cross-frequency coupling. Hanslmayr et al. (2005) calculated the phase-locking index using Gabor wavelet analysis with a frequency resolution of 0.5 Hz during a continuous visual target stimulus processing task. The results suggested that focused attention will cause a large phase locking of alpha wave without amplitude change. Szczepanski et al. (2014) found increases in power of high gamma (70–250 Hz) in electrocorticography (ECOG) signal during allocation of visuospatial attention, and these high gamma power increases were modulated by the phase of the ongoing delta/theta (2–5 Hz) phase. Borhani et al. (2019) applied brain network connectivity analysis based on Granger causality on event-related selective attention tasks, and found that the flow of information between independent neural components on the left occipital cortex and the right supplementary motor area became highly coupled on alpha waves during the selective attention tasks. The dynamical analysis is an extension of the EEG measures with a focus on the time sequence of EEG index or functional connectivity networks (Bola and Sabel, 2015; Pagnotta et al., 2020). Previous studies based on perfusion functional magnetic resonance imaging found that the connectivity strength of frontal-parietal network represented by topological characteristics dynamically changes to compensate the cognitive decline during long-term sustained attention task (Taya et al., 2018). The dynamical analysis represents the evolution of neural activity, mostly used is dynamic functional connectivity. At present, the correlation between dynamical sequence of EEG complexity measures and sustained attention performance in the aspect of helping assess attention during sustained attention tasks still need research proof.

To date, little research has been done on the identification of multiple attention levels using continuous EEG feature analysis instead of superposition of event related potentials. Previous studies mostly allowed subjects to control or self-assess their attention level by subjective methods (Li et al., 2013; Chen et al., 2017). Due to the reliable of continuous attention test, and the neural mechanism of continuous attention has been widely recognized (Rosvold et al., 1956), we performed an AX-continuous performance test (AX-CPT) task to get EEG segments of different attention levels. The entropy and multiscale entropy based dynamical complexity analysis were performed for discriminating attention states in four levels. The quantified complexity indices were used as a feature vector to classify different attention levels through Extreme Gradient Boosting (XGBoost). Furthermore, the relationship between EEG complexity indices and task response performance was studied to assess the effectiveness of dynamical complexity analysis in attention recognition. Our hypotheses for this study were that (1) during sustained attention tasks, the reaction time becomes faster or slower is corresponded to different attention state, and this change of attention level can be reflected in the variation of EEG dynamical complexity; (2) the multi-scale nonlinear method will provide more information than the single-scale complexity in recognition of different attention states, and (3) the EEG dynamical complexity-based attention recognition may be more sensitive and effective in frontal region than other brain regions.

MATERIALS AND METHODS

Experiments

Participants

This study included 42 right-handed subjects (16 males, 26 females) with age ranging from 20 to 26 years (mean age 24.26 ± 1.17). All participants had no history of neurological or psychological disorder, and they all had normal vision or were corrected to normal vision (self-report). This study was approved by IEC for clinical research of Zhongda Hospital, affiliated to Southeast University (No. 2019ZDSYLL073-P01). All participants signed the informed written consent form before participating in the experiment.

Sustained Attention Task

In this study, the attention continuous performance test (CPT) task was employed to assess the sustained attention capability of the participants, while their continuous EEG signals were recorded.

The version of the CPT task used in this study was the AX-CPT (Cohen et al., 1999). During the AX-CPT task (see **Figure 1**), the participants were presented with a series of letters of English alphabet randomly, in which they were instructed to inhibit their response when the target sequence is "X" preceded by an "A" and to make a response as fast as possible for other target sequences different from sequence "AX." There were 192 trials, with 144 "AX" sequences (Left mouse button response) and 48 other sequences (right mouse button response), each sequence contained two letters. Each letter stimulus was presented for 250 ms with an inter-stimulus interval (ISI) of 250 ms and an inter-trial interval (ITI) of 1,500, 2,000, or 2,500 ms randomly set. The task was performed in a dimly lit and quiet room.

Questionnaires

All questionnaires were administered to participants after EEG recording, including demographic (age, gender, and handedness), Pittsburgh Sleep Quality Index (PSQI) questionnaire (Buysse et al., 1989), Profile of Mood States (POMS) questionnaire (Curran et al., 1995), and Cognitive Failures Questionnaires (CFQ) (Broadbent et al., 1982).

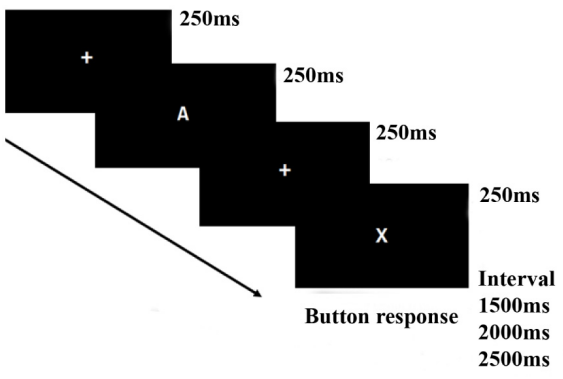


FIGURE 1 | AX-Continuous performance test (AX-CPT) task flow.

EEG Recording and Preprocessing

The EEG data were recorded by 32 electrode cap (Easycap) based on the international 10-10 system (Fp1, Fp2, Af3, Af4, F7, F3, Fz, F4, F8, Fc5, Fc1, Fc2, Fc6, T7, C3, Cz, C4, T8, Cp5, Cp1, CP2, Cp6, P7, P3, Pz, P4, P8, Po3, Po4, O1, Oz, O2) and digitized at 1,000 Hz using the Neuroscan Synamp2. The reference electrode is located near Cz. Eye movements were recorded using two bipolar electrodes (one electrode superior to the right eye, another electrode to the right of the orbital fossa). The impedance of each electrode was below 10 kΩ.

A notch filter at 50 Hz to suppress the remained power-line noise, and a band-pass filter at 0.3–70 Hz using a FIR filter with a 3rd order Butterworth window was used to eliminate movement artifacts. Independent Component Analysis (ICA) was performed for removing EEG ocular artifacts (Vigário, 1997). The 32-channels EEG was decomposed into 30 independent components (ICs) by ICA, and then the Electro-oculogram (EOG) were automatically recognized by calculating the correlation between each IC and two EOG signals. The noise ICs were set to zero, and the other ICs were reconstructed to EEG without ocular noise. The clean EEG was segmented into 3 s sections according to each AX-CPT trial for feature extraction.

Previous studies have shown that response time may be an indicator of attention level (Gunawan et al., 2017). When people are sustaining a high level of concentration, they usually respond to visual stimuli immediately and quickly. Conversely, people with low attention levels usually make slower response. In fact, the ability to respond to tasks rapidly varies with different individuals, so it would happen that some subjects respond relatively slowly on all trials, and some subjects respond relatively quickly on all trials. In this case, evaluation and definition of different attention levels directly by absolute reaction time in a cross-subject attention recognition will not be accurate enough. Hence, for the purpose of cross-subject attention states recognition, we corrected the task response times based on the average reaction time of each person to obtain calibrated response times (C-RTs). Since the AX-CPT task is composed of 192 trials and each trial corresponds to a calibrated response time, the EEG data is segmented according to the event of the task trial with a window length of 3 s. After removing the outliers of C-RTs, an average of 190.76 (±2.55) task epochs segments (1 epoch = 3 s) for each subject were remained. As we can see in **Figure 2**, the distribution of C-RTs in the histogram is very similar to the log-normal distribution.

In this context, we set a critical value α based on the C-RTs to match EEG segments with different attention states. we defined task data epochs with C-RTs significantly shorter than the critical value as "high attention state," data epochs with C-RTs significantly longer than the critical value as "low attention state," and other data segments with medium C-RTs were considered as "medium attention state." For example, α = 0.25 defined that the data segments with C-RTs in the top 25% (shorter response times) represented "high attention state," the data segments with C-RTs in the bottom 25% (longer response times) represented "low attention state," and other segments represented "medium attention state." To increase the reliability of this definition, different significance level α (0.05–0.35) were picked up to

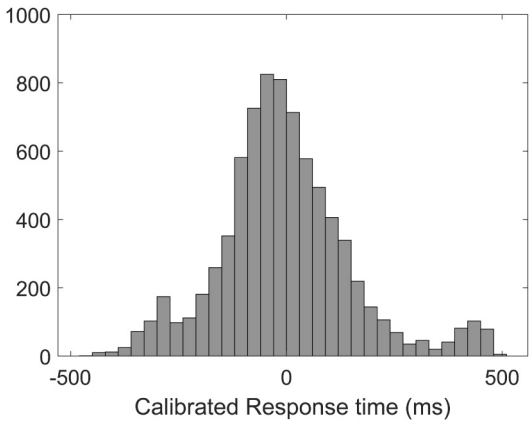


FIGURE 2 | The distribution of the calibrated response times (C-RT).

represent different attention levels in Results section. Besides the task states, we also took the 3-min resting state EEG of each subject for analysis and obtained 60 resting data epochs (1 epoch = 3 s) for each subject.

Frequency Domain Features

Wavelet Packet Decomposition (WPD) (Coifman and Wickerhauser, 1992) projects the time series onto the space of orthogonal wavelet basis functions and decomposes the signal into low frequency and high frequency. Compared with wavelet analysis, it not only decomposes the low-frequency part of the signal, but also re-decomposes the high-frequency part. In this paper, 7-layer WPD was used to obtain theta (θ, 4–8 Hz), alpha (α, 8–13 Hz), and beta (β, 13–30 Hz) band oscillations. The power ratio features β/θ, β/α, β/(α+θ) calculated by Welch's power spectral density estimate methods (Welch, 1967) were seen as classical EEG features for attention recognition in this study.

Single-Scale Complexity

Approximate Entropy

Approximate entropy (ApEn) (Pincus, 1991) uses a non-negative number to represent the complexity of a time series and reflect the occurrence of new information in the time series. For a given time series $[u(1), u(2), ..., u(N)]$, two parameters m and r are defined to compute ApEn. First, take m consecutive points in sequence to form vectors sequence

$$U(i) = \{u(i), u(i+1), ..., u(i+m-1)\} \quad (1)$$

where $i = 1, 2, ..., N-m+1$ then define the distance between the two vectors $U(i)$ and $U(j)$ as

$$d[U(i), U(j)] = \max_{k=0,1,2,..,m-1} \left|u(i+k) - u(j+k)\right| \quad (2)$$

where $i = 1, 2, ..., N-m+1$. For a given threshold r, count the number of $d[U(i), U(j)] < r$ as $N_m(i)$, and the ratio of $N_m(i)$ to the total number of distances N-m+1 is recorded as $C_m^r(i)$,

$$C_m^r(i) = N_m(i)/(N-m+1) \quad (3)$$

Take the logarithm of $C_m^r(i)$ and calculate its average values for all i, denoted as $\phi_m(r)$,

$$\phi_m(r) = (N-m+1)^{-1} \sum_{i=1}^{N-m+1} lnC_m^r(i) \quad (4)$$

Increase the dimension to $m+1$, and recalculate $\phi_{m+1}(r)$ according to the above steps, the approximate entropy of the sequence is

$$ApEn(m, r, N) = \phi_m(r) - \phi_{m+1}(r) \quad (5)$$

The parameters for ApEn was set as $m = 2$,$r = 0.2$.

Sample Entropy

Sample entropy (SampEn) is an improved complexity measurement based on the concept of ApEn (Richman and Moorman, 2000). The initial calculation process is the same as ApEn, but there is an additional restriction that $i \neq j$ for the calculation of $d[U(i), U(j)]$ in formula (2).

SampEn is calculated as

$$SampEn(m, r, N) = \ln(\phi_m(r)/\phi_{m+1}(r)) \quad (6)$$

The parameter for SampEn was set as $m = 2$,$r = 0.15$.

Fuzzy Entropy

Fuzzy entropy (FE) uses a fuzzy membership function to measure the degree of similarity of vectors (Chen et al., 2007), so the calculated values are smooth, continuous, and more robust. The phase-space reconstruction is performed on time series $[u(1), u(2), ..., u(N)]$, and a set of m-dimensional ($m \leq N$-2) vectors are reconstructed as follows:

$$X_i^m = \{u(1), u(2), ..., u(i+m-1)\} - u_m(i), i = 1, 2, ..., N-m+1 \quad (7)$$

where $u_m(i)$ represents the mean of this m-dimensional vector.

The maximum Euclidean distance between X_i^m and X_j^m was defined as d_{ij}^m, given n and r, the degree of similarity of two vectors can be calculated according to the fuzzy membership function:

$$A_{ij}^m = u(d_{ij}^m, n, r) = \exp(-(d_{ij}^m)^n/r) \quad (8)$$

Define the function $\varphi^m(n, r)$ as

$$\varphi^m(n, r) = \frac{1}{N-m} \sum_{i=1}^{N-m} [\frac{1}{N-m-1} \sum_{j=1, j\neq i}^{N-m} A_{ij}^m] \quad (9)$$

The fuzzy entropy for the given time series can be defined as

$$FuzzyEn(m, r, n, N) = ln\varphi^m(n, r) - ln\varphi^{m+1}(n, r) \quad (10)$$

The parameters for FuzzyEn was set as $m = 2$, $r = 0.15$, $n = 2$.

Multiscale Complexity

Multiscale Sample Entropy

Multiscale Entropy analysis is a method to investigate the complexity of time series at multiple time scales (Costa et al., 2005). Multiscale sample entropy (MSE) extends sample entropy to multiple time scales or resolutions. The basic principle of MSE involves coarse-graining and entropy calculation procedure. For a given time series$[u(1), u(2), ..., u(N)]$, the range of scale factor τ is defined from 1 to s, and the original time series was coarse-grained according to the scale factor τ as follows:

$$y_j^{\tau} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} u(i), j = 1, 2, ..., int\ (N/\tau) \qquad (11)$$

Calculate the SampEn for each coarse-grained series $\{y_j^{\tau}\}$ in different scale factors.

$$MSE(\tau, m, r) = SampEn_{y_j^{\tau}}(m, r, N), \tau = 1, 2, ..., s \qquad (12)$$

Multiscale Fuzzy Entropy

The calculation principle of multiscale fuzzy entropy (MFE) (Zheng et al., 2014) are similar to the steps of MSE. After the coarse-graining procedure in all scale factors, MFE is defined as:

$$MFE(\tau, m, r) = FuzzyEn_{y_j^{\tau}}(m, r, N), \tau = 1, 2, ..., s \qquad (13)$$

When analyzing multiscale complexity, we used $m = 2$, $r = 0.15$ for MSE and $m = 2$, $r = 0.15$, $n = 2$ for MFE, and scales were set from 1 to 50.

Extreme Gradient Boosting

Extreme gradient boosting (XGBoost) model was used as a classifier to distinguish different levels of sustained attention. XGBoost is an efficient and distributed implementation of the Gradient Boosting algorithm (Chen and Guestrin, 2016). The main advantage of XGBoost lies in its scalability, which allows parallel and distributed computing, and makes learning and model exploration faster. And approximate algorithm used in XGBoost finds the candidate set of cutting points according to the quantile of the feature distribution, and then traverses all the sub-sets to determine the best split point. This method replaces the greedy algorithm that needs to traverse all samples when looking for the best segmentation point in training. In addition, XGBoost proposed a more regularized model formalization to prevent over-fitting, so that the performance of XGBoost is better than the conventional Boosting algorithm. Apart from that, XGBoost is an unexplored algorithm in the field of attention recognition. So here in this work we explored this algorithm to get better accuracy. The subject leave-one-out cross-validation (LOOCV) and 5-fold-cross validation were both performed to evaluate the model's predictive performance. XGBoost library through the python package was used to complete this work.

Statistical Analysis

The D'Agostino-Pearson normality test was performed to check whether the data satisfied the normal distribution. The parametric ANOVA test was used for the data that satisfied normality. If the data are not normally distributed, a non-parametric Kruskal-Wallis test was used.

To uncover the associations between the dynamical analyzed indices and behavioral parameters, Spearman's correlation coefficient was performed between the EEG features and calibrated response time (C-RTs). The significant difference was defined as the p-value < 0.05. All statistical analysis methods were performed in MATLAB R2018a.

RESULTS

Behavior Analysis

Descriptive statistics were presented in **Table 1**, behavioral results were reported in terms of mean reaction time (RT) (ms) and mean errors (%) for all subjects. The average PSQI score (mean = 5.31, SD = 2.09) can be used as evidence to support that the currently selected group of subjects does not have insomnia disorder (Dietch et al., 2016). The average Total Mood Disturbance (TMD) score of POMS (mean = 108.31, SD = 20.60) shown that the emotional state of participants is normal and there is no negative mood swing [TMD score range from 68 to 268, a higher TMD score is indicative of greater mood disturbance (McNair et al., 1992)]. The average CFQ score (mean = 52.90, SD = 12.47) suggested that the concentration of participants in daily life is at a normal range (Wallace et al., 2002). No abnormal value was found in the PSQI, POMS, and CFQ questionnaires.

EEG Dynamical Complexity Between Resting and Attention States

We analyzed the ApEn, SampEn, FuzzyEn, MSE, and MFE of all 32 channels for all EEG epochs to investigate the single-scale and multiscale complexity, and the group difference of EEG complexity between resting state (eyes opened) and sustained

TABLE 1 | Mean task results of all subjects.

Characteristics	Mean (standard deviation)	n	%
Subject			
Male		16	38.10
Female		26	61.90
Age (years)	24.26 (1.17)		
Education level	college student	42	100
Handedness	right-handed	42	100
AX-CPT task			
Mean RT (ms)	405.04 (96.52)		
Mean errors (%)	8 (6.65)		
Questionnaires			
PSQI	5.31 (2.09)		
POMS	108.31 (20.60)		
CFQ	52.90 (12.47)		

attention state was presented in **Figure 3**. The three single-scale complexity values during the sustained attention task were significantly higher than resting state (**Figure 3A**, ApEn, $p < 0.001$, SampEn, $p < 0.001$, and FuzzyEn, $p < 0.001$). Distinct increases of multiscale complexity curves (scales 1–50) measured by MSE and MFE were observed during sustained attention state (**Figures 3B,C**). These promising results suggest that complexity has the potential to distinguish different levels of attention states.

EEG Dynamical Complexity Compared Among Four Attention Levels

Single-Scale EEG Complexity

The single-scale EEG complexity for different attention levels was explored using three kinds of entropy calculation methods, i.e., ApEn, SampEn, and FuzzyEn. Four different levels of attention states were defined in this study, including high attention (HA), medium attention (MA), low attention (LA), and resting state (RS). The definition criterion for different attention levels was introduced in section "EEG Recording and Preprocessing." Here, we first took $\alpha = 0.25$ as the representative case in this section, and finally $\alpha = 0.05$–0.35 were evaluated accordingly in section "Effects of Significance Level Alpha."

The brain topography of averaged ApEn, SampEn, and FuzzyEn for all 42 subjects was shown in **Figure 4**, representing HA, MA, LA, and RS, separately. These topography maps demonstrate that (i) consistently in all three entropy methods, higher attention levels showed higher single-scale EEG complexity than lower attention levels in frontal and central brain regions, whereas resting state showed lowest EEG complexity than keeping focused (see **Figure 4A**), and (ii) these differences among the four attention levels in frontal and central regions were highly significant (evaluated by nonparametric Kruskal-Wallis test, see **Figure 4B**).

Furthermore, as illustrated in **Figures 4C,D**, we took Fz and F4 as representative results to show the group differences in frontal regions. FuzzyEn achieved highest statistical difference than ApEn and SampEn among HA, MA, and LA states, representatively shown in **Figures 4C,D**, ApEn (Fz: $p < 0.05$, F4: $p < 0.01$), SampEn (Fz: $p < 0.01$, F4: $p < 0.01$), and FuzzyEn (Fz: $p < 0.01$, F4: $p < 0.01$). Approximately equal significant group differences were found among four attention levels ($p < 0.001$ for both SampEn and FuzzyEn measures in both Fz and F4). Similar group differences were found in other EEG channels (i.e., channels marked with the red asterisk in **Figure 4B**).

Thus, the comparison results of single-scale complexity measured by ApEn, SampEn, and FuzzyEn showed that most of the frontal lobe electrodes (e.g., FP1, FP2, AF3, AF4, F7, F3, Fz, F4, F8, Fc1, Fc2) and part of the parietal lobe electrodes (e.g., C3, C4) all have significant differences among different attention levels, suggesting that characterization of different attention levels is more associated with EEG complexity in the frontal and central regions than other brain regions.

Multiscale EEG Complexity

The dynamical complexity measured from the temporal regularity of EEG oscillations across different time scales may provide more information. The multiscale EEG complexity of three different attention levels and resting state were analyzed by multiscale measures of both MSE and MFE. The averaged MSE and MFE curves of F4 were shown in **Figure 5**. As in the enlarged detail plots at scales 1–10, more obvious difference was found in MFE curves (**Figure 5B**) than MSE curves (**Figure 5C**) among four attention states, showing that the higher the attention level is, the higher the entropy value at this scale. In **Figure 5A**, the mean sample entropy values of MA and LA are approximately equal across different time scales. At time scales 11–50, the curves of four attention levels are aliased together in both **Figures 5A,B**. Hence, the area under MSE and MFE curves at time scales 1–10 was defined as multiscale sample entropy index (MSEI) and multiscale fuzzy entropy index (MFEI). **Figure 5C** displayed the group differences among both three different attention levels (i.e., HA, MA and LA, MSEI: $p < 0.01$ and MFEI: $p < 0.001$) and four attention

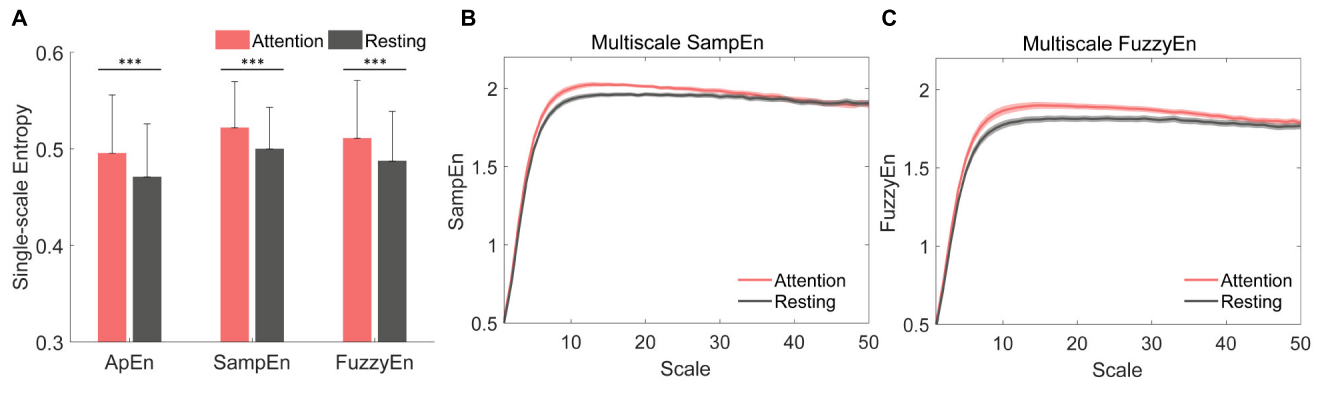


FIGURE 3 | Comparison of complexity measured separately by approximate entropy (ApEn), sample entropy (SampEn), fuzzy entropy (FuzzyEn), Multiscale sample entropy (MSE), and Multiscale fuzzy entropy (MFE) between resting state and sustained attention state at Fz. **(A)** The averaged values of single-scale complexity (ApEn, SampEn, and FuzzyEn) for 42 subjects [***$p < 0.001$, the error bar represented standard deviation (SD)]. **(B)** The averaged Multiscale complexity measured by MSE and **(C)** MFE for 42 subjects [shadows represented standard error (SE)].

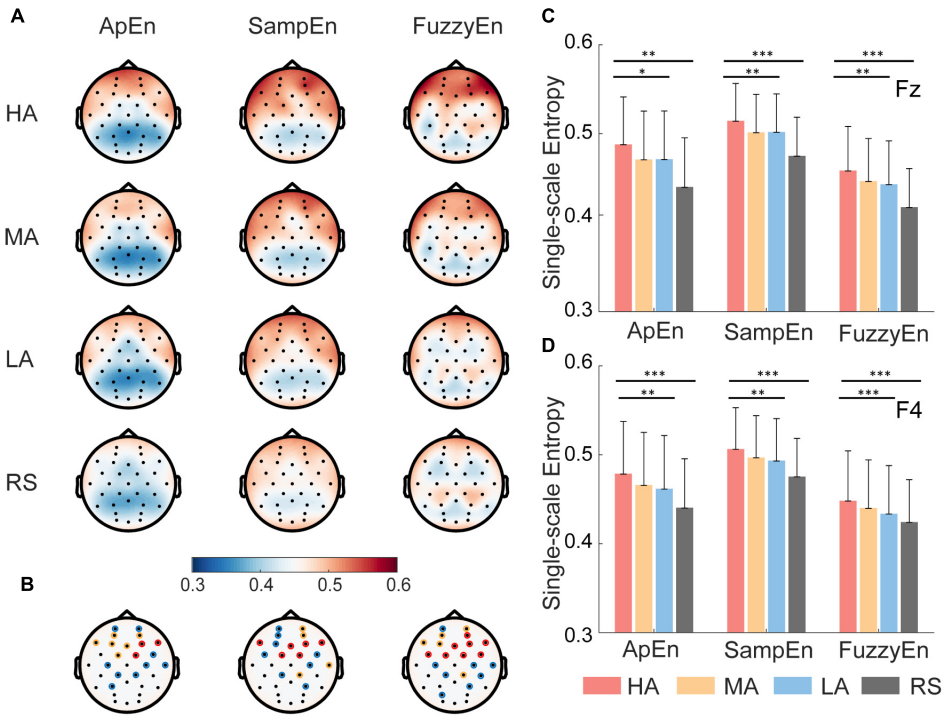


FIGURE 4 | Brain topography and representative results of single-scale EEG complexity. **(A)** Topography of approximate entropy (ApEn), sample entropy (SampEn), and fuzzy entropy (FuzzyEn) for all 42 subjects at four attention states (i.e., high attention—HA, medium attention—MA, low attention—LA, and resting state—RS). The red color means high entropy, and blue means low entropy. **(B)** The statistical significance of 32 channels using Kruskal-wallis test. The color of circle indicates significance level α, blue represents α in the range of (0.01, 0.05], yellow represents α in the range of (0.001, 0.01], red represents α smaller than 0.001. **(C,D)** Representative results of single-scale complexity measured group differences for three attention levels (HA, MA, and LA) and four different attention levels (HA, MA, LA, and RS) in frontal region at Fz and F4. Nonparametric Kruskal-Wallis test, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Error bars = ±1 SD.

levels (i.e., HA, MA, LA, and RS, MSEI: $p < 0.001$ and MFEI: $p < 0.001$).

Recognition of Multi-Level Attention

Classification Results

For the dynamical complexity method, both single-scale (i.e., ApEn, SampEn, FuzzyEn) and multiscale complexity indices (i.e., MSEI, MFEI) extracted from EEG recordings in the frontal region (i.e., FP1, FP2, AF3, AF4, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6) were fed into the classification model. To prove the efficiency and advantage of complexity-based features, the conventionally used power ratios of different frequency bands which have been widely considered to be related with attention level (Barry et al., 2009; Putman et al., 2010), e.g., β/θ, β/α, β/(α+θ), were applied as classical method to recognize different level of attention states. For the classical method, 3 power ratio features β/θ, β/α, β/(α+θ) was extracted from the same frontal electrodes. Hence, the feature dimension of each sample in the complexity-based model is 65 (5*13) and for the classical model, each sample is described by 39 (3*13) features. To eliminate the influence of individuality on classification, the features of all data segments were separately normalized to 0–1 in all subjects, respectively, before classification.

Furthermore, in order to further evaluate the distinguish performance of the Complexity-XGBoost, SVM, and random forest (RF) were also performed, which have been commonly used in EEG analysis. The parameters optimizing process was performed for each classifier model during training.

Table 2 demonstrated the classification results of multi-level attention recognition obtained from both the LOOCV and 5-fold cross-validation. With LOOCV, the proposed Complexity-XGBoost model achieved accuracy of 64.69, 70.49, and 76.39%, respectively for four-level attention states classification (HA, MA, LA, and RS), three-level attention states classification (HA, MA, and LA), and two-level attention states classification (attention state (AS), and RS). With 5-fold cross-validation, the Complexity-XGBoost model achieved accuracy of 81.39, 80.42, and 95.36%, respectively.

Compared with classical methods, complexity-based methods resulted in best performance in all two validations and three classification strategies, and this advantage is more distinct when using the 5-fold cross-validation. When both LOOCV and 5-fold cross-validation was conducted, the performance in XGBoost outperformed those of other two machine learning methods. The results of the SVM were not much different from those of the RF.

The Receiver Operating Characteristic (ROC) curves of Complexity-XGBoost and Classical-XGBoost methods for three classification strategies are showed in **Figure 6**. The area under curve (AUC) of Complexity-XGBoost method are 0.95, 0.91, and 0.99 respectively, for four-level, three level, and two-level

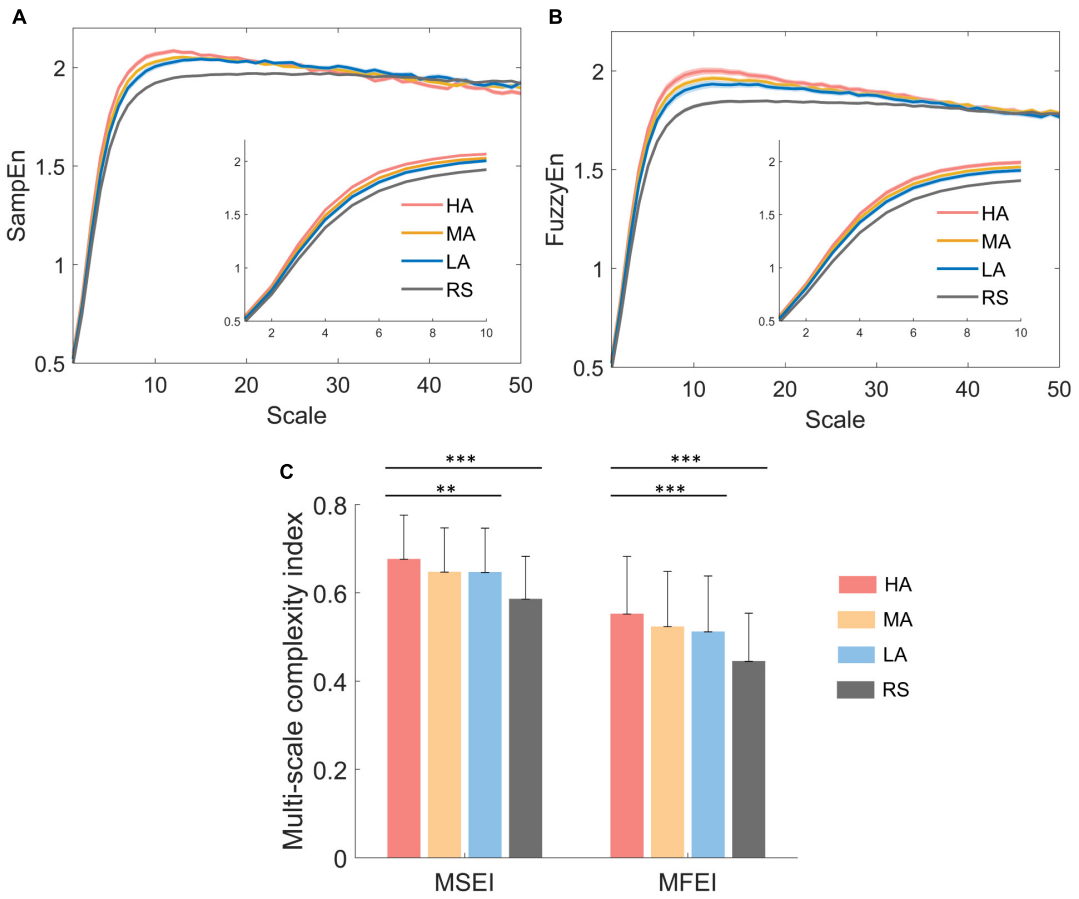


FIGURE 5 | Representative results of multiscale complexity analysis on electrodes F4. **(A)** The averaged multiscale sample entropy curves and **(B)** averaged multiscale fuzzy entropy curves of four attention states for 42 subjects, and the shadow represents SE. **(C)** Group differences between MSEI and MFEI among different attention levels. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. Error bars = SD.

TABLE 2 | The performance between XGBoost and other machine learning methods in the LOOCV and 5-fold cross-validation.

Methods	Model	LOOCV			5-Fold cross validation		
		4-Level	3-Level	2-Level	4-Level	3-Level	2-Level
Classical methods	SVM	59.23 ± 8.32	65.14 ± 7.79	73.82 ± 8.72	62.42 ± 0.51	72.00 ± 1.73	82.32 ± 0.81
	RF	60.19 ± 5.21	64.12 ± 4.96	72.34 ± 7.46	63.52 ± 0.65	73.39 ± 1.10	81.71 ± 1.37
	XGBoost	63.15 ± 4.63	65.24 ± 3.86	73.39 ± 4.47	67.45 ± 0.48	72.75 ± 1.03	83.46 ± 1.83
Complexity analysis	SVM	60.13 ± 12.32	70.34 ± 9.76	75.43 ± 13.45	78.46 ± 1.12	78.25 ± 1.19	94.34 ± 1.18
	RF	63.77 ± 8.63	72.33 ± 10.22	73.70 ± 11.68	77.27 ± 1.73	77.68 ± 0.90	94.12 ± 0.34
	XGBoost	64.69 ± 6.20	70.49 ± 4.59	76.30 ± 9.24	81.39 ± 1.47	80.42 ± 0.84	95.36 ± 2.31

attention states classification. And the AUC of Classical-XGBoost method are 0.83, 0.81, and 0.88, respectively, for three kinds pf classification strategies.

Feature Importance

The complexity-based features and the classical PSD features from 32 channels were combined to conduct a Combined-XGBoost classifier. One advantage of this approach is that we can retrieve the importance score of each feature after constructing the gradient boosted trees to obtain the importance ranking of the feature. The top 10 important features and their electrode positions are shown in **Figure 7**. The top 10 important features were all complexity-based features. MFEI appears 4 times in the top 10 important features, FuzzyEn 3 times, MSEI twice, and SampEn once, indicating that the complexity-based features in identifying different attention levels were far better than the PSD features. Meanwhile, the electrode positions where these features are located also showed that the frontal lobe brain area is of great significance for identifying different level of attention.

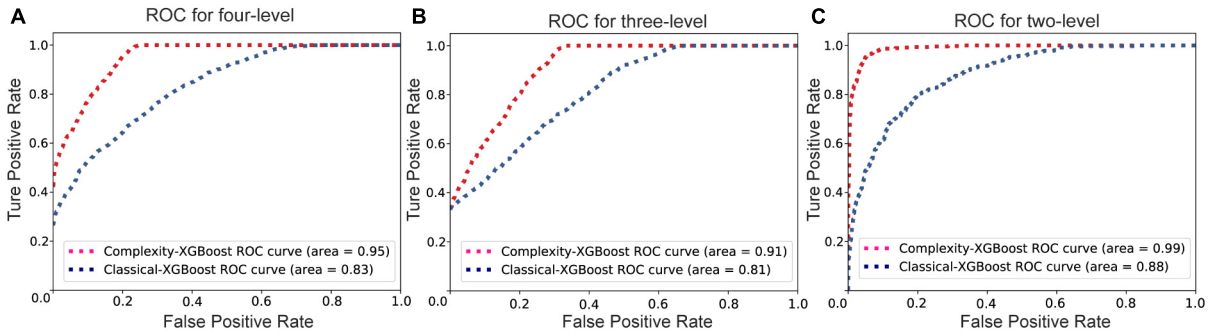


FIGURE 6 | The Receiver Operating Characteristic (ROC) curves of Complexity-XGBoost and Classical-XGBoost methods for three classification strategies. **(A)** The ROC curve for four-level classification. **(B)** The ROC curve for three-level classification. **(C)** The ROC curve for two-level classification.

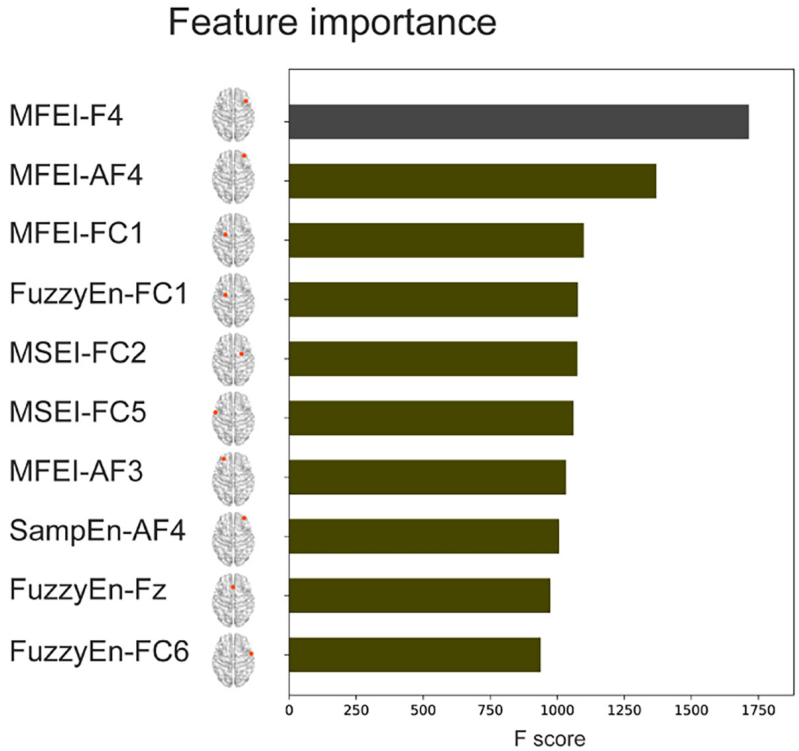


FIGURE 7 | Feature importance of Combined-XGBoost (top 10 features).

Effects of Significance Level Alpha

The above results were obtained based on the definition criterion $\alpha = 0.25$ for different attention levels (introduced in sections "EEG Dynamical Complexity Compared Among Four Attention Levels" and "Recognition of Multi-Level Attention"). Here, to investigate the reliability of complexity-based features for characterizing specific attention level, we evaluated the multi-level attention recognition accuracy under different significance level $\alpha = 0.05$–0.35 (see **Figure 8**, 4-level attention classification). With the increase of α-level, the classification accuracy of classical methods based on the power ratio tends to decrease dramatically from 80.59% ($\alpha = 5\%$) to 64.40% ($\alpha = 30\%$), while the methods based on complexity analysis have much higher accuracies (73.91–95.01%) for the recognition of the four attention states. These findings further confirmed the advantages of nonlinear complexity analysis for attention-related EEG recognition. Hence, in this study we took $\alpha = 0.25$ as the representative threshold to define HA, MA, and LA is reliable.

Relationship Between EEG Dynamical Features and Real-Time Response During Sustained Attention Task

Brain dynamics research has highlighted the contributions of the ongoing EEG to behavioral responses. In this section, we

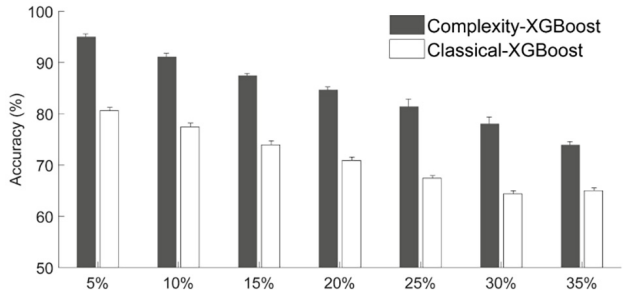


FIGURE 8 | The effects of the significance level α to classification accuracy.

examined the correlation effects of state-related EEG changes on stimulus-response efforts during sustained attention task. First, the averaged time-varying C-RTs across subjects were calculated. As for multi-channels EEG features, we performed principal component analysis (PCA) method to realize dimensionality reduction. PCA and related techniques have been applied to describe the fluctuation of EEG measurements during the resting state (Leonardi et al., 2013), continuous movie-watching task (Demirtaş et al., 2019), and whole-brain connectivity dynamics (Allen et al., 2014; Zhu et al., 2020). PCA is a method accepted by many researches to reduce the dimensionality of multi-channel or whole brain features, and then to study dynamic fluctuation. PCA was performed on the dynamical EEG features of 13 electrodes in the frontal brain area within each subject, then the first principal component (PC1) was selected as the representative EEG features for each subject. Finally, the averaged time-varying PC1 of 42 subjects were calculated. These correlation effects were estimated using a Spearman's correlation test between averaged EEG features and averaged C-RTs of all subjects along with data segments over time.

Five complexity-based features (ApEn, SampEn, FuzzyEn, MSEI, and MFEI) and 3 power ratio features [β/θ, β/α, and β/(α+θ)] are all took into consideration separately. Significant negative correlations were found between the involved EEG

TABLE 3 | Comparison with previous studies on attention recognition.

Authors	Attention task (levels)	Subjects	Methods	Window (seconds)	Brain regions (channels)	Validation	Accuracy (%)		
							2-Levels		
Chen et al. (2017)	Continuous performance task (high-attention, low-attention)	10	Temporal and entropy features—SVM	Trial length	Prefrontal (1)	3/4 train, 1/4 test	91.60		
							3-Levels		
Hu et al. (2018)	Randomly selected learning task (high, neutral, low)	10	Linear and nonlinear features—CFS+KNN	180	Central and temporal (6)	10 times 3-Fold CV	80.84		
							2-Levels	**3-Levels**	
Gaume et al. (2019)	Continuous performance task (easy, medium, and hard)	14	Power features—LDA	5	Whole brain (16)	Leave-one-subject-out	75	51.8	
				30			85	64.8	
							2-Levels (in-ear)	**2-Levels (prefrontal)**	
Jeong and Jeong (2020)	Psychomotor vigilance tasks (attention, rest)	6	Temporal and spectral features—Echo State Network	0.5	In-ear (2) prefrontal (2)	Within-subject	81.16	82.44	
						Cross-subject	64.00	65.70	
						10-Fold CV	74.15	73.73	
							2-Levels	**3-Levels**	**4-Levels**
Our study	AX-CPT (rest, LA, MA, HA)	42	Complexity—XGBoost	3	Frontal (13)	Leave-one-subject-out	76.30	70.49	64.69
						5-Fold CV	95.36	80.42	81.39

SVM, support vector machine; CFS, correlation-based feature selection; KNN, k-nearest-neighbor; LDA, Linear discriminant analysis; ESN, Echo State Network; CV, cross-validation.

features (except the classical features β/α) and C-RTs, MFEI had the highest correlation ($r = -0.35$, $p < 0.001$).

DISCUSSION AND CONCLUSION

This study proposed to use a calibrated response times (C-RTs) to obtain multi-level attention states during an AX-CPT sustained attention test, which can truly reflect the changes in attention without the influence of individuality on response. The proposed entropy-based Complexity-XGBoost model achieved outstanding performance, respectively, in recognizing four, three, and two levels of attention states relative to a similar model trained on conventional power spectral based measures. Furthermore, we found significant correlation relationships between complexity-based EEG features and C-RTs over time.

EEG signal is a nonlinear coupling of large number of nerve cells. The linear EEG analysis can evaluate the communication between neural networks in the same oscillating frequency band or similar neuron firing patterns. However, it is not clear how much information is missing since the behavior of neural network can be highly nonlinear and nonstationary (Yang et al., 2018; He and Yang, 2021). Thus, nonlinear analysis methods like entropy and complexity are more suitable for EEG feature extraction than the power spectral based linear analysis. The dynamical complexity of the neural network should correlate with the conscious state of the subject (Tononi and Edelman, 1998). The neural network based on automatic behavior or low-control behavior should have lower dynamical complexity than the neural network that consciously controls behavior, such as controlling oneself to maintain a high attention state.

Compared with the performance of previous studies on the EEG application of attention state monitoring (Chen et al., 2017; Hu et al., 2018; Gaume et al., 2019; Jeong and Jeong, 2020), we dealt with the recognition up to four levels of attention states and our performance is higher than them using Complexity-XGBoost (see **Table 3**). The Complexity-XGBoost achieved the accuracy of 81.39 ± 1.47% for four-level attention states (HA, MA, LA, and RS), 80.42 ± 0.84% for three-level attention states (HA, MA, and LA), and 95.36 ± 2.31% for two-level attention states (AS and RS) when using 5-fold cross-validation. With LOOCV, the accuracies were 64.69 ± 6.20%, 70.49 ± 4.59%, and 76.30 ± 9.24%, respectively. The performance of the three-level attention classification (without RS) is still relatively high, indicating that the accuracy of the four-level classification is not affected by the obvious difference between resting state and attention state.

The proposed attention recognition model based on the XGBoost algorithm adds a regularization step to the traditional gradient enhancement algorithm, which can reduce the degree of overfitting of training and improve the performance of cross-subject classification. Moreover, the Complexity-XGBoost model supports multi-threaded parallel computing, and the approximate algorithm is used to replace the greedy algorithm when looking for the best segmentation point, thereby greatly improving the computing efficiency of the algorithm and reducing the computational cost in real-time training. In terms of the recognition of small and medium-sized data, such as attention recognition using EEG features, algorithms based on XGBoost are by far one of the best ways for an application purpose. So attention recognition based on Complexity-XGBoost model has great application potential in actual portable brain-computer interface applications.

Furthermore, the interpretable property of the XGBoost algorithm also showed the importance of complexity-based features and frontal brain region. The presented attention recognition results in this study were obtained using EEG features derived from frontal brain region instead of the whole brain. The frontal region is involved in the generation of top-down control signals for attention transition, especially the prefrontal lobe area plays a vital role in the ability to switch attention control based on changing task requirements (Rossi et al., 2009). It is suggested that the frontal region played an important role in attention regulation (Daffner et al., 2000; Paneri and Gregoriou, 2017). Both the brain topography of three single-scale complexity indices (see **Figure 4**) and feature importance ranking with electrode locations (see **Figure 7**) showed that frontal channels could distinguish different levels of attention states more significantly than any other brain area. The frontal cortex, the control center for most cognitive functions, is considered a higher order area that controls several executive functions including taking charge of the brain's attention and controlling relevant parts of the visual cortex (Baldauf and Desimone, 2014). However, previous studies found that other brain regions, including parietal and occipital, also reflect attentional modulation (van Schouwenburg et al., 2017; Magosso et al., 2019; Misselhorn et al., 2019). These studies focused on the relationship of attentional modulation and alpha oscillation or alpha power, which all used linear analysis methods. It appears that Hu et al. (2018) excluded frontal electrodes (they used C3, C4, Cz, P3, P4, Pz), whereas the this study focused on frontal electrodes. The mechanism that the linear analysis method and the nonlinear analysis method behave differently in different brain regions is worthy of further investigation. In this study, we proposed and verified the effective dynamical complexity indices for attention evaluation based on frontal EEG, which is a good impetus for the application of portable prefrontal EEG devices to promote real-time attention state assessment. Attention recognition based on Complexity-XGBoost could applied to the neural feedback treatment of diseases that affect cognitive function such as attention deficit hyperactivity disorder, mild cognitive impairment, or Alzheimer's disease.

This study also provides an instance of EEG dynamical correlation analysis to investigate the effectiveness of EEG features in attention assessment. Our results showed that dynamical complexity measures were related to the changing process of response, e.g., dynamical complexity measure MFEI ($r = -0.35$, $p < 0.001$) were significantly correlated to the task performance during sustained attention task. Previous studies also demonstrated that attention dynamically modulates brain rhythms (Liu et al., 2020; McCusker et al., 2020; Pagnotta et al., 2020). McCusker et al. (2020) observed multi-spectral oscillatory robust effected by attention dynamically for both the directed and divided attention experiments in a MEG study.

Liu et al. (2020) applied an complex analysis framework composed of weighted phase lag index and tensor component analysis and they found dynamic organizations of frequency-specific function connectivity can track the decrement and motivation of attention in sustained task. In a simplified perspective, dynamical complexity analysis conducted in this study may offer additional predictive value for attention.

One limitation of this study is that the time course of AX-CPT tasks was not sufficiently long, resulting in the reaction time not being able to progress overtime to produce more obvious changes. In future research, multi-session variable-speed AX-CPT tasks and longer experimental time will be performed to verify the advantages of nonlinear complexity methods for attention recognition. In addition, we also hope to use cross-frequency neural coupling measurement for attention recognition.

The present study investigated brain dynamical complexity concurrently during rest and a task characterized by sustained attention. The present findings demonstrated that dynamical complexity and XGBoost achieved great performance for different levels of attention states recognition, and significant differences were observed in the frontal regions.

DATA AVAILABILITY STATEMENT

The datasets for this article are not publicly available. Requests to access the datasets should be directed to XC, cuixr@seu.edu.cn.

ETHICS STATEMENT

This study was approved by the IEC for clinical research of Zhongda Hospital, affiliated to the Southeast University (No. 2019ZDSYLL073-P01). The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

ZGu and XC conceived and designed the study. WW and ZGa performed the experiment and analyzed the data. WW drafted the manuscript. XC supervised the analysis, reviewed, and editing the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported in part by the National Natural Science Foundation of China under Grant No. 61807007 and in part by National Key Research and Development Program of China under Grant No. 2018YFC2001100.

ACKNOWLEDGMENTS

We thank all the subjects who participated in this study, the members from the Wearable Technology, and Signal Analysis Lab in the Southeast University for helping with data collection.

REFERENCES

Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., and Calhoun, V. D. (2014). Tracking whole-brain connectivity dynamics in the resting state. *Cereb. Cortex* 24, 663–676. doi: 10.1093/cercor/bhs352

Aoki, F., Fetz, E. E., Shupe, L., Lettich, E., and Ojemann, G. A. (1999). Increased gamma-range activity in human sensorimotor cortex during performance of visuomotor tasks. *Clin. Neurophysiol.* 110, 524–537. doi: 10.1016/S1388-2457(98)00064-9

Baldauf, D., and Desimone, R. (2014). Neural mechanisms of object-based attention. *Science* 344, 424–427. doi: 10.1126/science.1247003

Barry, R. J., Clarke, A. R., Johnstone, S. J., McCarthy, R., and Selikowitz, M. (2009). Electroencephalogram θ/β ratio and arousal in attention-deficit/hyperactivity disorder: evidence of independent processes. *Biol. Psychiatry* 66, 398–401. doi: 10.1016/j.biopsych.2009.04.027

Bob, P., Golla, M., Epstein, P., and Konopka, L. (2011). EEG complexity and attentional processes related to dissociative states. *Clin. EEG Neurosci.* 42, 175–179. doi: 10.1177/155005941104200306

Bola, M., and Sabel, B. A. (2015). Dynamic reorganization of brain functional networks during cognition. *NeuroImage* 114, 398–413. doi: 10.1016/j.neuroimage.2015.03.057

Borhani, S., Abiri, R., Jiang, Y., Berger, T., and Zhao, X. (2019). Brain connectivity evaluation during selective attention using EEG-based brain-computer interface. *Brain Comput. Interfaces* 6, 25–35. doi: 10.1080/2326263X.2019.1651186

Broadbent, D. E., Cooper, P. F., FitzGerald, P., and Parkes, K. R. (1982). The cognitive failures questionnaire (CFQ) and its correlates. *Br. J. Clin. Psychol.* 21, 1–16. doi: 10.1111/j.2044-8260.1982.tb01421.x

Buysse, D. J., Reynolds, C. F. III, Monk, T. H., Berman, S. R., and Kupfer, D. J. (1989). The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatry Res.* 28, 193–213. doi: 10.1016/0165-1781(89)90047-4

Chen, C. M., Wang, J. Y., and Yu, C. M. (2017). Assessing the attention levels of students by using a novel attention aware system based on brainwave signals. *Br. J. Educ. Technol.* 48, 348–369. doi: 10.1111/bjet.12359

Chen, T., and Guestrin, C. (2016). "XGBoost: a scalable tree boosting system," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (San Francisco, CA: ACM), 785–794. doi: 10.1145/2939672.2939785

Chen, W., Wang, Z., Xie, H., and Yu, W. (2007). Characterization of surface EMG signal based on fuzzy entropy. *IEEE Trans. Neural Syst. Rehabil. Eng.* 15, 266–272. doi: 10.1109/TNSRE.2007.897025

Coelli, S., Barbieri, R., Reni, G., Zucca, C., and Bianchi, A. M. (2018). EEG indices correlate with sustained attention performance in patients affected by diffuse axonal injury. *Med. Biol. Eng. Comput.* 56, 991–1001. doi: 10.1007/s11517-017-1744-5

Cohen, J. D., Barch, D. M., Carter, C., and Servan-Schreiber, D. (1999). Context-processing deficits in schizophrenia: converging evidence from three theoretically motivated cognitive tasks. *J. Abnorm. Psychol.* 108, 120–133. doi: 10.1037/0021-843X.108.1.120

Coifman, R. R., and Wickerhauser, M. V. (1992). Entropy-based algorithms for best basis selection. *IEEE Trans. Inform. Theory* 38, 713–718. doi: 10.1109/18.119732

Costa, M., Goldberger, A. L., and Peng, C. K. (2005). Multiscale entropy analysis of biological signals. *Phys. Rev. E Stat. Nonlin. Soft. Matter. Phys.* 71(Pt 1):021906. doi: 10.1103/PhysRevE.71.021906

Curran, S. L., Andrykowski, M. A., and Studts, J. L. (1995). Short form of the profile of mood states (POMS-SF): psychometric information. *Psychol. Assess.* 7, 80–83. doi: 10.1037/1040-3590.7.1.80

Daffner, K. R., Mesulam, M. M., Scinto, L. F., Acar, D., Calvo, V., Faust, R., et al. (2000). The central role of the prefrontal cortex in directing

attention to novel events. *Brain* 123(Pt 5), 927–939. doi: 10.1093/brain/123.5.927

Demirtaş, M., Ponce-Alvarez, A., Gilson, M., Hagmann, P., Mantini, D., Betti, V., et al. (2019). Distinct modes of functional connectivity induced by movie-watching. *Neuroimage*. 184, 335–348. doi: 10.1016/j.neuroimage.2018.09.042

Dietch, J. R., Taylor, D. J., Sethi, K., Kelly, K., Bramoweth, A. D., and Roane, B. M. (2016). Psychometric evaluation of the PSQI in U.S. college students. *J. Clin. Sleep Med.* 12, 1121–1129. doi: 10.5664/jcsm.6050

Fortenbaugh, F. C., Degutis, J., and Esterman, M. (2017). Recent theoretical, neural, and clinical advances in sustained attention research. *Ann. N. Y. Acad. Sci.* 1396, 70–91. doi: 10.1111/nyas.13318

Gao, Z., Cui, X., Wan, W., and Gu, Z. (2019). Recognition of emotional states using multiscale information analysis of high frequency EEG oscillations. *Entropy* 21:609. doi: 10.3390/e21060609

Gaume, A., Dreyfus, G., and Vialatte, F. B. (2019). A cognitive brain–computer interface monitoring sustained attentional variations during a continuous task. *Cogn. Neurodyn.* 13, 257–269. doi: 10.1007/s11571-019-09521-4

Gunawan, F. E., Wanandi, K., Candra, S., Soewito, B., and Sekishita, N. (2017). "Detecting the early drop of attention using EEG signal," in *Proceedings of the International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, Yogyakarta, 19–21. doi: 10.1109/EECSI.2017.8239175

Hanslmayr, S., Klimesch, W., Sauseng, P., Gruber, W., Doppelmayr, M., Freunberger, R., et al. (2005). Visual discrimination performance is related to decreased alpha amplitude but increased phase locking. *Neurosci. Lett.* 375, 64–68. doi: 10.1016/j.neulet.2004.10.092

He, F., and Yang, Y. (2021). Nonlinear system identification of neural systems from neurophysiological signals. *Neuroscience* 458, 213–228. doi: 10.1016/j.neuroscience.2020.12.001

Hu, B., Li, X., Sun, S., and Ratcliffe, M. (2018). Attention recognition in EEG-based affective learning research using CFS+KNN algorithm. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 15, 38–45. doi: 10.1109/TCBB.2016.2616395

Jeong, D. H., and Jeong, J. (2020). In-ear EEG based attention state classification using echo state network. *Brain Sci.* 10:321. doi: 10.3390/brainsci10060321

Ke, Y., Chen, L., Fu, L., Jia, Y., Li, P., Zhao, X., et al. (2014). Visual attention recognition based on nonlinear dynamical parameters of EEG. *BioMed. Mater. Eng.* 24, 349–355. doi: 10.3233/BME-130817

Leonardi, N., Richiardi, J., Gschwind, M., Simioni, S., Annoni, J. M., Schluep, M., et al. (2013). Principal components of functional connectivity: a new approach to study dynamic brain connectivity during rest. *Neuroimage* 83, 937–950. doi: 10.1016/j.neuroimage.2013.07.019

Li, W., Ming, D., Xu, R., Ding, H., Qi, H., and Wan, B. (2013). "Research on visual attention classification based on EEG entropy parameters," in *Proceedings of the World Congress on Medical Physics and Biomedical Engineering May 26-31, 2012, Beijing, China. IFMBE Proceedings*, Vol. 39, ed. M. Long (Berlin: Springer), 1553–1556. doi: 10.1007/978-3-642-29305-4_408

Liu, J., Zhu, Y., Sun, H., Ristaniemi, T., and Cong, F. (2020). Sustaining attention for a prolonged duration affects dynamic organizations of frequency-specific functional connectivity. *Brain Topogr.* 33, 677–692. doi: 10.1007/s10548-020-00795-0

Magosso, E., Crescenzio, F. D., Ricci, G., Piastra, S., and Ursino, M. (2019). EEG alpha power is modulated by attentional changes during cognitive tasks and virtual reality immersion. *Comput. Intell. Neurosci.* 2019:7051079. doi: 10.1155/2019/7051079

McCusker, M. C., Wiesman, A. I., Schantell, M. D., Eastman, J. A., and Wilson, T. W. (2020). Multi-spectral oscillatory dynamics serving directed and divided attention. *NeuroImage.* 217:116927. doi: 10.1016/j.neuroimage.2020.116927

McNair, D. M., Lorr, M., and Droppleman, L. F. (1992). *POMS Manual for the Profile of Mood States*. San Diego, CA: Educational and Industrial Testing Service.

Misselhorn, J., Friese, U., and Engel, A. K. (2019). Frontal and parietal alpha oscillations reflect attentional modulation of cross-modal matching. *Sci. Rep.* 9:5030. doi: 10.1038/s41598-019-41636-w

Müller, M. M., Gruber, T., and Keil, A. (2000). Modulation of induced gamma band activity in the human EEG by attention and visual information processing. *Int. J. Psychophysiol.* 38, 283–299. doi: 10.1016/S0167-8760(00)00171-9

Pagnotta, M. F., Pascucci, D., and Plomp, G. (2020). Nested oscillations and brain connectivity during sequential stages of feature-based attention. *NeuroImage* 223:117354. doi: 10.1016/j.neuroimage.2020.117354

Palva, S., and Palva, J. M. (2007). New vistas for α-frequency band oscillations. *Trends Neurosci.* 30, 150–158. doi: 10.1016/j.tins.2007.02.001

Paneri, S., and Gregoriou, G. G. (2017). Top-down control of visual attention by the prefrontal cortex. Functional specialization and long-range interactions. *Front. Neurosci.* 11:545. doi: 10.3389/fnins.2017.00545

Pincus, S. M. (1991). Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci. U.S.A.* 88, 2297–2301. doi: 10.1073/pnas.88.6.2297

Putman, P., van Peer, J., Maimari, I., and van der Werff, S. (2010). EEG theta/beta ratio in relation to fear-modulated response-inhibition, attentional control, and affective traits. *Biol. Psychol.* 83, 73–78. doi: 10.1016/j.biopsycho.2009.10.008

Rezaeezadeh, M., Shamekhi, S., and Shamsi, M. (2020). Attention deficit hyperactivity disorder diagnosis using non-linear univariate and multivariate EEG measurements: a preliminary study. *Phys. Eng. Sci. Med.* 43, 577–592. doi: 10.1007/s13246-020-00858-3

Richman, J. S., and Moorman, J. R. (2000). Physiological time-series analysis using approximate and sample entropy. *Am. J. Physiol. Heart Circ. Physiol.* 278, H2039–H2049. doi: 10.1152/ajpheart.2000.278.6.H2039

Rossi, A. F., Pessoa, L., Desimone, R., and Ungerleider, L. G. (2009). The prefrontal cortex and the executive control of attention. *Exp. Brain Res.* 192, 489–497. doi: 10.1007/s00221-008-1642-z

Rosvold, H. E., Beck, L. H., Bransome, E. D. Jr., Mirsky, A. F., and Sarason, I. (1956). A continuous performance test of brain damage. *J. Consult. Psychol.* 20, 343–350. doi: 10.1037/h0043220

Schu, M. (1999). Are cognitive processes manifested in event-related gamma, alpha, theta and delta osciTheta wavelengths are between 4 – 8 Hertz and have been found to be associated with low levels of alertness as well as activities associated with cognitive processing. *Neurosci. Lett.* 259, 165–168.

Srinivasan, R., Thorpe, S., Deng, S., Lappas, T., D'Zmura, M., et al. (2009). "Decoding attentional orientation from eeg spectra," in *Human-Computer Interaction. New Trends. HCI 2009. Lecture Notes in Computer Science*, Vol. 5610, ed. J. A. Jacko (Berlin: Springer), 176–183. doi: 10.1007/978-3-642-02574-7_20

Stam, C. J. (2005). Nonlinear dynamical analysis of EEG and MEG: review of an emerging field. *Clin. Neurophysiol.* 116, 2266–2301. doi: 10.1016/j.clinph.2005.06.011

Szczepanski, S. M., Thorpe, S., Deng, S., Lappas, T., and D'Zmura, M. (2014). Dynamic changes in phase-amplitude coupling facilitate spatial attention control in Fronto-parietal cortex. *PLoS Biol.* 12:e1001936. doi: 10.1371/journal.pbio.1001936

Taya, F., Dimitriadis, S. I., Dragomir, A., Lim, J., Sun, Y., Wong, K. F., et al. (2018). Fronto-parietal subnetworks flexibility compensates for cognitive decline due to mental fatigue. *Hum. Brain Mapp.* 39, 3528–3545. doi: 10.1002/hbm.24192

Teplan, M. (2002). Fundamentals of EEG measurement. *Meas. Sci. Rev.* 2, 1–11.

Tononi, G., and Edelman, G. M. (1998). Consciousness and complexity. *Science* 282, 1846–1851. doi: 10.1126/science.282.5395.1846

van Schouwenburg, M. R., Zanto, T. P., and Gazzaley, A. (2017). Spatial attention and the effects of frontoparietal alpha band stimulation. *Front. Hum. Neurosci.* 10:658. doi: 10.3389/fnhum.2016.00658

Vigário, R. N. (1997). Extraction of ocular artefacts from EEG using independent component analysis. *Electroencephalogr. Clin. Neurophysiol.* 103, 395–404. doi: 10.1016/S0013-4694(97)00042-8

Wallace, J. C., Kass, S. J., and Stanny, C. J. (2002). The cognitive failures questionnaire revisited: dimensions and correlates. *J. Gen. Psychol.* 129, 238–256. doi: 10.1080/00221300209602098

Welch, P. D. (1967). The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE Trans. Audio Electroacoust.* 15, 70–73. doi: 10.1109/TAU.1967.1161901

Yang, Y., Dewald, J. P. A., van der Helm, F. C. T., and Schouten, A. C. (2018). Unveiling neural coupling within the sensorimotor system: directionality and nonlinearity. *Eur. J. Neurosci.* 48, 2407–2415. doi: 10.1111/ejn.13692

Zheng, J., Junsheng, C., Chen, M.-J., and Yang, Y. (2014). Multiscale fuzzy entropy and its application in rolling bearing fault diagnosis. *Zhendong Gongcheng Xuebao* 27, 145–151.

Zhu, Y., Liu, J., Ristaniemi, T., and Cong, F. (2020). Distinct patterns of functional connectivity during the comprehension of natural, narrative speech. *Int. J. Neural Syst.* 30, 1–15. doi: 10.1142/S0129065720500070

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.

frontiers | Frontiers in Human Neuroscience

ORIGINAL RESEARCH
published: 29 April 2022
doi: 10.3389/fnhum.2022.838338

Attentional Prioritization of Complex, Naturalistic Stimuli Maintained in Working-Memory–A Dot-Probe Event-Related Potentials Study

*Natalia Rutkowska, Łucja Doradzińska and Michał Bola**

Laboratory of Brain Imaging, Nencki Institute of Experimental Biology, Polish Academy of Sciences, Warsaw, Poland

Recent studies suggest that a stimulus actively maintained in working memory (WM) automatically captures visual attention when subsequently perceived. Such a WM-based guidance effect has been consistently observed for stimuli defined by simple features, such as color or orientation, but studies using more complex stimuli provided inconclusive results. Therefore, we investigated whether the WM-based guidance effect occurs also for naturalistic stimuli, whose identity is defined by multiple features and relations among them, specifically for faces and houses. The experiment consisted of multiple blocks in which participants (N = 28) either memorized or merely saw (WM or exposure condition, respectively) a template stimulus and then performed several dot-probe trials, with pairs of stimuli (template and control) presented laterally as distractors and followed by a target-asterisk. Analysis of reaction-times (RT) in the dot-probe task shows that memorized stimuli were prioritized by attention and points toward attention-hold, rather than capture, as a mechanism of attentional prioritization. Consistent with this interpretation, memorized items did not evoke a lateralized N2pc ERP component, thought to indicate attention shifts. However, in an exploratory ERP analysis we found evidence for a very early (100–200 ms post-stimulus) prioritization specific to the memorized faces, which is in line with the sensory recruitment theory of WM. In conclusion, our data provide evidence that complex stimuli are prioritized by attention when maintained in WM, and that the mechanism of such prioritization is based on a prolonged hold of spatial attention.

Keywords: attention, dot-probe, naturalistic stimuli, N2pc, working memory

OPEN ACCESS

Edited by:
Lutz Jäncke,
University of Zurich, Switzerland

Reviewed by:
Summer Sheremata,
Florida Atlantic University, United States
Yu-Chin Chiu,
Purdue University, United States

****Correspondence:***
Michał Bola
m.bola@nencki.edu.pl

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

Received: *17 December 2021*
Accepted: *17 March 2022*
Published: *29 April 2022*

Citation:
Rutkowska N, Doradzińska Ł and Bola M (2022) Attentional Prioritization of Complex, Naturalistic Stimuli Maintained in Working-Memory–A Dot-Probe Event-Related Potentials Study. Front. Hum. Neurosci. 16:838338. doi: 10.3389/fnhum.2022.838338

INTRODUCTION

Contemporary theories of memory emphasize its role in the prospective guidance of perception and action (Nobre and Stokes, 2019). Particularly, the working memory (WM) system is currently recognized as the key component of a pro-active, top-down selection mechanism (Desimone and Duncan, 1995; Nobre and Stokes, 2019). While WM plays an important role in the volitional control of attention, it influences attentional selection also in an involuntary way. Specifically, stimuli encoded and actively maintained in visual WM automatically attract attention upon a

subsequent presentation (review: Soto et al., 2008). This effect has been revealed, first, by dot-probe experiments, in which responses to probes presented at the location of the WM-maintained stimulus were faster, than to probes following an unfamiliar, control stimulus (Downing, 2000). Second, by visual search experiments, showing that search times are increased when a WM-maintained stimulus appears in the search array as a distractor (e.g., Soto et al., 2005; Olivers et al., 2006; Soto and Humphreys, 2007, 2009). Third, by eye-tracking studies indicating that eye-movements are automatically attracted by visual items matching the WM content (Hollingworth et al., 2013; Schneegans et al., 2014; Silvis and Van der Stigchel, 2014). Finally, by electrophysiological experiments revealing that the WM-maintained stimuli evoke an N2pc component, which is a classic index of covert attention shifts (Kumar et al., 2009; Carlisle and Woodman, 2013). Importantly, such a WM-based attention-capture effect is not a form of priming, as it was observed only when a stimulus was actively maintained in WM, but not when it was merely seen; and is considered automatic and involuntary, as it occurred even when detrimental to the task performance (review: Soto et al., 2008).

The automatic guidance of attention from WM has been so far demonstrated mainly with the use of simple stimuli, defined either by color or orientation (review: Soto et al., 2008). Stimuli varying on a single dimension of one basic feature are generally most effective in guiding bottom-up attention, as they can be processed pre-attentively and result in a pop-out search (Wolfe and Horowitz, 2017). Therefore, a question arises whether complex stimuli, whose identity is typically defined by multiple features and relations among them, are able to cause a similar WM-based attention guidance effect. The seminal dot-probe study by Downing (2000) revealed that images of faces, abstract geometric shapes, and line drawings of real life objects captured attention when maintained in WM. However, subsequent experiments using visual search paradigms did not replicate these findings. First, Houtkamp and Roelfsema (2006) used drawings of real-life objects as stimuli in a visual search task and found no evidence–neither in accuracy or RT scores, nor in the eye-tracking data–indicating that such complex items captured attention when held in WM. Second, Downing and Dodds (2004) also used a visual search task with complex artificial shapes as stimuli and found no interference on search performance–as indexed by accuracy and RT in response to targets–from items concurrently held in WM. Third, Peters et al. (2009) conducted an event-related potentials (ERP) study–also using complex artificial shapes as WM items–and showed not only that the memory-matching distractors did not influence the search task accuracy, but also that complex stimuli held in WM do not differ from control stimuli in terms of evoked ERP activity. Finally, a key study was conducted by Zhang et al. (2010), who used the same procedure as Peters et al. (2009) to directly compare two different sets of stimuli–one consisting of simple shapes [previously used by Soto et al. (2005)], and another of more complex, artificial shapes [used by Downing and Dodds (2004)]. Zhang et al. (2010) found that simple stimuli captured attention when maintained in WM, but complex stimuli did not cause such an effect. Therefore, they concluded that the WM-based attentional guidance critically depends on the stimulus features, with only the simple ones being effective in guiding attention from visual WM.

In light of such conflicting findings–with the classic study of Downing (2000) supporting the WM-based attention-capture by complex stimuli, but several subsequent studies challenging his conclusions (Downing and Dodds, 2004; Houtkamp and Roelfsema, 2006; Peters et al., 2009; Zhang et al., 2010)–we designed a study to provide further evidence, either in favor or against the discussed effect. Specifically, we investigated whether two types of complex, naturalistic stimuli–images of faces and houses–are prioritized by attention when maintained in visual WM. We chose these two categories as both are defined by multiple features and thought to exhibit similar levels of complexity, and thus are often compared in visual perception studies (Filliter et al., 2016). In the conducted experiment images of faces and houses were used as template stimuli, which were either memorized for later recollection (WM condition) or merely seen without the need to memorize (exposure condition), in separate blocks. Next, within each block participants performed a sequence of dot-probe trials, in which the target dot followed either a template (congruent trials) or a control image (incongruent trials), which were presented as task-irrelevant distractors (MacLeod et al., 1986). We analyzed a difference in reaction-times (RT) between congruent and incongruent trials, which is a primary index of attention capture in the dot-probe task, and compared it between experimental conditions. Further, to obtain a measure of attentional prioritization that is time-resolved and independent of behavioral response, we recorded electro-encephalographic (EEG) activity and analyzed two lateralized ERP components. First, the N2 posterior contralateral (N2pc), which is considered to represent covert attention shifts (Luck and Hillyard, 1994; Eimer, 1996; Kiss et al., 2008). N2pc is defined as more negative amplitude of signals recorded at posterior electrodes contralateral to the presented stimulus (in comparison to ipsilateral electrodes) in the early time-window (starting c.a. 200 ms after stimulus onset). Second, the Sustained Posterior Contralateral Negativity (SPCN), defined in a similar manner but occurring later, around 300–400 ms after stimulus onset, and thought to reflect maintenance and manipulation of information in visual WM (Jolicæur et al., 2008; Emrich et al., 2009; Clark et al., 2015).

We hypothesized that template stimuli will attract attention in the WM condition–as indicated by shorter RT in the congruent dot-probe trials, and by greater amplitude of the N2pc ERP component–but we did not expect to observe these effects in the mere exposure condition. Further, because we assumed the attentional prioritization of WM items to be automatic and involuntary, we did not expect to observe the SPCN ERP component, neither in the WM, nor in the exposure condition. Finally, considering preferential processing of faces in the visual system, in comparison to houses, we expected them to benefit from attentional and memory advantage (in line with: Farah et al., 1998; Tsao and Livingstone, 2008; Curby et al., 2009).

MATERIALS AND METHODS

Participants

The study was conducted with the approval of the human ethics committee of the SWPS University of Social Sciences and Humanities (Warsaw, Poland). All participants declared normal or corrected-to-normal vision and no history of mental or neurological disorders. All participants provided written informed consent and received monetary compensation for their time (100 PLN = c.a. 25 EUR).

We analyzed data of 28 participants (18 females, mean age = 24.2, SD = 2.59 years, range: 19–28, 2 left-handed). Data of six additional participants were collected, but they were excluded from the analysis: four participants due to the technical problems during an EEG recording procedure; one participant did not comply with the dot-probe task instruction; and one participant due to insufficient number of epochs remaining after EEG signal pre-processing (detailed criteria are described in the Section "Electro-Encephalographic Recording and Analysis").

Stimuli

Two sets of stimuli were used. First, 60 pictures of faces with neutral expression (30 male, 30 female; all Caucasian) selected from the Karolinska Directed Emotional Faces stimulus set (KDEF; Lundqvist et al., 1998). From their original format, the face photographs were converted to grayscale using the Gnu Image Manipulation Program (GIMP[1]). Second, 60 pictures of houses from the DalHouses stimulus set (Filliter et al., 2016). All houses' images were originally in grayscale, presented on a white background, thus no modifications were introduced. Identifiers of stimuli used in the present study can be found in the project description at OSF.[2]

Procedure

The experimental procedure was written in the Presentation software (Neurobehavioral Systems, Albany, CA, United States) and presented on a FlexScan EV-2450 (Hakusan, Ishikawa, Japan) screen through an Intel Core i3 computer. Participants were seated comfortably in a dimly lit room with a viewing distance of 57 cm, which was maintained by a chinrest.

The procedure started with a display providing participants with the task instructions and information about the trial structure. The procedure consisted of two tasks: a working-memory (WM) task and a mere exposure task; and involved two stimuli types: faces and houses. Thus, there were four conditions: a face WM condition; a house WM condition; a face exposure condition; and a house exposure condition; which were presented to participants as separate blocks, the order of which was randomized. Each condition was further sub-divided into 32 memory or exposure blocks. For each memory/exposure block one template and one control stimulus were randomly chosen from the pool of all houses or faces. Additionally, in the face conditions the face stimuli were gender-matched, i.e., female and male template images were paired only with, respectively, female and male control images. In both, a face WM condition and a face exposure condition, female faces were used in half of the blocks, and male faces in the other half. All stimuli were presented against a black background.

Each of those 32 blocks started with a central presentation of a template stimulus for 5000 ms (**Figure 1**). The instruction–either "Memorize this picture" (in the WM condition) or "Take a look at this picture" (in the exposure condition)–was displayed above the image. The face images subtended 7.4° × 10.0° of the visual angle, while house images varied in size and subtended from 6.1° to 9.3° × 7.8° of the visual angle. After the display of the template stimulus, a sequence of dot-probe trials was presented. Each dot-probe trial started with a fixation cross (subtending 0.9° × 0.9° of the visual angle) displayed in the center of the screen. The fixation cross remained on-screen throughout the trial. After 1000 ms a pair of stimuli were presented bilaterally for 200 ms–the template stimulus on one side and control stimulus on the other. Face stimuli were presented with their inner edge 4.4° left and right from the fixation cross, while house stimuli with the inner edge from 3.8° to 5.3° left and right from the fixation cross. Next, a target asterisk subtending 0.7° × 0.7° of the visual angle was presented for 150 ms in the location of the center of either the template stimulus (congruent trial) or the control one (incongruent trial). Participants were instructed to maintain their gaze on the centrally presented fixation cross, ignore the laterally appearing stimuli, and indicate the side of the target asterisk presentation (left or right) by pressing one of two buttons using index fingers of their left or right hand. Participants were asked to respond as quickly and accurately as possible. The response time to the target asterisk was limited to 3000 ms and the next trial started immediately after the manual response. Within each dot-probe sequence the template stimulus was presented on the left side in half of the dot-probe trials, and on the right side in the other half. Further, half of the dot-probe trials were congruent and half were incongruent. The order of trials within each sequence was randomized.

In the WM condition, participants were given a memory test at the end of each memory block (i.e., after completing the dot-probe trials sequence). Either a template image or a different image (neither a template, nor a control stimulus) was presented centrally. Participants had to indicate whether the presented stimulus is the same or different from the one they were previously asked to memorize. Above the image a question "Is this the same image?" was displayed. In half of the WM blocks the stimulus was the template image (correct answer "yes"), and in the other half it was a different image (correct answer "no"). The answer "yes" was displayed in the left corner of the screen and the answer "no" in the right corner. Participants responded by pressing one of two buttons (left for "yes," right for "no"). The response time in the memory test was limited to 5000 ms and the next WM block started immediately after the manual response. In the mere exposure condition participants were not tested for stimulus recognition, but immediately after completion of the dot-probe sequence the next exposure block started.

In total 192 dot-probe trials were presented per condition. Within each condition half of the WM/exposure blocks

[1] Available at http://www.gimp.org/

[2] https://osf.io/9rc4j/

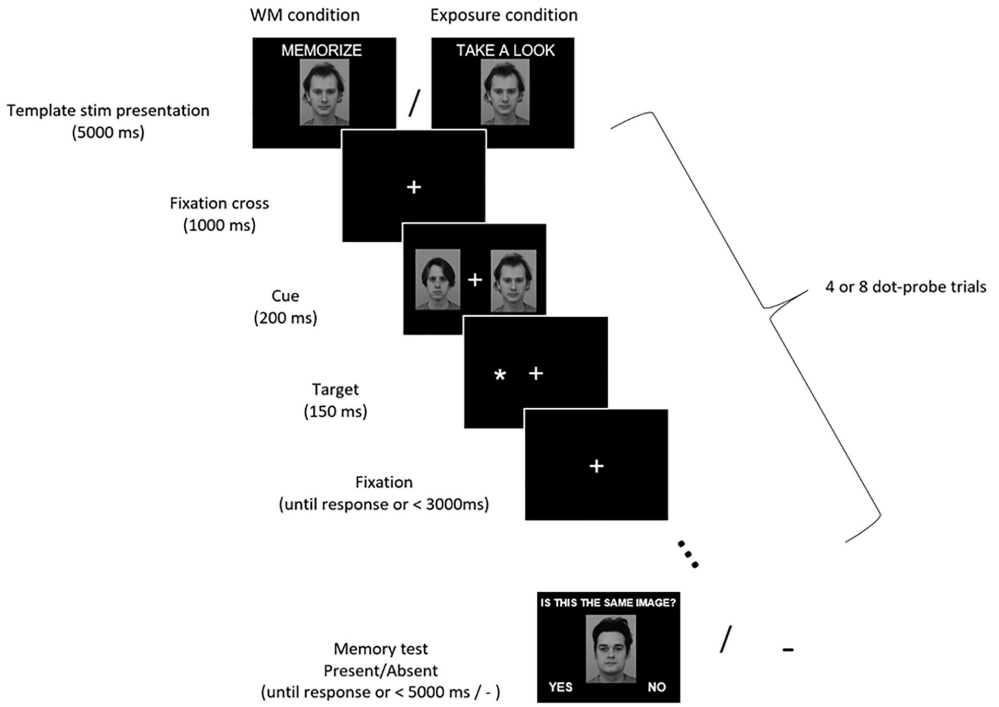


FIGURE 1 | Experimental procedure. Participants were asked to memorize (WM condition) or just look (exposure condition) at the presented stimulus (either a face or a house). Then they performed 4 or 8 dot-probe trials in which their task was to indicate with a button press the location of a target asterisk (left or right). Asterisks were preceded by pairs of stimuli, which participants were supposed to ignore. A memorized/seen stimulus was presented on one side, and a control one on the other side. In the WM condition participants completed a memory test, in which they had to decide if the presented stimulus is the one they were instructed to memorize (in the exposure condition there was no memory test). Images from KDEF stimulus set reproduced with permission from Lundqvist et al. (1998). IDs of images included in the Figure: AM13NES and AM22NES.

comprised 4 dot-probe trials, and the other half comprised 8 dot-probe trials (the order of blocks was random). The number of trials varied in order to prevent participants from expecting the exact moment of a memory test, and thus encourage them to maintain the WM active throughout the block. Participants had a self-paced break five times per condition.

Analysis of Behavioral Data

All analyses of behavioral data were conducted using custom-made Python scripts. Accuracy of responses to the presentation side of the target-dot was calculated as a percentage of correct responses. The obtained values are presented in the Results section, but due to ceiling level performance in the majority of participants this measure was not analyzed statistically. Therefore, analysis of the dot-probe task data was focused on establishing whether reaction times (RT) of manual responses to the target asterisk differ between two types of trials: those in which the asterisk was presented on the same side as the potentially attention-grabbing stimulus (memorized/seen face or house; congruent trials) and those in which the asterisk was presented on the neutral stimulus side (control face or house; incongruent trials). Mean reaction times were calculated only for the correct responses. For the WM condition the accuracy of memorizing a template stimulus was calculated as a percentage of correct responses in the memory test.

Electro-Encephalographic Recording and Analysis

During the experiment, EEG signal was recorded with 64 Ag-AgCl electrically shielded electrodes mounted on an elastic cap (ActiCAP, Munich, Germany) and positioned according to the extended 10–20 system. Vertical (VEOG) and horizontal (HEOG) electro-oculograms were recorded using bipolar electrodes placed at the supra- and sub-orbit of the right eye and at the external canthi. Electrode impedances were kept below 10 kΩ. The data were amplified using a 128-channel amplifier (QuickAmp, Brain Products, Enschede, Netherlands) and digitized with BrainVisionRecorder® software (Brain Products, Munich, Germany) at a 500 Hz sampling rate. The EEG signal was recorded against an average of all channels calculated by the amplifier hardware.

Electro-encephalographic and EOG data were analyzed using EEGlab 14 functions and Matlab 2016b. First, all signals were filtered using a high-pass (0.5 Hz) and a low-pass (45 Hz) Butterworth IIR filter (filter order = 2; Matlab functions: *butter* and *filtfilt*). Then data were re-referenced to the average of signals recorded from left and right earlobes, and down-sampled to 250 Hz. All data were divided into 768 dot-probe epochs (192 epochs per condition; [−200, 1200] ms with respect to the faces/houses images onset) and the epochs were baseline-corrected by subtracting the mean of the pre-stimulus period

(i.e., [−200, 0 ms]). Further, epochs were rejected based on the following criteria (all values Mean ± SEM): (i) when there was no manual response to the target dots until 1.2 s after the onset (18.9 ± 6.0; range [0, 120] epochs per participant); (ii) when activity of the HEOG electrode in the time-window [−200, 500] ms exceeded −40 or 40 uV (105.5 ± 19.1; range [11, 352] epochs per participant); (iii) when activity of the P7 or P8 electrode in the time-window [−200, 600] ms exceeded −80 or 80 uV (none of the epochs rejected). Thus, after applying the described criteria the average number of analyzed epochs per participant was: 643.6 ± 22.0; range [355, 752].

A participant was excluded if the number of epochs in any condition was <60. This criterion resulted in excluding 1 participant out of 29 (but additional 5 participants were excluded due to other criteria, as described in the Section "Participants"). The numbers of epochs provided above were calculated based on the final sample of 28 participants.

Next, each EEG-EOG data-set was decomposed into 50 components using Independent Component Analysis as implemented in the EEGlab *pop_runica* function. To remove residual oculographic artifacts from the data the following procedure was used: time-course of each component was correlated with time-courses of HEOG and VEOG electrodes and in case the Spearman correlation coefficient exceeded −0.3 or 0.3 a component was subtracted from the data. Using this procedure 3.0 ± 0.2 components (range [1, 6]) per participant were removed.

To ensure that there was no difference in the number of trials with the target presented on the WM (congruent) and control side (incongruent) which were retained in the data after preprocessing, we conducted a 3-way repeated-measures ANOVA with *congruency* (congruent, incongruent trials), *task* (WM, exposure), and *stimulus* (face, house) as factors. We confirmed that *congruency* had neither main effect [$F_{(1,27)} = 1.386$, $p = 0.249$, $\eta_p^2 = 0.049$], nor interacted with other factors [*congruency* × *task*: $F_{(1,27)} = 0.649$, $p = 0.428$, $\eta_p^2 = 0.023$, *congruency* × *stimulus*: $F_{(1,27)} = 0.042$, $p = 0.839$, $\eta_p^2 = 0.002$].

Both N2pc and SPCN components are defined as a difference between the contralateral and ipsilateral activity recorded at posterior electrodes after a stimulus presentation (Luck and Hillyard, 1994; Jolicæur et al., 2008). As N2pc is typically maximal at posterior electrodes sites (Luck, 2012), signals from P8 and P7 electrodes were used to calculate both components, similarly to other studies using N2pc and SPCN as markers of attention shifts (Kappenman et al., 2014; Furtak et al., 2020; Bola et al., 2021). Epochs were divided with respect to the condition and presentation side of the template stimulus in the following way: when template stimulus was presented on the left side, P8 was the contralateral electrode and P7 was the ipsilateral electrode; when template stimulus was presented on the right side, P7 was the contralateral electrode and P8 was the ipsilateral electrode. For each condition contralateral and ipsilateral signals were first concatenated and then averaged, to obtain contralateral and ipsilateral waveforms. These waveforms were averaged within the 200–400 ms time-window for the N2pc analysis (e.g., Woodman et al., 2009; Reutter et al., 2017; Wójcik et al., 2019) and in the 400–600 ms time-window for the SPCN analysis (Clark et al., 2015). Further, based on the visual inspection of the obtained ERP waveforms (**Figure 3**) an exploratory analysis was conducted on lateralized activity observed in the early, 100–200 ms time-window.

Statistical Analysis

Statistical analyses were conducted in the JASP software and cross-checked with Statcheck.[3] The values are reported as Mean ± SEM, unless stated otherwise. For all statistical tests probability values were reported (p) and the standard 0.05 alpha level was used as a threshold for refuting the null hypothesis.

To test for the presence of the behavioral (RT) and electrophysiological dot-probe task effects repeated–measures (rm) ANOVA models were used. Specific to the analysis of RT was the factor of *congruency*, defined by the asterisk presentation side with respect to the memorized/seen item (congruent vs. incongruent trials). Specific to the electrophysiological analysis was the factor of *side*, defined as the side on which ERP activity was recorded, with respect to the memorized/seen item (ipsi- vs. contra-lateral). The *side* effect was analyzed separately for activity recorded in the three analyzed time-windows (100–200 ms, 200–400 ms, and 400–600 ms). The factors of a *stimulus* (faces vs. houses) and *task* (memory vs. mere exposure) were included in all models. The simple main effects analyses were conducted in case of a significant interaction. Results were reported as F(df) and partial eta-squared, the indicator of the effect size, was reported as η_p^2.

To compare the accuracy scores in the memory test between faces and houses conditions, the data distribution was first tested with the Shapiro–Wilk test and, as it deviated from normality, a non-parametric two-tailed Wilcoxon test was used. The statistic was reported as a sum of positive ranks (W), together with the matched rank biserial correlation (r_{rb}) as a measure of the effect size.

Data Availability

The data used in the statistical analysis can be accessed from the OSF (see text footnote 2). Raw EEG data, and scripts used for presentation of the experimental procedure and data analysis will be shared by authors per request.

RESULTS

Memory Accuracy

We observed high working-memory (WM) accuracy scores for both faces (97.2 ± 0.7%) and houses (94.2 ± 1.6%), which indicates that participants were actively maintaining the template stimulus in working memory. Comparing the WM accuracy between face and houses we found better memory performance for face images ($W = 172.500$, $p = 0.010$, $r_{rb} = -0.150$). This result confirms our hypothesis and is in line with previous studies (Curby et al., 2009), but due to the ceiling level performance and low effect size it should be treated with caution.

[3]http://statcheck.io/index.php

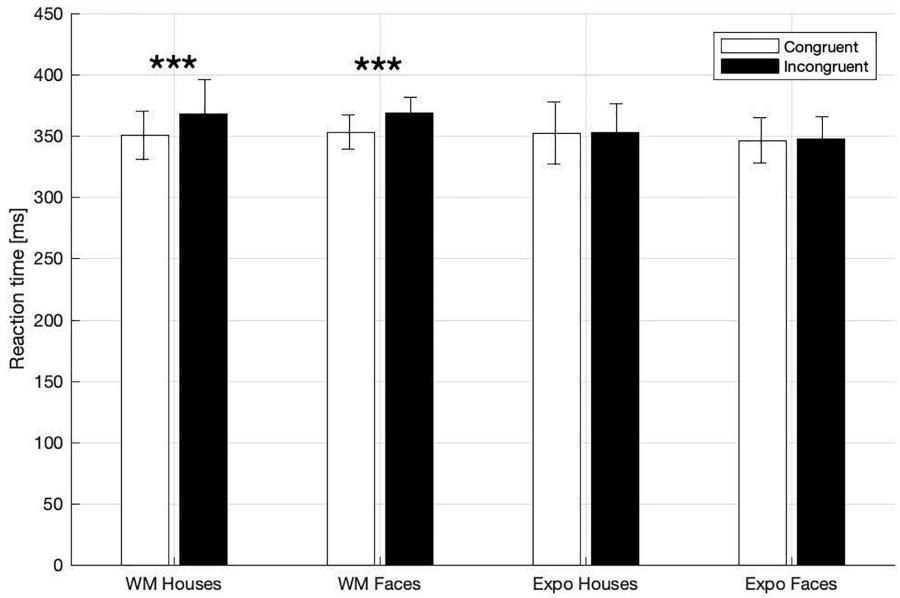


FIGURE 2 | Mean RTs in dot-probe task for the WM and exposure conditions and both types of stimuli (House, Face). Asterisks indicate statistically significant differences between congruent and incongruent trials ($p < 0.001$). Error bars represent 2SEM.

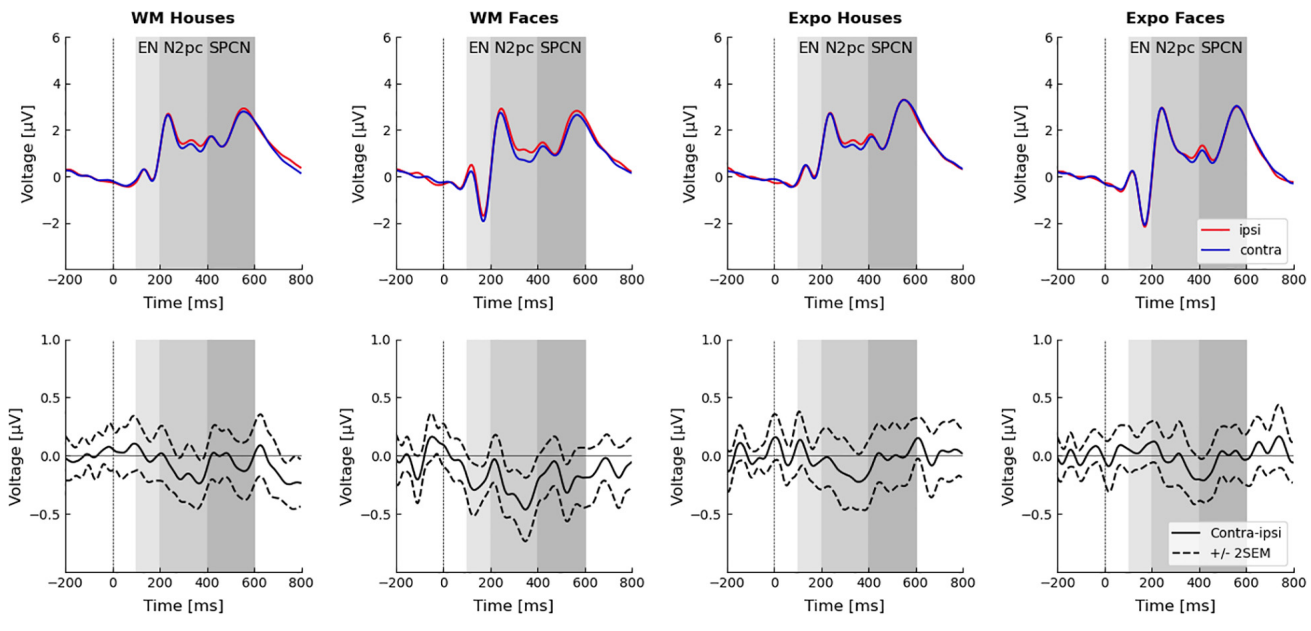


FIGURE 3 | Event related potentials in the dot-probe task. Electrodes P7/P8 were chosen for the analysis. Waveforms recorded ipsi- and contra-laterally with respect to the seen or memorized stimulus are presented in the top row. Difference waveforms (i.e., contra–ipsi-lateral side) are presented in the bottom-row. Time windows of the three analyzed components–the early negativity (EN; 100–200 ms), N2pc (200–400 ms), and SPCN (400–600 ms)–are highlighted.

Dot-Probe Task–Behavioral Results

In the dot-probe task participants exhibited ceiling-level accuracy (i.e., in indicating the target-dot presentation side), with the percentage of correct responses being: 97 ± 0.6% in the house WM condition; 97 ± 0.6% in the face WM condition; 97 ± 0.8% in the house exposure condition; 98 ± 0.5% in the face exposure condition. Correct responses to the target asterisk were analyzed in terms of reaction times (RT), which was our primary index of attention capture. Specifically, we investigated whether RT were shorter when the target

followed a potentially attention-grabbing template stimulus (i.e., congruent trials), in comparison to trials when it followed a control stimulus (i.e., incongruent trials; **Figure 2**). In a three-way rm-ANOVA analysis we found a significant main effect of *congruency* [$F_{(1,27)}$ = 15.49, $p < 0.001$, η_p^2 = 0.365] and *task* [WM vs. exposure; $F_{(1,27)}$ = 9.05, p = 0.006, η_p^2 = 0.251], and a significant interaction between those two factors [$F_{(1,27)}$ = 14.24, p = 0.001, η_p^2 = 0.345]. With regard to this interaction there was a significant simple main effect of *congruency* in the memory condition [congruent trials: 352.06 ± 8.67 ms; incongruent 368.54 ± 10.86 ms; $F_{(1,27)}$ = 21.42, $p < 0.001$, η_p^2 = 0.442], but not in the exposure condition [congruent trials: 349.33 ± 11.14 ms; incongruent 350.55 ± 10.56 ms; $F_{(1,27)}$ = 0.265, p = 0.611, η_p^2 = 0.010]. The main effect of *stimulus* [$F_{(1,27)}$ = 0.052, p = 0.822, η_p^2 = 0.002] and other interactions [*task* × *stimulus*: $F_{(1,27)} = 0.830$, $p = 0.370$, $\eta_p^2 = 0.03$, *stimulus* × *congruency:* $F_{(1,27)}$ = 0.033, p = 0.858, η_p^2 = 0.001, *congruency* × *task* × *stimulus:* $F_{(1,27)}$ = 0.12, p = 0.746, η_p^2 = 0.004] did not reach significance. Therefore, in line with our hypothesis, we found attentional prioritization of stimuli that were actively maintained in WM (but not of stimuli that were merely seen), and such a WM-based effect was observed irrespective of the stimulus type.

Dot-Probe Task–Electrophysiological Results

Results of the three-way rm-ANOVA for the N2pc ERP component indicate a significant *side* effect (ipsilateral amplitude: 1.64 ± 0.44 μV; contralateral: 1.49 ± 0.44 μV), but interactions between *side* and *task,* or *side* and *stimulus* were not significant (**Figure 3** and **Table 1**). Therefore, we conclude that our manipulations did not affect the N2pc component, and that our hypothesis stating that larger N2pc will be observed in the WM task was not confirmed.

Similarly, results of the three-way rm-ANOVA for the SPCN ERP component indicate a significant *stimulus* effect only. This effect indicates that activity evoked by faces (1.77 ± 0.37 μV) had lower amplitude than activity evoked by houses (2.15 ± 0.33 μV). However, because none of the interactions involving the *side* factor was significant, we conclude that our manipulation did not affect the SPCN component, which is in line with our hypothesis.

Inspection of the obtained ERP waveforms prompted us to conduct an unplanned, exploratory analysis of the lateralized activity in an earlier, 100–200 ms time-window. This analysis yielded a significant interaction between *side*, *task*, and *stimulus* (**Figure 3** and **Table 1**). The simple main effects analysis showed that an early contralateral negativity is present in the memory condition for faces [ipsilateral: −0.51 ± 0.39 μV; contralateral: −0.74 ± 0.38 μV; $F_{(1,27)}$ = 10.13, p = 0.004, η_p^2 = 0.273], but not for houses [ipsilateral: 0.15 ± 0.43 μV; contralateral: 0.15 ± 0.44 μV; $F_{(1,27)}$ = 0.004, p = 0.951, $\eta_p^2 < 0.001$]. In the mere exposure conditions the simple main effects were not significant neither for faces [ipsilateral: −0.85 ± 0.42 μV; contralateral: −0.79 ± 0.41 μV; $F_{(1,27)}$ = 1.02, p = 0.322, η_p^2 = 0.036] nor for houses [ipsilateral: 0.23 ± 0.42 μV; contralateral: 0.23 ± 0.43 μV; $F_{(1,27)}$ = 0.01, p = 0.971, $\eta_p^2 < 0.001$]. Therefore, our analysis revealed that the very early lateralized activity was evoked only by face images maintained in WM.

DISCUSSION

The present study examined whether complex naturalistic stimuli that are actively maintained in visual WM are automatically prioritized by attention upon subsequent presentation. Such a WM-guided attentional selection has been consistently observed for simple stimuli (review: Soto et al., 2008), but inconsistent findings were reported when more complex stimuli were used (Downing, 2000; Downing and Dodds, 2004; Houtkamp and Roelfsema, 2006; Peters et al., 2009; Zhang et al., 2010). Therefore, in the present study images of faces and houses were either memorized or merely seen by participants, and subsequently presented in a dot-probe task. To test a hypothesis that such complex, naturalistic stimuli are automatically prioritized by attention when actively maintained in WM we analyzed both behavioral (RT) and electrophysiological (N2pc) indexes of attentional prioritization.

Attentional Prioritization of Complex Stimuli

Our main finding from the dot-probe task is that RTs were significantly shorter when a target asterisk followed a memorized face or a house, in comparison to the situation when it followed a control stimulus. Importantly, the RT effect was not observed, neither for faces nor for houses, when participants merely saw the template images, without an instruction to memorize them. This indicates that stimuli actively maintained in WM were prioritized by attention, but those that were merely seen did not cause such an effect. Thus, our behavioral results replicate the findings of Downing (2000), who found a similar effect for a range of complex stimuli–including complex artificial shapes, line drawings, and images of faces–in a similar dot-probe task. However, his findings were challenged by subsequent visual search experiments, which used drawings of real-life objects and complex shapes and did not find a similar effect (Downing and Dodds, 2004; Houtkamp and Roelfsema, 2006; Peters et al., 2009). Of special relevance is work by Zhang et al. (2010), who compared two sets of artificial shapes–simple and complex ones–directly in the same study and using the same visual search task. They found only the simple stimuli to be prioritized by attention when maintained in WM and concluded that the attentional guidance critically depends on the stimulus features, with a stronger effect for simple than for complex stimuli. Thus, their conclusion is at odds with results of both Downing (2000) and our study, which show that even stimuli defined by multiple features and by relations among them can induce a WM-based attentional bias. Importantly, our data provide evidence that the mere complexity of a stimulus is not likely to be a critical factor in the investigated phenomenon. We rather argue that the images used in our study evoked the WM-based guidance effect because they were processed holistically and perceived as meaningful, in contrast to the complex artificial shapes [e.g.,

TABLE 1 | Rm-ANOVA analysis of the electrophysiological effects.

	Time window: 100–200 ms			Time window: 200–400 ms			Time window: 400–600 ms		
Factor	***F*(1,27)**	***p***	**η_p^2**	***F*(1,27)**	***p***	**η_p^2**	***F*(1,27)**	***p***	**η_p^2**
Side	1.556	0.223	0.054	10.194	0.004	0.274	1.717	0.201	0.060
Task	0.661	0.423	0.024	0.032	0.860	0.001	0.587	0.450	0.021
Stimulus	31.910	<0.001	0.542	1.023	0.321	0.037	10.061	0.004	0.271
Side × Task	4.923	0.035	0.154	1.899	0.180	0.066	1.414	0.245	0.050
Side × Stim	0.066	0.321	0.036	0.069	0.795	0.003	0.672	0.420	0.024
Task × Stim	2.341	0.138	0.080	0.027	0.871	0.001	0.394	0.536	0.014
Side × Task × Stim	4.923	0.035	0.154	2.032	0.165	0.070	0.224	0.640	0.008

Analysis was conducted separately for three time windows (early negativity: 100–200 ms; N2pc: 200–400 ms; and SPCN: 400–600 ms). Each model included the following three factors: side (recording from contralateral/ipsilateral electrodes), task (memory/exposure), stimulus (face/house).

used by Zhang et al. (2010)]. Our interpretation is in line with Xu (2017), arguing that visual WM is not typically used to encode features of a single dimension, but rather to store integrated representations of meaningful objects.

Notably, apart from differences in stimuli properties, the discussed studies differ also in terms of paradigms used–while the studies demonstrating a WM-based prioritization of complex stimuli used a dot-probe task (Downing, 2000, and our study), studies finding no evidence for such an effect used a visual search task (Downing and Dodds, 2004; Houtkamp and Roelfsema, 2006; Peters et al., 2009; Zhang et al., 2010). Potential differences in sensitivity between these two procedures might thus account, at least partially, for differences in reported results. More data has to be collected by future studies to address which factor plays a key role in the discussed effect.

Two opposing accounts of the mechanism behind the WM-based guidance effect have been proposed. First one emphasizes the role of verbal (and perhaps semantic) representations in WM maintenance and subsequent directing of visual attention. It is based on studies showing that verbalization by itself can induce attention guidance (Soto and Humphreys, 2007), and that when visual stimuli are used as memory items the articular suppression task impairs the guidance effect (Downing and Dodds, 2004; Woodman and Luck, 2007; Soto and Humphreys, 2008). In contrast, the second view assumes that the guidance effect relies predominantly on visual representations. It is supported by experiments revealing the effect only when stimuli were defined by small and hard to verbalize differences in their attributes (e.g., hues of one color or slightly differing shapes) but not when easy to verbalize categorical differences were used (Olivers et al., 2006). Importantly, the stimuli used in our study were also difficult to verbalize and required maintaining a predominantly visual representation. Thus, our results provide further support for the latter view.

Attentional Prioritization–Capture or Hold?

The majority of previous studies investigated the WM-based guidance effect using behavioral methods, and thus relevant EEG or fMRI data is scarce. Therefore, in the present study we collected EEG data, with the main aim of using them to evaluate the time-course of attentional prioritization. However, the N2pc ERP component–a classic index of attention capture (Luck and Hillyard, 1994; Eimer, 1996; Kiss et al., 2008)–was not affected by the memory manipulation. Thus, in our study we found a robust behavioral (RT) effect of attentional prioritization, but at the same time no related electrophysiological effect in the form of N2pc. Lack of N2pc is thus at odds with previous studies that found N2pc in similar WM tasks (Kumar et al., 2009; Carlisle and Woodman, 2013), albeit it is important to emphasize these previous studies used simple stimuli. Importantly, the series of previous studies have shown that N2pc can be observed when stimuli maintained in WM are subsequently presented as task-irrelevant distractors [Kumar et al., 2009; but see also Carlisle and Woodman (2011)], but its amplitude is four times higher when stimuli constitute task-relevant targets (Carlisle and Woodman, 2013). Thus, while absence of the N2pc component in our study might be related to the fact that complex stimuli were used, it might also stem from task-irrelevance of the memory items. Future studies will investigate the effect of naturalistic stimuli in situations when they are task-relevant, which would also more closely reflect daily life situations.

Importantly, the classic theory of attention proposed by Posner et al. (1987) differentiates two independent functionalities of spatial attention orienting–attention shifts (or capture) defined as movement of attention from its current location to a new one; and attention engagement (or hold) described as an involvement in processing of a stimulus and a transient inability to switch to a new location. Therefore, another potential explanation of the dissociation between RT and N2pc observed in our study is that the memorized stimuli did not automatically capture attention (thus no N2pc effect was observed), but rather held and engaged attention for a longer time. What further supports this interpretation is that we observed an elongation of RT in the incongruent trials, rather than shortening of RT in the congruent ones–this is evident when the WM and exposure conditions are compared (**Figure 2**). Previous visual search studies using simple stimuli and including valid, neutral, and invalid conditions have provided conflicting results on the matter of capture vs. hold. Some found longer RTs in invalid as relative to neutral trials, but no difference between valid and neutral conditions (which would be indicative of an attention-hold by memorized items; Soto et al., 2007). However, others show both shorter RT in valid

and longer RT in invalid trials, in comparison to the neutral ones (which would be indicative of both capture and hold; Soto et al., 2006). Importantly, data collected by Downing (2000) do not speak to the matter at hand because of the between-subject design used in his study (the group performing the control task had in general significantly longer RTs to probes than the group performing WM task). Thus, further studies are required to elucidate the precise mechanism of attentional prioritization of the WM-maintained items.

Early Prioritization Specific to Faces

Even though our WM manipulation did not influence the N2pc or SPCN components, in an exploratory analysis we did find electrophysiological evidence suggesting a very early prioritization of the memorized faces. Specifically, we observed a contralateral negativity in response to the memorized face already between 100 and 200 ms after the stimulus onset. Thus, it is not clear whether contralateral negativity occurring so early can be termed N2pc, as N2pc is considered to occur around 175–200 ms after the stimulus onset (i.e., co-occurring with N2; e.g., Woodman et al., 2009; Reutter et al., 2017; Wójcik et al., 2019). Nevertheless, observing contralateral negativity already around 100 ms after the stimulus onset (i.e., co-occurring with the P100 component) suggests it represents the early and perceptual stages of processing. Enhanced activity of the occipital area in response to a stimulus held in WM has been already reported (Tan et al., 2014, 2015), but while Tan and colleagues analyzed the P100 amplitude, here we analyzed a difference between contralateral and ipsilateral activity (i.e., in our study the WM-maintained stimulus was presented always in pair with control stimulus, thus analysis of P100 is not possible). However, other studies, which used simple stimuli, did not observe any evidence for such an early WM-associated activity (e.g., Kumar et al., 2009; Telling et al., 2010). It is thus important to emphasize that in our study the early effect was present for faces, but not for houses. While the mechanisms of such an early electrophysiological effect remain to be investigated, the fact that it was observed only for faces is in line with several lines of evidence. First, due to their evolutionary and social importance, faces are processed in a largely automatic and holistic manner (Farah et al., 1998; Tsao and Livingstone, 2008). Second, due to holistic encoding strategies, faces benefit from a WM advantage (which is observed also in our data; Curby et al., 2009). Third, continuous flash suppression (CSF) studies show that faces actively maintained in WM break the CFS faster than faces that were merely seen (Pan et al., 2014). The fact that in the CFS paradigm WM can bias face perception outside of awareness is in line with the automatic and involuntary (possibly pre-attentive) nature of the effect found here. Thus, the attention prioritization revealed in the present study is a plausible mechanism accounting for the CFS effects. Finally, face recognition is performed by a specialized set of brain regions (Haxby et al., 2000; Kanwisher and Yovel, 2006) with the initial stages of face categorization occurring as early as 80–150 ms post-stimulus (Herrmann et al., 2005), which is in line with the observed early effect.

The presence of such early visual cortex activity in response to the WM-maintained stimuli is relevant to the ongoing debate on neuronal mechanisms of visual working memory. Here, two opposing theories have been proposed: first, the top-down amplification hypothesis, which assumes that visual WM items are maintained by fronto-parietal interactions (Bettencourt and Xu, 2016; Riley and Constantinidis, 2016; Christophel et al., 2018; Thigpen et al., 2019; review: Xu, 2017); second, the sensory-recruitment hypothesis, assuming that visual WM items are stored and maintained in the visual cortex (i.e., that perception and visual WM share the same neural substrate; Postle, 2006). The latter view might particularly effectively account for the automatic interactions between perception and the WM contents, which were observed in our and other studies (e.g., Silvanto and Cattaneo, 2010; Albers et al., 2013; Gayet et al., 2017; Teng and Kravitz, 2019). Importantly, given that ERP components observed in the 100–200 ms time-range are generated by sensory brain regions and reflect perceptual processing (Nusslock, 2016), such an early prioritization of the memorized faces provides support for the sensory recruitment theory. Further, such an early effect was not found in previous dot-probe studies using very salient and relevant emotional faces (Holmes et al., 2009) or self-faces of participants (Wójcik et al., 2019; Bola et al., 2021), which further indicates it might specifically reflect a match between the WM-maintained representation and an incoming stimulus. However, because this analysis was exploratory, the conclusion should be treated with caution.

Limitations and Conclusion

There are two main limitations of our study that should be pointed out. First, we are not able to definitely exclude the possibility of strategic allocation of attention to the WM-maintained stimuli. In such a scenario participants would focus on the template image in order to improve (refresh) its representation in WM and perform better in the subsequent memory test. However, the design of our study–including brief presentation time of distractor stimuli (200 ms), the minimal demands on WM for the memory test (i.e., participants performance exhibited a ceiling-level), and the fact that occurrence of the observed WM-based attention capture effect was detrimental to the dot-probe task performance–discouraged such strategic and volitional effects. Further, considering that robust N2pc is observed when task-relevant stimuli are attended strategically (Kumar et al., 2009), and that SPCN's amplitude increases during retention and processing of stimuli in the WM (Jolicæur et al., 2008; Emrich et al., 2009; Clark et al., 2015), the fact that neither component was observed in our data is also an argument against the strategic resampling. Second limitation is that WM and exposure conditions differed in terms of cognitive effort necessary to encode and maintain the presented stimulus. As this might be important for disentangling the exact mechanisms behind attentional effects, future studies should use the retro-cueing procedure (e.g., Gayet et al., 2017; Gayet and Peelen, 2019).

In conclusion, our study provides evidence that attentional prioritization of WM-maintained stimuli can be observed for complex and naturalistic stimuli, and thus encourages further investigations of this effect in more ecological conditions. The pattern of RT results and the dissociation between RT and

N2pc suggests that the observed attentional prioritization might reflect hold, rather than capture of attention. Further, our electrophysiological results provide evidence for a very early prioritization of the memorized face images, which is in line with the sensory-recruitment theories of WM.

DATA AVAILABILITY STATEMENT

The data used in the statistical analysis can be accessed from the OSF (https://osf.io/9rc4j/). Raw EEG data and scripts used for the presentation of the experimental procedure and data analysis will be shared by the authors per request.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Human Ethics Committee of the SWPS University of Social Sciences and Humanities (Warsaw, Poland). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

MB: conceived the study. NR: collected the data and drafted the manuscript. ŁD and MB: revised the manuscript. All authors designed the study, analyzed the data, and approved the submitted version.

FUNDING

This study was supported by a National Science Center Poland grant (2018/29/B/HS6/02152).

REFERENCES

Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., and De Lange, F. P. (2013). Shared representations for working memory and mental imagery in early visual cortex. *Curr. Biol.* 23, 1427–1431. doi: 10.1016/j.cub.2013.05.065

Bettencourt, K. C., and Xu, Y. (2016). Decoding the content of visual short-term memory under distraction in occipital and parietal areas. *Nat. Neurosci.* 19:150. doi: 10.1038/nn.4174

Bola, M., Paź, M., Doradzińska, Ł, and Nowicka, A. (2021). The self-face captures attention without consciousness: evidence from the N2pc ERP component analysis. *Psychophysiology* 58:e13759. doi: 10.1111/psyp.13759

Carlisle, N. B., and Woodman, G. F. (2011). When memory is not enough: electrophysiological evidence for goal-dependent use of working memory representations in guiding visual attention. *J. Cogn. Neurosci* 23, 2650–2664. doi: 10.1162/jocn.2011.21602

Carlisle, N. B., and Woodman, G. F. (2013). Reconciling conflicting electrophysiological findings on the guidance of attention by working memory. *Atten. Percept. Psychophys.* 75, 1330–1335. doi: 10.3758/s13414-013-0529-7

Christophel, T. B., Allefeld, C., Endisch, C., and Haynes, J. D. (2018). View-independent working memory representations of artificial shapes in prefrontal and posterior regions of the human brain. *Cereb. Cortex* 28, 2146–2161. doi: 10.1093/cercor/bhx119

Clark, K., Appelbaum, L. G., van den Berg, B., Mitroff, S. R., and Woldorff, M. G. (2015). Improvement in visual search with practice: mapping learning-related changes in neurocognitive stages of processing. *J. Neurosci.* 35, 5351–5359. doi: 10.1523/JNEUROSCI.1152-14.2015

Curby, K. M., Glazek, K., and Gauthier, I. (2009). A visual short-term memory advantage for objects of expertise. *J. Exp. Psychol. Hum. Percept. Perform.* 35:94. doi: 10.1037/0096-1523.35.1.94

Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. *Ann. Rev Neurosci.* 18, 193–222. doi: 10.1146/annurev.ne.18.030195.001205

Downing, P., and Dodds, C. (2004). Competition in visual working memory for control of search. *Visual Cogn.* 11, 689–703. doi: 10.1080/13506280344000446

Downing, P. E. (2000). Interactions between visual working memory and selective attention. *Psychol. Sci.* 11, 467–473. doi: 10.1111/1467-9280.00290

Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalogr. Clin Neurophysiol.* 99, 225–234. doi: 10.1016/0013-4694(96)95711-9

Emrich, S. M., Al-Aidroos, N., Pratt, J., and Ferber, S. (2009). Visual search elicits the electrophysiological marker of visual working memory. *PLoS One* 4:e8042. doi: 10.1371/journal.pone.0008042

Farah, M. J., Wilson, K. D., Drain, M., and Tanaka, J. N. (1998). What is" special" about face perception? *Psychol. Rev.* 105:482. doi: 10.1037/0033-295x.105.3.482

Filliter, J. H., Glover, J. M., McMullen, P. A., Salmon, J. P., and Johnson, S. A. (2016). The DalHouses: 100 new photographs of houses with ratings of typicality, familiarity, and degree of similarity to faces. *Behav. Res. Methods* 48, 178–183. doi: 10.3758/s13428-015-0561-8

Furtak, M., Doradzińska, Ł., Ptashynska, A., Mudrik, L., Nowicka, A., and Bola, M. (2020). Automatic attention capture by threatening, but not by semantically incongruent natural scene images. *Cereb. Cort.* 30, 4158–4168. doi: 10.1093/cercor/bhaa040

Gayet, S., Guggenmos, M., Christophel, T. B., Haynes, J. D., Paffen, C. L., Van der Stigchel, S., et al. (2017). Visual working memory enhances the neural response to matching visual input. *J. Neurosci.* 37, 6638–6647. doi: 10.1523/JNEUROSCI.3418-16.2017

Gayet, S., and Peelen, M. V. (2019). Scenes modulate object processing before interacting with memory templates. *Psychol. Sci.* 30, 1497–1509. doi: 10.1177/0956797619869905

Haxby, J. V., Hoffman, E. A., and Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends Cogn. Sci.* 4, 223–233. doi: 10.1016/s1364-6613(00)01482-0

Herrmann, M. J., Ehlis, A. C., Ellgring, H., and Fallgatter, A. J. (2005). Early stages (P100) of face perception in humans as measured with event-related potentials (ERPs). *J. Neural Transm.* 112, 1073–1081. doi: 10.1007/s00702-004-0250-8

Hollingworth, A., Matsukura, M., and Luck, S. J. (2013). Visual working memory modulates rapid eye movements to simple onset targets. *Psychol. Sci.* 24, 790–796. doi: 10.1177/0956797612459767

Holmes, A., Bradley, B. P., Kragh Nielsen, M., and Mogg, K. (2009). Attentional selectivity for emotional faces: Evidence from human electrophysiology. *Psychophysiology* 46, 62–68. doi: 10.1111/j.1469-8986.2008.00750.x

Houtkamp, R., and Roelfsema, P. R. (2006). The effect of items in working memory on the deployment of attention and the eyes during visual search. *J. Exp. Psychol.: Hum. Percept. Perform.* 32:423. doi: 10.1037/0096-1523.32.2.423

Jolicæur, P., Brisson, B., and Robitaille, N. (2008). Dissociation of the N2pc and sustained posterior contralateral negativity in a choice response task. *Brain Res.* 1215, 160–172. doi: 10.1016/j.brainres.2008.03.059

Kanwisher, N., and Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philos. Trans. R. Soc. B, Biol. Sci.* 361, 2109–2128. doi: 10.1098/rstb.2006.1934

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

Kiss, M., Van Velzen, J., and Eimer, M. (2008). The N2pc component and its links to attention shifts and spatially selective visual processing. *Psychophysiology* 45, 240–249. doi: 10.1111/j.1469-8986.2007.00611.x

Kumar, S., Soto, D., and Humphreys, G. W. (2009). Electrophysiological evidence for attentional guidance by the contents of working memory. *Eur. J. Neurosci.* 30, 307–317. doi: 10.1111/j.1460-9568.2009.06805.x

Luck, S. J. (2012). "Electrophysiological correlates of the focusing of attention within complex visual scenes: N2pc and related ERP components," in *Oxford Handbook of Event-Related Potential Components*, eds S. J. Luck and E. S. Kappenman (New York, NY: Oxford University Press), 329–360.

Luck, S. J., and Hillyard, S. A. (1994). Spatial filtering during visual search: evidence from human electrophysiology. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 1000–1014. doi: 10.1037/0096-1523.20.5.1000

Lundqvist, D., Flykt, A., and Öhman, A. (1998). *The Karolinska Directed Emotional Faces - KDEF, CD ROM from Department of Clinical Neuroscience, Psychology section*. Solna: Karolinska Institutet.

MacLeod, C., Mathews, A., and Tata, P. (1986). Attentional bias in emotional disorders. *J. Abnorm. Psychol.* 95:15. doi: 10.1037/0021-843x.95.1.15

Nobre, A. C., and Stokes, M. G. (2019). Premembering experience: a hierarchy of time-scales for proactive attention. *Neuron* 104, 132–146. doi: 10.1016/j.neuron.2019.08.030

Nusslock, R. (2016). "Neurophysiology and neuroimaging," in *APA Handbook of Clinical Psychology Theory and Research*, Vol. 2, eds J. C. Norcross, G. R. VandenBos, and D. K. Freedheim (Washington DC: American Psychological Association).

Olivers, C. N., Meijer, F., and Theeuwes, J. (2006). Feature-based memory-driven attentional capture: visual working memory content affects visual attention. *J. Exp. Psychol. Hum. Percept. Perform.* 32:1243. doi: 10.1037/0096-1523.32.5.1243

Pan, Y., Lin, B., Zhao, Y., and Soto, D. (2014). Working memory biasing of visual perception without awareness. *Atten. Percept. Psychophys.* 76, 2051–2062. doi: 10.3758/s13414-013-0566-2

Peters, J. C., Goebel, R., and Roelfsema, P. R. (2009). Remembered but unused: The accessory items in working memory that do not guide attention. *J. Cogn. Neurosci.* 21, 1081–1091. doi: 10.1162/jocn.2009.21083

Posner, M. I., Inhoff, A. W., Friedrich, F. J., and Cohen, A. (1987). Isolating attentional systems: a cognitive-anatomical analysis. *Psychobiology* 15, 107–121. doi: 10.3758/bf03333099

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience* 139, 23–38. doi: 10.1016/j.neuroscience.2005.06.005

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

Riley, M. R., and Constantinidis, C. (2016). Role of prefrontal persistent activity in working memory. *Front. Syst. Neurosci.* 9:181. doi: 10.3389/fnsys.2015.00181

Schneegans, S., Spencer, J. P., Schöner, G., Hwang, S., and Hollingworth, A. (2014). Dynamic interactions between visual working memory and saccade target selection. *J. Vis.* 14, 9–9. doi: 10.1167/14.11.9

Silvanto, J., and Cattaneo, Z. (2010). Transcranial magnetic stimulation reveals the content of visual short-term memory in the visual cortex. *Neuroimage* 50, 1683–1689. doi: 10.1016/j.neuroimage.2010.01.021

Silvis, J. D., and Van der Stigchel, S. (2014). How memory mechanisms are a key component in the guidance of our eye movements: evidence from the global effect. *Psychon. Bull. Rev.* 21, 357–362. doi: 10.3758/s13423-013-0498-9

Soto, D., Heinke, D., Humphreys, G. W., and Blanco, M. J. (2005). Early, involuntary top-down guidance of attention from working memory. *J. Exp. Psychol. Hum. Percept. Perform.* 31:248. doi: 10.1037/0096-1523.31.2.248

Soto, D., Hodsoll, J., Rotshtein, P., and Humphreys, G. W. (2008). Automatic guidance of attention from working memory. *Trends Cogn. Sci.* 12, 342–348. doi: 10.1016/j.tics.2008.05.007

Soto, D., and Humphreys, G. W. (2007). Automatic guidance of visual attention from verbal working memory. *J. Exp. Psychol. Hum Percept. Perform.* 33:730. doi: 10.1037/0096-1523.33.3.730

Soto, D., and Humphreys, G. W. (2008). Stressing the mind: The effect of cognitive load and articulatory suppression on attentional guidance from working memory. *Percept. Psychophys.* 70, 924–934. doi: 10.3758/pp.70.5.924

Soto, D., and Humphreys, G. W. (2009). Automatic selection of irrelevant object features through working memory: evidence for top-down attentional capture. *Exp. Psychol.* 56, 165–172. doi: 10.1027/1618-3169.56.3.165

Soto, D., Humphreys, G. W., and Heinke, D. (2006). Dividing the mind: the necessary role of the frontal lobes in separating memory from search. *Neuropsychologia* 44, 1282–1289. doi: 10.1016/j.neuropsychologia.2006.01.029

Soto, D., Humphreys, G. W., and Rotshtein, P. (2007). Dissociating the neural mechanisms of memory-based guidance of visual selection. *Proc. Natl. Acad. Sci.* 104, 17186–17191. doi: 10.1073/pnas.0703706104

Tan, J., Zhao, Y., Wang, L., Tian, X., Cui, Y., Yang, Q., et al. (2015). The competitive influences of perceptual load and working memory guidance on selective attention. *PLoS One* 10:e0129533. doi: 10.1371/journal.pone.0129533

Tan, J., Zhao, Y., Wu, S., Wang, L., Hitchman, G., Tian, X., et al. (2014). The temporal dynamics of visual working memory guidance of selective attention. *Front. Behav. Neurosci.* 8:345. doi: 10.3389/fnbeh.2014.00345

Telling, A. L., Kumar, S., Meyer, A. S., and Humphreys, G. W. (2010). Electrophysiological evidence of semantic interference in visual search. *J. Cogn. Neurosci.* 22, 2212–2225. doi: 10.1162/jocn.2009.21348

Teng, C., and Kravitz, D. J. (2019). Visual working memory directly alters perception. *Nat. Hum. Behav.* 3, 827–836. doi: 10.1038/s41562-019-0640-4

Thigpen, N., Petro, N. M., Oschwald, J., Oberauer, K., and Keil, A. (2019). Selection of visual objects in perception and working memory one at a time. *Psychol. Sci.* 30, 1259–1272. doi: 10.1177/0956797619854067

Tsao, D. Y., and Livingstone, M. S. (2008). Mechanisms of face perception. *Annu. Rev. Neurosci.* 31, 411–443. doi: 10.1146/annurev.neuro.30.051606.094238

Wójcik, M. J., Nowicka, M. M., Bola, M., and Nowicka, A. (2019). Unconscious detection of one's own image. *Psychol. Sci.* 30, 471–480. doi: 10.1177/0956797618822971

Wolfe, J. M., and Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nat Hum Behav.* 1, 1–8.

Woodman, G. F., Arita, J. T., and Luck, S. J. (2009). A cuing study of the N2pc component: An index of attentional deployment to objects rather than spatial locations. *Brain Res.* 1297, 101–111. doi: 10.1016/j.brainres.2009.08.011

Woodman, G. F., and Luck, S. J. (2007). Do the contents of visual working memory automatically influence attentional selection during visual search? *J. Exp. Psychol. Hum. Percept. Perform.* 33, 363. doi: 10.1037/0096-1523.33.2.363

Xu, Y. (2017). Reevaluating the sensory account of visual working memory storage. *Trends Cogn. Sci.* 21, 794–815. doi: 10.1016/j.tics.2017.06.013

Zhang, B., Zhang, J. X., Kong, L., Huang, S., Yue, Z., and Wang, S. (2010). Guidance of visual attention from working memory contents depends on stimulus attributes. *Neurosci. Lett.* 486, 202–206. doi: 10.1016/j.neulet.2010.09.052

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.

frontiers | Frontiers in Human Neuroscience

TYPE Original Research
PUBLISHED 13 March 2023
DOI 10.3389/fnhum.2023.1145282

OPEN ACCESS

EDITED BY
Hidehiko Okamoto,
International University of Health and Welfare (IUHW), Japan

REVIEWED BY
Mitsuru Kikuchi,
Kanazawa University, Japan
Toshio Higashi,
Nagasaki University, Japan

*CORRESPONDENCE
Ryouhei Ishii
ishii@psy.med.osaka-u.ac.jp

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

RECEIVED 16 January 2023
ACCEPTED 27 February 2023
PUBLISHED 13 March 2023

CITATION
Ueno K, Ishii R, Ueda M, Yuri T, Shiroma C, Hata M and Naito Y (2023) Frontal midline theta rhythm and gamma activity measured by sheet-type wearable EEG device. *Front. Hum. Neurosci.* 17:1145282. doi: 10.3389/fnhum.2023.1145282

Frontal midline theta rhythm and gamma activity measured by sheet-type wearable EEG device

Keita Ueno[1], Ryouhei Ishii[1,2]*, Masaya Ueda[1], Takuma Yuri[3,4], China Shiroma[1], Masahiro Hata[2] and Yasuo Naito[1]

[1]Department of Occupational Therapy, Graduate School of Rehabilitation Science, Osaka Metropolitan University, Osaka, Japan, [2]Department of Psychiatry, Osaka University Graduate School of Medicine, Osaka, Japan, [3]Rehabilitation Unit, Kyoto University Hospital, Kyoto, Japan, [4]Department of Health Sciences, Graduate School of Health Sciences, Yamagata Prefectural University of Health Sciences, Yamagata, Japan

Introduction: The current study measured the frontal midline theta rhythm (Fmθ), which appears in the frontal midline region during the attentional focus state, using the sheet-type wearable electroencephalograph (EEG) device HARU-1, and examined the modulation of frontal gamma band activity by cognitive tasks.

Methods: We measured the frontal EEG of 20 healthy subjects using HARU-1 for 2 min during the rest eyes-closed condition and simple mental calculation task condition, respectively. Statistical analyses were conducted using permutation testing based on *t*-test and cluster analysis to compare the results between the resting state and the task condition.

Results: Twelve of 20 subjects showed Fmθ during the task condition. The 12 subjects with Fmθ showed significantly higher activity of the theta and gamma bands, and significantly low activity of the alpha band during the task condition compared to the resting condition. In the eight subjects without Fmθ were significantly low activity of the alpha and beta bands and no significant activity in the theta and gamma band activity during the task condition compared to the resting condition.

Discussion: These results indicate that it is possible to measure Fmθ using HARU-1. A novel finding was the gamma band activity appearing with Fmθ in the left and right frontal forehead regions, suggesting that it reflects the function of the prefrontal cortex in working memory tasks.

KEYWORDS

EEG, frontal midline theta rhythm (Fmθ), gamma activities, focused attention, calculation, HARU-1, sheet-type EEG device, wearable device

1. Introduction

Electroencephalography (EEG) is a leading method for non-invasive measurement of human brain activity and has been used in various fields, from the diagnosis of epilepsy and sleep disorders to brain-computer interface (Teplan, 2002; Padfield et al., 2019). However, EEG amplitude is smaller than electromyography (Raez et al., 2006), electrocardiography (Velayudhan and Peter, 2016), and electro-oculography (Choudhury et al., 2005). Therefore, when measuring EEG, it is necessary to avoid these artifacts. Additionally, power line interference can cause artifacts, limiting the environment. Furthermore, a conventional EEG

device has been found to be cumbersome, time-consuming, painful, and uncomfortable to use, requiring skin preparation, gel-electrode application, mounting many wired sensors, and connecting electrodes to the main acquisition unit and personal computer (Mihajlovic et al., 2015).

In the last decade, various types of wearable EEG devices have been developed, and some of them are commercially available. However, most non-clinical EEG solutions are designed for general-use EEG applications with a lack of support for sophisticated signal processing and effective feedback generation, and most of them are used for entertainment such as games (Mihajlovic et al., 2015). In this context, PGV Inc. has released a sheet-type wearable EEG device "HARU-1" (PGV Inc., n.d.a). HARU-1 applies sheet-type EEG electrodes developed by Professor Tsuyoshi Sekitani (Araki et al., 2019). The electrode sheet used in HARU-1 is manufactured by printing electrodes on a flexible and stretchable sheet, which adapts to the movement of the skin surface and does not cause stress or discomfort to the subject (Matsumori et al., 2022). They are low-cost and disposable. The main unit (amplifier, AD converter, wireless signal transmission unit, battery, etc.) is small and light enough to fit in the front forehead. Furthermore, the system is completely wireless, transmitting EEG signals to smartphones and tablets *via* Bluetooth, making it possible to measure EEG during activities without restricting the wearer's activities (PGV Inc., n.d.a). Li et al. (2019) reported that repeated and longitudinal recording of EEG using sheet-type EEG will contribute to the detection of specific EEG patterns for patients with sleep/wake-related problems to dementia. Matsumori et al. (2022) presented a novel automatic sleep scoring system with HARU-1, and the results indicated that the system is feasible to score the sleep stage with an accuracy comparable to that achieved by clinical polysomnography devices.

One of the characteristic EEG components that appears during mental activity is the frontal midline theta rhythm (Fmθ). Ishihara and Yoshii (1972) discovered Fmθ, which appears in the frontal midline of healthy subjects by using a 13-channel EEG system. Fmθ is induced in an attentional focus, such as during psychological tasks (Ishihara and Yoshii, 1972; Asada et al., 1999; Ishii et al., 2014), and is particularly sensitive to immersion in a task, or when attention is focused on a single task. Fmθ appears in tasks such as the Uchida-Kraepelin test (Ishihara and Yoshii, 1972), rifle shooting preparation (Doppelmayr et al., 2008), and car driving (Laukka et al., 1995), supporting the idea that Fmθ is related to the attentional focus state. Thus, measuring Fmθ during work activities can objectively evaluate the state of attentional focus. In the field of occupational therapy, the appearance of Fmθ during craft activities has been shown to be related to the autonomic nervous responses, suggesting that focusing attention on craft activities can have relaxing effects (Shiraiwa et al., 2020). Also, during cognitive tasks, active mobilization of cortical areas has been reported to be translated into a local increase in the gamma bandwidth (Tallon-Baudry and Bertrand, 1999). Furthermore, the prefrontal cortex (PFC) may influence the posterior sensory areas where working memory representations are maintained (Lara and Wallis, 2015). Using spectral analysis and beamforming, Roux et al. (2012) suggests that gamma oscillations in PFC are critically implicated in the maintenance of relevant working memory information. Previous studies that measured intracerebral EEG during working memory tasks have also reported modulation of gamma-band activity (Mainy et al., 2007; Meltzer et al., 2008). Based on these findings, gamma-band activity may be modulated by cognitive tasks that induce Fmθ.

Since HARU-1 can be used to measure EEG during situations involving movement, it is a suitable device for measuring EEG during work activities and monitoring the subject's attentional state. However, no attempt has been made to use HARU-1 to measure Fmθ induced by the attentional state, nor gamma-band activity modulated during working memory tasks. This study aimed to validate the utility of HARU-1 for measuring Fmθ and gamma band activity, during a calculation task confirmed to be validated in previous studies.

2. Materials and methods

2.1. Subjects

The subjects were 20 healthy young volunteers (mean age: 24.6 ± 8.7 years, 8 males and 12 females).

2.2. Experimental procedure

Electroencephalography measurements were performed in the sitting position. The room in which the EEG was measured was not a shielded room. However, EEG was measured after confirming that no HAM noise was introduced by real-time spectral analysis using the measurement application HARU-Measure (PGV Inc., Tokyo, Japan). The stretchable electrode sheets (weight = 0.5 g, thickness = 80 μm, stretchability ≤ 150%) used in this study were fabricated by screen-printing an Ag particle-containing paste onto the elastomer substrate that had a moisture permeability of up to 2,700 g m^{-2} day^{-1} (25 μm thickness at 40°C and 90% humidity), which can help avoid bacterial growth even in a wet state (Araki et al., 2019). The reliability of this EEG sensor platform has already been evaluated by comparison with the International 10–20 system, showing a voltage resolution not inferior to that from a high-end device for fixed EEG (Araki et al., 2019). The electrode sheet was attached with a special conductive gel, and the skin of the front forehead and left mastoid region of the subject was wiped with an alcohol swab. The electrode sheet was attached in the same position as AFz, Fp1, and Fp2 based on the International 10–10 system, and the reference electrode was positioned at the left mastoid. The sampling rate was set to 250 Hz (**Figure 1**).

Electroencephalography signals were transmitted *via* Bluetooth to the measurement tablet LAVIE Tab PC-TE507JAW (NEC Corporation, Japan) and recorded using the measurement application HARU-Measure (PGV Inc., Tokyo, Japan).

Below is an overview of the tasks (**Figure 2**).

Electroencephalography signals were recorded under the following two conditions.

1. Resting eyes-closed condition.

The subjects were instructed to close their eyes while sitting on a chair. The eyes were then kept closed for 2 min.

2. Simple mental calculation task condition.

We adopted a calculation task as an Fmθ-induced task, as in previous studies (Mizuki et al., 1980; Iramina et al., 1996;

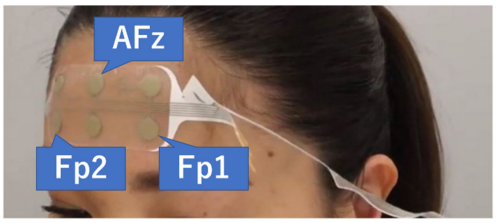


FIGURE 1
The electrode sheet is attached. The three channels used are AFz, Fp1, and Fp2. The reference electrode was positioned at the left mastoid. PGV Inc. (n.d.b).

Ishii et al., 1999). Subjects were instructed to subtract 7 from 1,000 for 2 min in a row with their eyes closed. If subjects calculated while saying the answer or using their fingers, they were instructed to "calculate in your mind."

2.3. Analysis

Electroencephalography data of 20 subjects were visually checked and evaluated whether Fmθ appeared or not by certificated medical doctor of the Japanese Society of Clinical Neurophysiology. The measured EEG data for the resting eyes-closed condition and the simple mental calculation task condition were analyzed using the Brain Electrical Source Analysis: BESA Research 6.0 (BESA GmbH, Germany) software. High Pass Filter of 4.0 Hz, Low Pass Filter of 75.0 Hz, and Notch Filter of 60.0 Hz were applied offline. Two minutes of EEG data in each of the two conditions were divided into 2 s. Artifacts such as blinks and electromyograms were visually removed. Additionally, artifact scans were performed with an amplitude threshold that accepted at least 85% of the remaining epochs. The amplitude thresholds for the two conditions were the same and were set for each subject. The mean value of the amplitude threshold was 102 ± 27 μV. The average number of remaining epochs in the resting eyes-closed condition was 54 ± 6, and the average number of remaining epochs in the simple mental calculation task condition was 50 ± 6. We performed time-frequency analysis in the frequency range of 4.0–75.0 Hz in 0.25 Hz-steps with a time sampling rate of 200 ms-steps by applying complex demodulation to transform time-domain EEG data into time-frequency data which was implemented in BESA (Hoechstetter et al., 2004). A *post hoc* power analysis was performed using G power 3.1 software with a medium effect size (0.5) and α of 0.05 (two-tailed paired *t*-test).

The two groups were set up by certificated medical doctor of the Japanese Society of Clinical Neurophysiology as follows: the Fmθ group (Subjects with theta waves for more than 1 s during simple mental calculation task) and the No-Fmθ group (except for the Fmθ group). Statistical analyses were conducted using BESA Statistics 1.0 (BESA GmbH, Germany) for permutation testing based on *t*-test and cluster analysis (Bornfleth et al., 2020). For the BESA Statistics analysis procedure, we first performed a preliminary test, the paired *t*-test, to define data clusters showing significant effects. After summing all the t-values for each cluster to derive the cluster value, we performed a permutation test based on the cluster value. This method addressed the issue of multiple comparisons. We set "Average over Time" to average the 2 s windows and investigated the statistical difference between the resting eyes-closed condition and the simple mental calculation task condition in each of the two groups. The significance level was set at 0.05.

3. Results

3.1. Subject classification based on the appearance of Fmθ

Twelve of the 20 subjects (mean age: 21.2 ± 1.2 years, 4 males, 8 females) showed Fmθ in the simple mental calculation task condition and were assigned to the Fmθ group (**Figure 3**). The remaining eight subjects (mean age: 29.6 ± 12.0 years, 4 males, 4 females), who did not show clear theta waves in the simple calculation task condition, were assigned to the No-Fmθ group.

3.2. Fmθ group

The Fmθ group showed significantly higher activity in theta band in all channels, beta band in all channels and gamma band in Fp1 and Fp2 and significantly lower activity in alpha band in AFz during the task condition compared to the resting condition (**Table 1**).

3.3. No-Fmθ group

The No-Fmθ group showed significantly low activity in theta band in Fp2 and AFz, alpha band in Fp2 and AFz, beta band in all channels and low-gamma band in AFz during the task condition compared to the resting condition (**Table 2**).

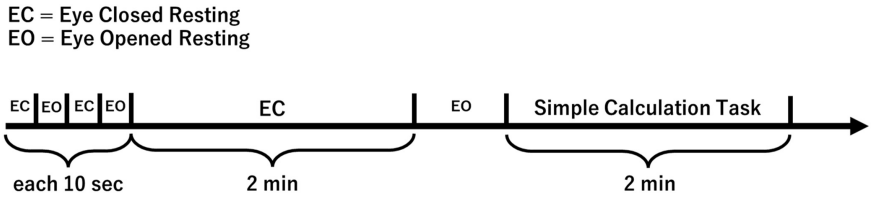


FIGURE 2
EEG measurement overview.

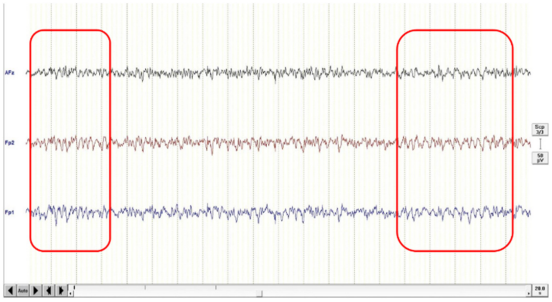


FIGURE 3
An example of Fmθ lasting more than 1 s during a simple calculation task.

4. Discussion

In this study, we tested the feasibility of measuring the Fmθ using the sheet-type wearable EEG device "HARU-1." We found that 12 of 20 subjects (60.0%) showed Fmθ for more than 1 s in the simple calculation task condition. Additionally, the activity of the gamma band was different between the Fmθ group and the No-Fmθ group.

4.1. The feasibility of measuring Fmθ using "HARU-1"

In a previous study by Ishihara, 14 university students were asked to perform the Uchida-Kraepelin test, intelligence test, mental calculation, and Tetris, and EEG measurements during the tasks showed that Fmθ was recorded in 5 (35.7%), and long-lasting Fmθ was recorded from 2 of them (Ishihara, 2020). Ishii et al. (2014) performed magnetoencephalography (MEG) recordings using a helmeted whole-head array of 64-channel SQUID sensors and reported that prominent Fmθ was observed in 8 of 11 subjects (72.7%) during the calculation task. Ishihara reported that the Fmθ elicitation rate varies greatly depending on the subject and the eliciting medium and is likely to vary depending on differences in attitude toward the task, the environment in which the subject is placed, and the subject's personality (Ishihara, 2020). Mizuki et al. (1984) investigated the relationship between the occurrence of Fmθ and personality and reported that subjects who were more anxious, more introverted, and had higher neurotic tendencies showed less or no Fmθ. Tani also reported that increasing motivation for work activities increased the occurrence of Fmθ (Tani, 1978). The subjects who did not exhibit Fmθ in this study were considered to have been unable to focus attention although they were able to perform the task due to the experimental environment, the type and difficulty of the task, and their personalities. In this study, Fmθ during the calculation task was observed in 12 of 20 subjects (60.0%). Our finding is similar to previous studies (Ishii et al., 2014; Ishihara, 2020). This suggests that HARU-1 can measure Fmθ and serve as an objective measure of attentional focus during activities.

4.2. EEG characteristics during calculation tasks and attentional focus

In this study, the Fmθ group showed low activity during the calculation task in the α band of AFz, and the No-Fmθ group showed low activity in the α and β bands in all channels. There is a difference in the spectral parameters of the EEG between resting and cognitive tasks, as well as during different mental tasks (Fernández et al., 1995). In particular, the decrease in alpha band activity during mental tasks has been a common observation since the EEG was firstly recorded (Compston, 2010). These previous studies support a significantly low activity of alpha band in all subjects in this study.

We found the gamma band activity only in the Fmθ group during the calculation task. The calculation task used in this study was simple continuous subtraction of 7 from 1,000 which involved several cognitive functions, such as focused attention, executive function, figure manipulation and working memory. Previous studies have reported that the PFC plays a role in maintenance of information, representation, attention, and inhibition in working memory tasks (Miller and Cohen, 2001; Curtis and D'Esposito, 2003). Kaiser and Lutzenberger (2005) reported that synchronized gamma band activity is involved in bottom-up driven perception and top-down driven functions such as selective attention and retention of information in memory. Using spectral analysis and beamforming, Roux et al. (2012) suggests that gamma oscillations in PFC are critically implicated in the maintenance of relevant working memory information. The frontal gamma band activity identified in this study may represent representational maintenance in working memory tasks.

Ishii et al. (2014) and Fitzgibbon et al. (2004) also reported gamma band activity that appears with Fmθ during calculation tasks. Both reports showed increased gamma band activity in the parietal or occipital regions, not sustained frontal gamma band activity associated with Fmθ. The coupling of theta and gamma oscillations has been reported to be associated with the formation of a neural code, recall of sequences of items from long-term memory, and encoding of short-term memory (Lisman and Buzsáki, 2008). Kaplan et al. (2014) suggest that theta phase coupling between medial PFC (mPFC) and medial temporal lobe and theta-gamma phase-amplitude coupling between mPFC and neocortical regions may play a role in human spatial memory retrieval. Therefore, the high gamma band power during the computational task is related to Fmθ and may be theta-gamma coupling that reflects the process of memory processing and recall. As future application, the gamma band activity in the left and right frontal regions observed at the time of Fmθ appearance needs to be verified by multichannel EEG and MEG.

The wearable EEG device is suitable as a tool for assessing and monitoring brain activity in actual daily life because it has few restrictions on activity. The results of this study indicate the possibility that HARU-1 can monitor the attentional state in activities of daily living (ADL). It can aid in the assessment and intervention of difficult ADL in patients with stroke, dementia, and subjects with attention-deficit hyperactivity disorder (ADHD).

TABLE 1 Results of permutation test in Fmθ group.

Channel	Cluster value	Mean for task	Mean for rest	Start frequency (Hz)	End frequency (Hz)	*p*-value
Fp1	32.12	4.18	3.15	4.0	7.0	$p < 0.01$
	4.97	2.90	2.52	13.0	13.3	$p < 0.05$
	5.20	1.30	1.20	23.8	24.1	$p < 0.05$
	7.00	1.24	1.14	24.7	25.3	$p < 0.05$
	4.99	1.08	1.01	28.0	28.3	$p < 0.05$
	8.61	0.99	0.89	35.5	36.1	$p < 0.02$
	65.44	0.90	0.79	38.8	45.1	$p < 0.01$
	16.43	0.84	0.74	45.7	47.2	$p < 0.01$
	7.61	0.81	0.73	48.4	49.0	$p < 0.02$
	249.20	0.65	0.57	50.2	73.3	$p < 0.01$
	13.90	0.57	0.49	73.9	75.1	$p < 0.01$
Fp2	16.31	4.13	3.22	4.6	6.1	$p < 0.01$
	5.08	2.98	2.63	13.0	13.3	$p < 0.05$
	12.55	1.37	1.22	24.4	25.3	$p < 0.01$
	17.91	1.20	1.06	27.7	29.5	$p < 0.01$
	28.46	1.15	0.99	30.1	33.1	$p < 0.01$
	19.57	1.10	0.95	34.3	36.1	$p < 0.01$
	400.89	0.79	0.66	39.1	75.1	$p < 0.01$
AFz	4.73	4.37	3.54	4.0	4.3	$p < 0.05$
	8.16	3.14	2.82	5.2	5.8	$p < 0.02$
	−15.53	3.59	4.39	10.3	11.5	$p < 0.03$
	8.01	1.25	1.16	24.4	25.0	$p < 0.02$

Frequency bands with positive cluster values indicate high activity in the task condition, and frequency bands with negative values indicate low activity in the task condition.

TABLE 2 Results of permutation test in No-Fmθ group.

Channel	Cluster value	Mean for task	Mean for rest	Start frequency (Hz)	End frequency (Hz)	*p*-value
Fp1	−33.99	1.28	1.50	19.3	22.3	$p < 0.01$
	−17.45	1.17	1.38	22.9	24.4	$p < 0.01$
	−21.25	1.07	1.24	27.1	29.2	$p < 0.01$
Fp2	−19.17	2.09	2.56	7.6	9.4	$p < 0.01$
	−24.16	1.23	1.43	16.6	19.0	$p < 0.01$
	−28.30	1.34	1.58	19.6	22.0	$p < 0.01$
	−16.83	1.15	1.42	23.5	24.7	$p < 0.01$
AFz	−15.32	2.04	2.77	8.5	9.7	$p < 0.01$
	−23.93	1.19	1.38	16.6	18.7	$p < 0.01$
	−24.66	1.26	1.48	19.3	21.4	$p < 0.01$
	−22.11	1.19	1.39	22.6	24.4	$p < 0.01$
	−11.67	0.87	0.99	32.8	33.7	$p < 0.05$
	−18.21	0.77	0.88	34.3	36.1	$p < 0.01$

4.3. Limitations of this study

A limitation of this study was the small number of subjects. The power of the Fmθ group analysis (n = 12) was 0.35 and the power of the No-Fmθ group analysis (n = 8) was 0.23. Statistical power may not be sufficient, a concept closely related to Type II error. The band of no significant differences between the resting condition and the task condition should be interpreted with caution. Additionally, since this study was conducted on young healthy subjects, it was impossible to examine the effects on the elderly and children, nor was it possible to examine the effects of diseases. And because of many artifacts in the low frequency range and setting the High Pass Filter at 4.0 Hz, frequency bands below the theta band could not be considered.

The tasks that tend to induce Fmθ are those with simple structures, single trials, and appropriate difficulty levels (Ishihara, 2020). Therefore, it is necessary to be careful in interpreting Fmθ when using it to monitor tasks that require the distribution of attention or require multiple trials.

5. Conclusion

The HARU-1 can be used to measure the Fmθ induced by the attentional focus state. As a novel finding, we found gamma band activity in the left and right frontal forehead regions that appeared with Fmθ. Future studies are desired that investigate the EEG activity in daily life with more subjects and examine the usefulness of wearable EEG devices in ADL.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by the Osaka Prefecture University Graduate School General Rehabilitation Studies Ethics Committee (2019–203). The patients/participants provided their written informed consent to participate in this study.

Author contributions

KU, RI, MU, CS, and YN contributed to the design, implementation of the research, and analysis of this study. KU wrote the manuscript with support from RI, MU, TY, and YN. MH instructed KU on how to use the sheet-type wearable EEG device "HARU-1". All authors contributed to the article and approved the submitted version.

Funding

This work was supported by JSPS KAKENHI Grant Numbers: JP22K11453 and JP20K11288.

Acknowledgments

We thank all the participants. We also appreciate Prof. Tsutomu Ishihara for his expertise related to Fmθ.

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

References

Araki, T., Uemura, T., Yoshimoto, S., Takemoto, A., Noda, Y., Izumi, S., et al. (2019). Wireless monitoring using a stretchable and transparent sensor sheet containing metal nanowires. *Adv. Mater.* 32:e1902684. doi: 10.1002/adma.201902684

Asada, H., Fukuda, Y., Tsunoda, S., Yamaguchi, M., and Tonoike, M. (1999). Frontal midline theta rhythms reflect alternative activation of prefrontal cortex and anterior cingulate cortex in humans. *Neurosci. Lett.* 274, 29–32. doi: 10.1016/s0304-3940(99)00679-5

Bornfleth, H., Cho, J. H., and Spangler, R. (2020). *BESA statistics 2.1 user manual*. Gräfelfing: BESA GmbH.

Choudhury, S. R., Soundararajan, V., Nemade, H. B., and Sahambi, J. S. (2005). Design and development of a novel EOG bio-potential amplifier. *Int. J. Bioelectromagn.* 7, 271–274.

Compston, A. (2010). The Berger rhythm: Potential changes from the occipital lobes in man. *Brain* 133(Pt 1), 3–6. doi: 10.1093/brain/awp324

Curtis, C. E., and D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends Cogn. Sci.* 7, 415–423. doi: 10.1016/s1364-6613(03)00197-9

Doppelmayr, M., Finkenzeller, T., and Sauseng, P. (2008). Frontal midline theta in the pre-shot phase of rifle shooting: Differences between experts and novices. *Neuropsychologia* 46, 1463–1467. doi: 10.1016/j.neuropsychologia.2007.12.026

Fernández, T., Harmony, T., Rodríguez, M., Bernal, J., Silva, J., Reyes, A., et al. (1995). EEG activation patterns during the performance of tasks involving different components of mental calculation. *Electroencephalogr. Clin. Neurophysiol.* 94, 175–182. doi: 10.1016/0013-4694(94)00262-j

Fitzgibbon, S. P., Pope, K. J., Mackenzie, L., Clark, C. R., and Willoughby, J. O. (2004). Cognitive tasks augment gamma EEG power. *Clin. Neurophysiol.* 115, 1802–1809. doi: 10.1016/j.clinph.2004.03.009

Hoechstetter, K., Bornfleth, H., Weckesser, D., Ille, N., Berg, P., and Scherg, M. (2004). BESA source coherence: A new method to study cortical oscillatory coupling. *Brain Topogr.* 16, 233–238. doi: 10.1023/b:brat.0000032857.55223.5d

Iramina, K., Ueno, S., and Matsuoka, S. (1996). MEG and EEG topography of frontal midline theta rhythm and source localization. *Brain Topogr.* 8, 329–331. doi: 10.1007/BF01184793

Ishihara, T. (2020). *EEG Fmθ: A review of 40 years of Fmθ research*. Osaka: Nagai Shoten Co., Ltd. (in Japanese).

Ishihara, T., and Yoshii, N. (1972). Multivariate analytic study of EEG and mental activity in Juvenile delinquents. *Electroencephalogr. Clin. Neurophysiol.* 33, 71–80. doi: 10.1016/0013-4694(72)90026-0

Ishii, R., Canuet, L., Ishihara, T., Aoki, Y., Ikeda, S., Hata, M., et al. (2014). Frontal midline theta rhythm and gamma power changes during focused attention on mental calculation: An MEG beamformer analysis. *Front. Hum. Neurosci.* 8:406. doi: 10.3389/fnhum.2014.00406

Ishii, R., Shinosaki, K., Ukai, S., Inouye, T., Ishihara, T., Yoshimine, T., et al. (1999). Medial prefrontal cortex generates frontal midline theta rhythm. *Neuroreport* 10, 675–679. doi: 10.1097/00001756-199903170-00003

Kaiser, J., and Lutzenberger, W. (2005). Human gamma-band activity: A window to cognitive processing. *Neuroreport* 16, 207–211. doi: 10.1097/00001756-200502280-00001

Kaplan, R., Bush, D., Bonnefond, M., Bandettini, P. A., Barnes, G. R., Doeller, C. F., et al. (2014). Medial prefrontal theta phase coupling during spatial memory retrieval. *Hippocampus* 24, 656–665. doi: 10.1002/hipo.22255

Lara, A. H., and Wallis, J. D. (2015). The role of prefrontal cortex in working memory: A mini review. *Front. Syst. Neurosci.* 9:173. doi: 10.3389/fnsys.2015.00173

Laukka, S. J., Järvilehto, T., Alexandrov, Y. I., and Lindqvist, J. (1995). Frontal midline theta related to learning in a simulated driving task. *Biol. Psychol.* 40, 313–320. doi: 10.1016/0301-0511(95)05122-q

Li, F., Egawa, N., Yoshimoto, S., Mizutani, H., Kobayashi, K., Tachibana, N., et al. (2019). Potential clinical applications and future prospect of wireless and mobile electroencephalography on the assessment of cognitive impairment. *Bioelectricity* 1, 105–112. doi: 10.1089/bioe.2019.0001

Lisman, J., and Buzsáki, G. (2008). A neural coding scheme formed by the combined function of gamma and theta oscillations. *Schizophr. Bull.* 34, 974–980. doi: 10.1093/schbul/sbn060

Mainy, N., Kahane, P., Minotti, L., Hoffmann, D., Bertrand, O., and Lachaux, J. P. (2007). Neural correlates of consolidation in working memory. *Hum. Brain Mapp.* 28, 183–193. doi: 10.1002/hbm.20264

Matsumori, S., Teramoto, K., Iyori, H., Soda, T., Yoshimoto, S., and Mizutani, H. (2022). HARU sleep: A deep learning-based sleep scoring system with wearable sheet-type frontal EEG sensors. *IEEE Access.* 10, 13624–13632. doi: 10.1109/ACCESS.2022.3146337

Meltzer, J. A., Zaveri, H. P., Goncharova, I. I., Distasio, M. M., Papademetris, X., Spencer, S. S., et al. (2008). Effects of working memory load on oscillatory power in human intracranial EEG. *Cereb. Cortex* 18, 1843–1855. doi: 10.1093/cercor/bhm213

Mihajlovic, V., Grundlehner, B., Vullers, R., and Penders, J. (2015). Wearable, wireless EEG solutions in daily life applications: What are we missing? *IEEE J. Biomed. Health Inform.* 19, 6–21. doi: 10.1109/JBHI.2014.2328317

Miller, E. K., and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202. doi: 10.1146/annurev.neuro.24.1.167

Mizuki, Y., Kajimura, N., Nishikori, S., Imaizumi, J., and Yamada, M. (1984). Appearance of frontal midline theta rhythm and personality traits. *Folia Psychiatr. Neurol. Jpn.* 38, 451–458. doi: 10.1111/j.1440-1819.1984.tb00794.x

Mizuki, Y., Tanaka, M., Isozaki, H., Nishijima, H., and Inanaga, K. (1980). Periodic appearance of theta rhythm in the frontal midline area during performance of a mental task. *Electroencephalogr. Clin. Neurophysiol.* 49, 345–351. doi: 10.1016/0013-4694(80)90229-1

Padfield, N., Zabalza, J., Zhao, H., Masero, V., and Ren, J. (2019). EEG-based brain-computer interfaces using motor-imagery: Techniques and challenges. *Sensors* 19:1423. doi: 10.3390/s19061423

PGV Inc. (n.d.a). *Homepage*. Available online at: https://www.pgv.co.jp/en/ (accessed May 6, 2022).

PGV Inc. (n.d.b). Electroencephalography sensor installation manual. Available online at: https://www.pgv.co.jp/technology-device/ (accessed November 16, 2021).

Raez, M. B. I., Hussain, M. S., and Mohd-Yasin, F. (2006). Techniques of EMG signal analysis: Detection, processing, classification and applications. *Biol. Proced. Online* 8, 11–35. doi: 10.1251/bpo115

Roux, F., Wibral, M., Mohr, H. M., Singer, W., and Uhlhaas, P. J. (2012). Gamma-band activity in human prefrontal cortex codes for the number of relevant items maintained in working memory. *J. Neurosci.* 32, 12411–12420. doi: 10.1523/JNEUROSCI.0421-12.2012

Shiraiwa, K., Yamada, S., Nishida, Y., and Toichi, M. (2020). Changes in electroencephalography and cardiac autonomic function during craft activities: Experimental evidence for the effectiveness of occupational therapy. *Front. Hum. Neurosci.* 14:621826. doi: 10.3389/fnhum.2020.621826

Tallon-Baudry, C., and Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends Cogn. Sci.* 3, 151–162. doi: 10.1016/s1364-6613(99)01299-1

Tani, K. (1978). Motivation to task and Fmθ. *Rinshou Nouha* 20, 115–120. (in Japanese).

Teplan, M. (2002). Fundamentals of EEG measurement. *Meas. Sci. Rev.* 2, 1–11.

Velayudhan, A., and Peter, S. (2016). Noise analysis and different denoising techniques of ECG signal: A survey. *IOSR J. Electron. Commun. Eng.* 1, 40–44.

frontiers | Frontiers in Human Neuroscience

TYPE Original Research
PUBLISHED 10 May 2023
DOI 10.3389/fnhum.2023.1160981

OPEN ACCESS

EDITED BY
Tetsuo Kida,
Institute for Developmental Research, Japan

REVIEWED BY
Juha M. Lahnakoski,
Helmholtz Association of German Research Centres, Germany
Avi Mendelsohn,
University of Haifa, Israel

*CORRESPONDENCE
Hossein Dini
hdi@create.aau.dk
Luis Emilio Bruni
leb@create.aau.dk

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

RECEIVED 08 February 2023
ACCEPTED 26 April 2023
PUBLISHED 10 May 2023

CITATION
Dini H, Simonetti A, Bigne E and Bruni LE (2023) Higher levels of narrativity lead to similar patterns of posterior EEG activity across individuals.
Front. Hum. Neurosci. 17:1160981.
doi: 10.3389/fnhum.2023.1160981

Higher levels of narrativity lead to similar patterns of posterior EEG activity across individuals

Hossein Dini[1*†], Aline Simonetti[2†], Enrique Bigne[2] and Luis Emilio Bruni[1*]

[1]The Augmented Cognition Lab, Aalborg University, Copenhagen, Denmark, [2]Department of Marketing and Market Research, University of Valencia, Valencia, Spain

Introduction: The focus of cognitive and psychological approaches to narrative has not so much been on the elucidation of important aspects of narrative, but rather on using narratives as tools for the investigation of higher order cognitive processes elicited by narratives (e.g., understanding, empathy, etc.). In this study, we work toward a scalar model of narrativity, which can provide testable criteria for selecting and classifying communication forms in their level of narrativity. We investigated whether being exposed to videos with different levels of narrativity modulates shared neural responses, measured by inter-subject correlation, and engagement levels.

Methods: Thirty-two participants watched video advertisements with high-level and low-level of narrativity while their neural responses were measured through electroencephalogram. Additionally, participants' engagement levels were calculated based on the composite of their self-reported attention and immersion scores.

Results: Results demonstrated that both calculated inter-subject correlation and engagement scores for high-level video ads were significantly higher than those for low-level, suggesting that narrativity levels modulate inter-subject correlation and engagement.

Discussion: We believe that these findings are a step toward the elucidation of the viewers' way of processing and understanding a given communication artifact as a function of the narrative qualities expressed by the level of narrativity.

KEYWORDS

narrativity level, engagement, inter-subject correlation, EEG, naturalistic stimuli

1. Introduction

Attention is involved in all cognitive and perceptual processes (Chun et al., 2011). To some degree, an attentive state toward an external stimulus implies the silencing of internally oriented mental processing (Dmochowski et al., 2012). Sufficiently strong attentive states can hamper conscious awareness of one's environment and oneself (Busselle and Bilandzic, 2009). In contrast to neutral stimuli, emotional stimuli attract greater and more focused attention (see Vuilleumier, 2005 for a review of the topic). As stories can generate and evoke strong feelings (Hogan, 2011), the pathos implicit in narratives could be considered attention-seeking stimuli. In fact, people tend to engage emotionally with stories (Hogan, 2011). Attentional focus to a narrative stimulates complex processing (Houghton, 2021), with narratives inducing "emotionally laden attention" (Dmochowski et al., 2012). In the

narrative domain, attentional focus and the sense of being absorbed into the story are part of narrative engagement (Busselle and Bilandzic, 2009). Indeed, researchers have suggested that narratives' inherent persuasiveness is related to the feelings of immersion they evoke, a phenomenon that has been termed "transportation effects" (Green and Brock, 2000). In this context, "transportation" indicates a combination of attention, feelings, and imagery where there is a convergent process of different perceptual, cognitive, and affective systems and capacities to the narrative events (Green and Brock, 2000). During narrative comprehension, involuntary autobiographical memories triggered by the story in question appear to impair attention only momentarily (Tchernev et al., 2021). This is not the case for either internally or externally generated daydreaming or distraction (Tchernev et al., 2021). Thus, narrative engagement can be hampered when thoughts unrelated to the narrative arise (Busselle and Bilandzic, 2009). In addition, the degree to which audiences engage with a narrative varies based on delivery modality (e.g., audio or visual) (Richardson et al., 2020), during narrative comprehension (Song et al., 2021a), and across individuals (Ki et al., 2016; Sonkusare et al., 2019). Nonetheless, narratives are naturally engaging (Sonkusare et al., 2019), and they reflect daily experiences, making them potentially useful devices for understanding cognitive processes such as engagement.

The use of simplified and abstract stimuli has been common in cognitive neuroscience for decades (e.g., Dini et al., 2022a). Although such stimuli enable highly controlled experiments and isolation of the study variables, they tend to lack ecological validity. Hence, recent studies have employed stimuli that simulate real-life situations, including narratives, termed "naturalistic stimuli" (see Sonkusare et al., 2019, for a review of the topic). From this perspective, advertising is considered a naturalistic stimulus that is designed to be emotionally persuasive (Sonkusare et al., 2019). Video ads provide narrative content in short period of time; hence, advertisers focus on delivering a key message through stories with different narrativity levels. In fact, companies use narrative-style ads because stories captivate, entertain, and involve consumers (Escalas, 1998; Coker et al., 2021). Researchers have found that, at the neural level, narratives (particularly more structured ones) tend to induce similar affective and cognitive states across viewers (Dmochowski et al., 2012; Song et al., 2021a). However, whether consumers perceive stories in similar ways is of great interest, as this might affect whether they ultimately engage with the advertisement as expected. To investigate the underlying cognitive processes generated by narratives, many studies have used traditional methods such as electroencephalogram (EEG) power analysis (Wang et al., 2016). Though valid, the metrics provided by traditional methods may not be ideal when considering narrative as a continuous stimulus, given that such methods could generally require stimulus repetition. In addition, they do not capture how the same information is processed across individuals.

Inter-subject correlation (ISC) is an appropriate neural metric for investigating shared neural responses, especially when using naturalistic stimuli, including media messages (see Schmälzle, 2022, for a discussion of the topic). This data-driven method assumes the occurrence of common brain reactions to a narrative, which improves the generalizability of the findings. By correlating neural data across individuals, this metric can identify localized neural activities that react to a narrative in a synchronous fashion (i.e., in a time-locked manner) (Nastase et al., 2019). ISC is well-suited to analysis of both functional magnetic resonance (fMRI) (Redcay and Moraczewski, 2020) and EEG (Petroni et al., 2018; Imhof et al., 2020) data. ISC obtained using naturalistic stimuli has been used to investigate episodic encoding and memory (Hasson et al., 2008; Cohen and Parra, 2016; Simony et al., 2016; Song et al., 2021a), social interaction (Nummenmaa et al., 2012), audience preferences (Dmochowski et al., 2014), information processing (Regev et al., 2019), narrative comprehension (Song et al., 2021b), and, similar to this study, attention and engagement (Dmochowski et al., 2012; Ki et al., 2016; Cohen et al., 2017; Poulsen et al., 2017; Imhof et al., 2020; Schmälzle and Grall, 2020; Grall et al., 2021; Song et al., 2021a; Grady et al., 2022).

Inter-subject correlation calculated from EEG data has been shown to predict levels of attentional engagement with auditory and audio-visual narratives (Ki et al., 2016; Cohen et al., 2017). Cohen et al. (2017) found that neural engagement with narratives measured through ISC was positively correlated with behavioral measurements of engagement, including real-world engagement. Another study demonstrated that dynamic ISC aligns with reported suspense levels of a narrative (Schmälzle and Grall, 2020). Attention was shown to modulate (attended) narrative processing at high levels of the cortical hierarchy (Regev et al., 2019). Ki et al. (2016) found that ISC was weaker when participants had to concurrently perform mental arithmetic and attend to a narrative than when participants only attended to the narrative. This was especially the case when participants attended to audiovisual narratives compared to auditory-only narratives. They also reported that the ISC difference between the two attentional states, as well as its magnitude, was more pronounced when the stimulus was a cohesive narrative than when it was a meaningless one. Moreover, they demonstrated that audiovisual narratives generated stronger ISC than audio-only narratives. In line with this finding, a recent fMRI study indicated that sustained attention network dynamics are correlated with engagement while attending audiovisual narratives (Song et al., 2021a). However, this is not the case for audio-only narratives. Nevertheless, in audio-only stimuli, ISC is higher for personal narratives compared to a non-narrative description or a meaningless narrative (i.e., a reversed version of the personal narratives) (Grall et al., 2021). Furthermore, Song et al. (2021a) found higher interpersonal synchronization of default mode network activity during moments of higher self-reported engagement with both audio-visual and audio-only narratives. In addition, highly engaging moments of a narrative evoke higher ISC in the mentalizing network compared to less engaging moments (Grady et al., 2022). Finally, ISC was observed to diminish when viewers watched video clips for the second time (Dmochowski et al., 2012; Poulsen et al., 2017; Imhof et al., 2020), which was supported by a study showing lower ISC when viewers were presented with a scrambled narrative for the second time compared to the first time (Song et al., 2021b).

In summary, while some previous studies have examined the relationship between ISC and narrative engagement (or its correlate, attention) (Dmochowski et al., 2012; Ki et al., 2016; Cohen et al., 2017; Poulsen et al., 2017; Regev et al., 2019; Schmälzle and Grall, 2020; Grall et al., 2021; Song et al., 2021a; Grady et al., 2022), they assumed that "being a narrative" is an either-or quality. For example, a cohesive narrative is considered a narrative, whereas a meaningless narrative (e.g., a scrambled narrative) is not considered one. From the narratological perspective, there

has been increased interest in "narrativity" as a scalar property; that is, artifacts (pictures, videos, stories, and other forms of representation) may have different degrees of narrativity. In other words, narrative may be a matter of more-or-less rather than either-or (Ryan, 2007). Therefore, how different levels of narrativity affect narrative engagement remains an open question. Moreover, whether narrativity levels lead to different degrees of ISC, and whether it is possible to predict narrativity level based on ISC values, has not yet been explored. Therefore, our main research question for this study is as follows: does narrativity level [high (HL) vs. low (LL)] modulate self-reported engagement and ISC? To answer this question, we asked 32 participants to watch 12 narrative, real video ads twice each (with sound removed) varying in their degree of narrativity. We defined "narrativity level" as the degree to which an artifact is perceived to evoke a complete narrative script (Ryan, 2007). Essentially, this refers the degree to which a story seems complete in all its structural elements, such as the presence of defined characters in a stipulated context, clear causal links between events, and a sense of closure resulting from the intertwining of these events (Ryan, 2007). While participants watched the videos, we recorded their EEG brain signals to investigate shared neural responses to the video ads across participants; that is, we investigated the interpersonal reliability of neural responses (expressed by ISC). We further assessed how narrativity level affected self-reported engagement with the ads. Thus, we asked participants to rate how much each ad captured their attention and how immersed in it they felt. Furthermore, we examined how engagement and ISC differ for video ads watched for the first time versus the second time. Based on the aforementioned studies, which explore how narratives engage individuals and indicate that ISC levels are higher during moments of higher narrative engagement and, we postulate that HL video ads will have higher ISC than LL video ads and that self-reported engagement will be higher for HL ads than for LL ads. Moreover, we expect that both engagement and ISC will be higher the first time the videos are seen compared to the second time.

2. Materials and methods

This study was approved by the local ethics committee (Technical Faculty of IT and Design, Aalborg University) and performed in accordance with the Danish Code of Conduct for Research and the European Code of Conduct for Research Integrity. All participants signed a written informed consent form at the beginning of the session, were debriefed at the end of the experiment, and were given a symbolic payment as a token of gratitude for their time and effort. Data collection took place throughout November 2020. This study was part of a larger study; here, we report only the information relevant to the present study.

2.1. Participants and stimuli

We recruited 32 (13 women) right-handed participants of 16 nationalities between the ages of 20 and 37 (M = 26.84, SD = 4.33). Regarding occupation, 69% were students, 16% were workers, and 15% were both. Regarding educational level, 12% had completed or were completing a bachelor's degree, and 88% had completed or were completing a master's degree. We requested that participants not drink caffeine products at least 2 h prior the experimental session.

To select the stimuli, we initially pre-selected 22 video ads (eight brands; **Figure 1A**) varying in the number of elements that constitute different narrative levels according to Ryan (2007). The narrativity levels of these video ads were then preliminarily rated from 0 to 100 by an expert in the narrative field. Based on these scores, we selected six brands that had the largest difference in score between two of their ads (to maximize the difference in narrativity level between two ads of the same brand), coinciding with an absolute score below 50 for one ad and above 50 for the other ad (**Figure 1B**). We then grouped the video ads into two categories: HL and LL narrativity. Although not categorizing the videos—thus considering continuity in narrativity level—would better align with a scalar model of narrativity, our results would suffer from statistical invalidity. Because the videos were real and therefore not created specifically for the study, they had idiosyncratic features separate from the narrativity level (e.g., different scenario, characters, and plot). These features could either cover a potential effect of narrativity level or lead to a misattribution of an effect. By categorizing the videos into two narrativity levels, we sought to mitigate the potential effects of those individual features and make the differences in narrativity level more salient. The final selection consisted of six video ads with HL narrativity (M = 83.17, SD = 10.70) and another six with LL narrativity (M = 33.33, SD = 10.70), with each of the six brands contributing one video to each of the two groups (**Figure 1C**). The videos can be seen here: https://youtu.be/RBMo0wvFKoY. To validate the categorization of the ads into these two categories, another independent expert in the narrative field rated the narrativity level of the 12 video ads from 0 to 100, and their ratings were consistent with those of first expert (HL: M = 81.00, SD = 13.70; LL: M = 37.33, SD = 15.37; **Figure 1D**). To assess whether the general public would also perceive a difference in narrativity level between the two categories of videos, we conducted an online validation test using the Clickworker platform[1] with an independent panel of participants, which confirmed that our classification was valid (see section "3. Results"). For the online test, a total of 156 participants were assigned to one of four groups. Each group was assigned to watch three video ads, each of which was from a distinct brand. To capture the degree of perceived narrativity in each video ad, participants were asked to answer five questions used by Kim et al. (2017), such as "the commercial tells a story" and "the commercial shows the main actors or characters in a story." All questions were rated from 0 (strongly disagree) to 100 (strongly agree). Average responses to these five questions showed whether the videos ads were indeed perceived to be in the HL (higher scores) or LL (lower scores) category. The responses confirmed that our initial categorization of the videos was valid (see results and **Figure 1E**).

The stimuli therefore comprised 12 2D video ads from six brands (Barilla, Coke, Disney, Kellogg's, Nike, and Oculus), with two video ads from each brand, one for each narrativity level. We selected different brands to mitigate the influence of brand and product category. The ads were

1 https://www.clickworker.com

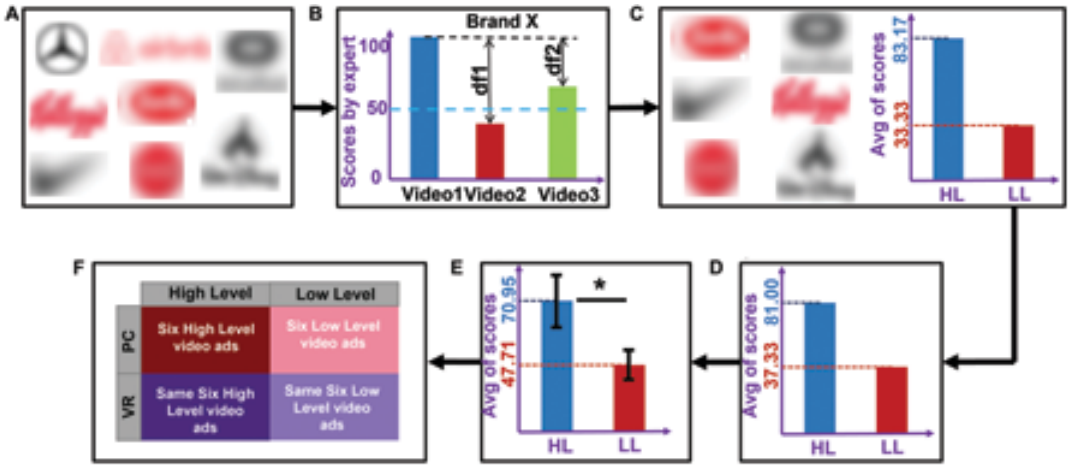


FIGURE 1
Video ads selection procedure, division into high (HL) and low (LL) categories, results of the ratings, and the resulting design. **(A)** The eight pre-selected brands from which we chose 22 video ads. **(B)** Process of selecting videos based on expert's rating. Suppose that, from brand X, we pre-selected three videos that were rated by the expert. The videos that had the maximum difference (df1) in score, for which one score was >50 and the other was <50, were selected as the final videos from brand X. **(C)** The six selected brands with average scores. For each brand, the video with a higher score was assigned to the HL category (six videos) and the video with a lower score was assigned to the LL category (six videos). **(D)** Average of the second expert's ratings across the six selected videos. **(E)** Average scores (across participants and stimuli) for narrativity level, rated by 124 participants. An independent sample *t*-test showed that the narrativity scores of the HL videos were significantly higher than those of the LL videos ($p < 0.001$). The black error bars represent the standard deviation of the sample. **(F)** The 2 × 2 experiment design consisted of the following conditions: HL–PC, HL–VR, LL–PC, and LL–VR. *$p < 0.05$.

real commercials retrieved from YouTube. We removed the audio and edited some of the videos slightly to adjust the length (which varied from 57 to 63 s). We selected non-verbal narratives in video format because motion picture narratives are less susceptible to interindividual differences and generate more homogeneous experiences across individuals than verbal (oral) narratives (Jajdelska et al., 2019). This may be because the visual images are directly related to the narrative content, which reduces personal interpretations of the story (Richardson et al., 2020).

2.2. Design, data collection, and task

We conducted a 2 × 2 full factorial within-subjects study with two levels of narrativity [high level (HL) vs. low level (LL)] and two media [computer screen (PC) vs. virtual reality (VR)]. The study included four conditions: six HL video ads presented on a PC (HL–PC), six HL video ads presented in VR (HL–VR), six LL video ads presented on a PC (LL–PC), and six LL video ads presented in VR (LL–VR). See **Figure 1F**.

Initially, the EEG device (32-channel, 10–20 system) was placed on each participant's scalp, and the impedance of the active electrodes was set to less than 25 kΩ using a conductive gel, according to the manufacturer's guidelines. The signals were recorded at a sampling rate of 500 Hz using Brain Products software. The HTC Vive Pro VR headset was placed on top of the EEG electrodes, and the impedance of the electrodes was checked again. The task comprised two sessions of approximately 25 min each, separated by a 20 min interval. During one session, participants watched ads on a PC; during the other session, they watched ads in VR. During each session, participants watched all of the video ads and answered a questionnaire after each. A 2-s fixation cross was shown before each new video. The videos were displayed first on a PC for half of the participants. The video presentation order was counterbalanced across participants but the same across media. The ten-item questionnaire included two questions of interest for this study: (i) "this commercial really held my attention" and (ii) "this ad draws me in"; both were scored from 0 (strongly disagree) to 100 (strongly agree). We retrieved these questions from the "being hooked" scale (Escalas et al., 2004), and they were meant to capture consumers' sustained attention to the advertisement (Escalas et al., 2004). High levels of focused attention and high levels of immersion (feelings of transportation) can indicate high levels of narrative engagement. Thus, we calculated our engagement metric by averaging attention and immersion scores.

2.3. Preprocessing

We performed all preprocessing and processing steps using Matlab R2020b (The Math Works, Inc, Natick, MA, USA) with in-house codes and tools from the FieldTrip 20210128[2] and EEGLAB 2021.0[3] toolboxes. To remove high and low frequency noise, we applied a third-order Butterworth filter with 1–40 Hz cut-off frequencies to the raw data. Next, we detected bad channels using an automated rejection process with voltage threshold +500 μV and, after confirmation by an expert, rejected them

2 http://fieldtriptoolbox.org
3 https://eeglab.org/

from the channel list. We then interpolated the removed channels using the spherical spline method based on the activity of six surrounding channels in the FieldTrip toolbox. The average number of rejected channels per participant was 1.92 + 1.48. One participant was excluded from the analyses due to having more than five bad channels, and two participants were excluded due to missing trials. Subsequently, we segmented the filtered EEG data corresponding to the 12 video ads and concatenated the data. We then excluded the EEG data corresponding to the time when participants were answering the questionnaires. Next, we conducted independent component analysis on the concatenated data to remove remaining noise. We estimated source activity using the second-order blind identification method. We then identified eye-related artifacts and other noisy components, which were confirmed by an expert and removed from the component list. Applying the inverse independent component analysis coefficients to the remaining components, we obtained the denoised data. Finally, we re-referenced the denoised data to the average activity of all electrodes.

2.4. Inter-subject correlation

To evaluate whether the evoked responses to the stimuli were shared among participants, we calculated the ISC of neural responses to the video ads by calculating the correlation of EEG activity among participants. Researchers have established that, for an evoked response across participants (or trials) to be reproduceable, it is necessary that each participant (trial) provides a reliable response (Ki et al., 2016). In this sense, ISC is similar to traditional methods that capture the reliability of a response by measuring the increase in magnitude of neural activity; the important difference is that, by measuring reliability across participants, we avoid presenting stimuli multiple times to a single participant. In addition, this method is compatible with continuous naturalistic stimuli such as video ads. Therefore, the goal of ISC analysis is to identify correlated EEG components that are maximally shared across participants. By "EEG component," we mean a linear combination of electrodes, which can be considered "virtual sources" (Ki et al., 2016). This correlated component analysis method is similar to principal component analysis, for which, instead of the maximum variance within a dataset, the maximum correlation among datasets is considered. Like other component extraction methods, this method identifies components by solving a eigenvalue problem (Parra and Sajda, 2003; de Cheveigné and Parra, 2014). Below, we explain our ISC calculation procedure for multiple stimuli [provided in Cohen and Parra (2016) and Ki et al. (2016)].

To construct the input data, we combined data from participants in all four conditions (six videos for each condition; see **Figure 2A** for one stimulus) and obtained 24 three-dimensional

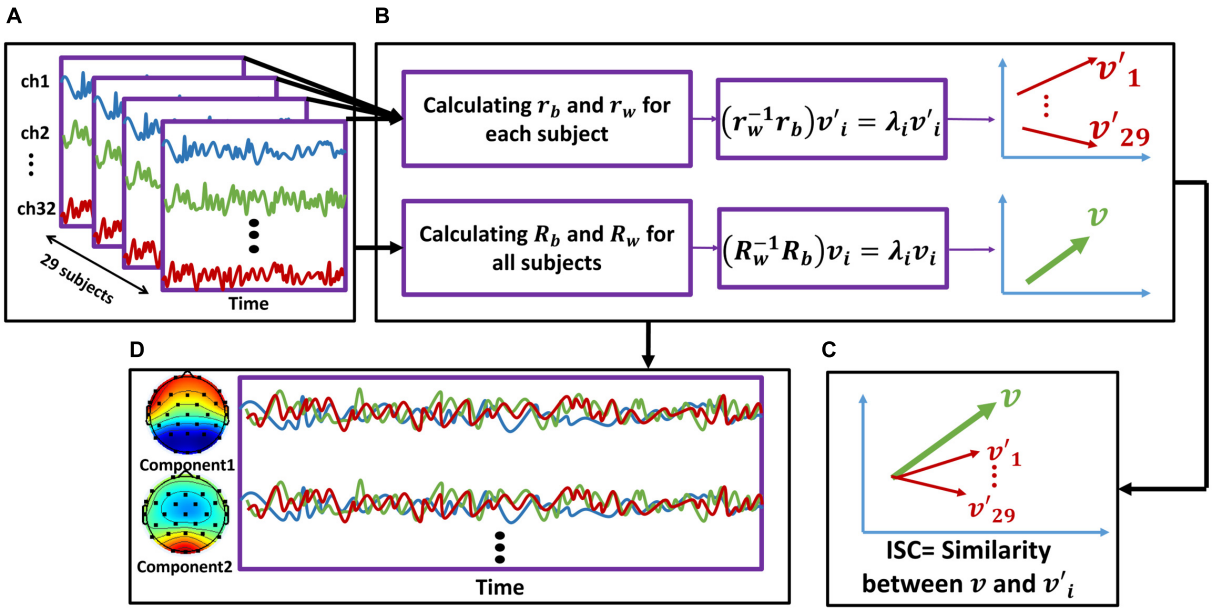


FIGURE 2
The inter-subject correlation (ISC) calculation and component activity estimation procedure. We repeated this procedure separately for each stimulus. **(A)** We first concatenated the EEG activity of all participants together. (**B**: upper line) For each subject, we calculated the r_b and r_w [using (Eqs. 1, 2), respectively] based on the cross-covariance matrices. By solving the eigenvalue problem (Eq. 3), we computed eigenvectors of each subject. The red arrows represent each subject's eigenvector (v′). Therefore, we obtained 29 v′, corresponding to the number of participants. Note that each v'_i is a 32-dimensional matrix corresponding to the number of electrodes (for simplicity, the vectors are shown in two dimensions). (**B**: lower line) We summed all the r_b across all participants to calculate R_b and did the same for r_w to calculate R_w. Then, by solving the eigenvalue problem, we obtained v_i, which represents maximal correlation across participants. Note that v is a 32-dimensional matrix corresponding to the number of electrodes (for simplicity, the vectors are shown in two dimensions). **(C)** We calculated the similarity between the representative vector of all participants (v) and each of the vectors corresponding to each subject (v'_1–v'_{29}) to obtain ISC. **(D)** Using the calculated eigenvectors and forward model, we calculated the scalp activity of each component, which are linear combinations of electrode activity (32 components were calculated with the order from strongest to weakest). In sum, the activity of first two strongest components is displayed here.

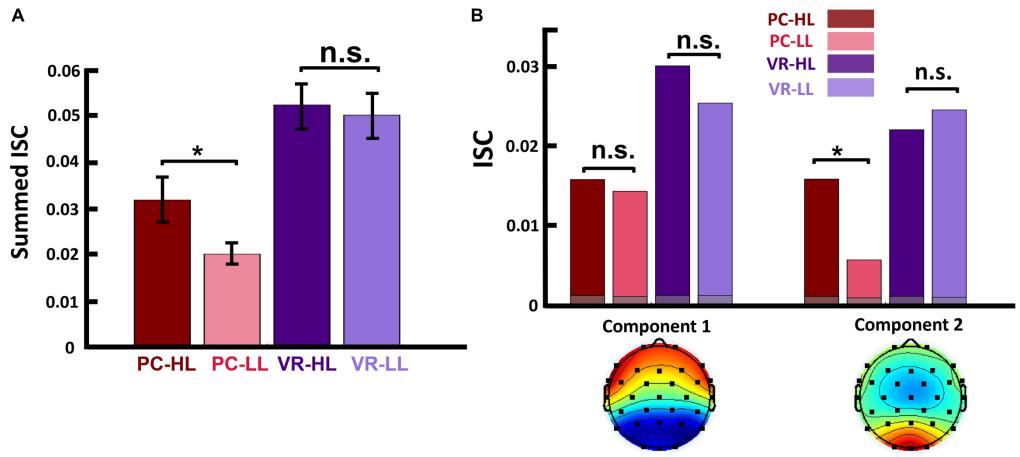


FIGURE 3
(A) Summed ISC over the first two strongest components in four conditions. This shows that, in the PC condition, the ISC of HL video ads was significantly higher than that of LL video ads (p = 0.029), and in the VR condition, the ISC of HL video ads was still higher than that of LL video ads, but not statistically significant (p = 0.638). (**B**-top) Calculated ISC separately for each component and within each condition. This shows that, for the first component, the difference between HL and LL video ads is not significant in either the PC or the VR condition. However, in the second component, the ISC of HL video ads was significantly higher than that of LL video ads (p = 0.002) in PC condition, but not in the VR condition (p = 0.313). The gray bars show the average over phase randomized iterations. (**B**-bottom) Scalp activity of first two strongest components. Activity in component 1 is throughout the anterior and posterior regions, while in component 2, activity is concentrated in the posterior region. In both figures, the colored dots represent the calculated ISC values for each participant. *p < 0.05.

EEG matrices (channel × data samples × participants). For each of these 24 matrices, we separately calculated the between-subject cross-covariance using Eq. 1:

$$R_b = \frac{1}{N(N-1)} \sum_{k=1}^{N} \sum_{l=1, l \neq k}^{N} R_{kl} \quad (1)$$

where R_{kl} is as follows:

$$R_{kl} = \sum_{t} (x_k(t) - \bar{x}_k)(x_l(t) - \bar{x}_l)^T$$

R_{kl} indicates cross-covariance among all electrodes of subject k with all electrodes of subject l. The matrix $x_k(t)$ contains 32 electrode activities (pre-processed data) of subject k measured in time, and $\bar{x}_k$ is the average of $x_k(t)$ over time. Additionally, we separately calculated the within-subject cross-covariance for each of the abovementioned matrices using Eq. 2:

$$R_w = \frac{1}{N} \sum_{k=1}^{N} R_{kk} \quad (2)$$

R_{kk} is calculated in an identical manner to R_{kl}, except it considers only the electrode activity of subject k. We then summed the calculated R_b and R_w of all 24 matrices to obtain the pooled within- and between- subject cross-covariances, representing data on all stimuli.

Then, to improve the robustness of our analysis against outliers, we used the shrinkage regularization method to regularize the within subject-correlation matrix (Blankertz et al., 2011) Eq. 3:

$$R_w \leftarrow (1-\gamma) R_w + \gamma \bar{\lambda} I \quad (3)$$

where $\bar{\lambda}$ is the average of eigenvectors of R_w, with γ equal to 0.5.

Next, by solving the eigenvalue problem for $R_w^{-1} R_b$ (Eq. 4), we obtained eigenvectors v_i. Such eigenvectors are projections indicating the maximum correlation among participants, from strongest to weakest, provided by eigenvalues λ_i.

$$\left(R_w^{-1} R_b\right) v_i = \lambda_i v_i \quad (4)$$

After this step, we obtained the projections that are maximally correlated among all participants considering all stimuli (**Figure 2B**). Next, to measure the reliability of individual participants' EEG responses, we calculated the correlation of projected data for each subject with projected data for the group separately for each stimulus (**Figure 2C**). This metric shows how similar the brain activity of a single subject was to that of all other participants. Next, we calculated the ISC using the correlation of such projections for each stimulus (video ads) and component, averaged across all possible combinations of participants. Thus, for each stimulus and each subject, we obtained a component matrix. By summing the first two strongest components, we obtained the ISC for each subject, as follows (Eq. 5):

$$ISC_k = \sum_{i=1}^{2} r_{ki} \quad (5)$$

where

$$r_{ki} = \frac{1}{N(N-1)} \sum_{l=1, l \neq k}^{N} \frac{\sum_t y_{ik}(t) y_{il}(t)}{\sqrt{\sum_t y_{ik}^2(t)} \sqrt{\sum_t y_{il}^2(t)}}$$

r_{ki} is the Pearson correlation coefficient averaged across all pairs of participants applied to component projection y_{ik}, which is defined as follows (Eq. 6):

$$y_{ik}(t) = v_i^T (x_{ik}(t) - x_k) \quad (6)$$

Next, we summed the first two strongest components to represent the ISC of each subject for each stimulus (i = 1:2) and

ignored the weaker components because they were below chance level in one condition. Finally, we summed the calculated ISC over the six stimuli of each condition to obtain the ISC of all participants in each of the four conditions. To illustrate the scalp activity of each component, we used the corresponding forward model following previous studies (Parra et al., 2005; Haufe et al., 2014; **Figure 2D**). In **Figure 3A**, the bars illustrate the sum of the two first components.

To determine chance-level ISC, we first built up phase-randomized EEG data using a method that randomizes EEG signal phases in the frequency domain (Theiler et al., 1992; Ki et al., 2016). Using this method, we obtained new time series where the temporal alterations were not necessarily aligned with the original signals and therefore not correlated across participants. We then implemented the aforementioned steps on the randomized data, identically: for each condition, we computed within- and between-subject cross-covariances, projected the data on eigenvectors, calculated ISC, and summed them over stimuli.

We generated 5000 sets of such randomized data and continued the process as described above to obtain a null distribution representing the random ISC activity. Next, we tested the actual ISC values against the null distribution to evaluate whether the actual ISCs are above the chance level. To do this, we used a two-tailed significance test with p = (1 + number of null ISC values $\geq$ empirical ISC) / (1 + number of permutations). See **Supplementary Figure 1**.

Some studies have reported that ISC values are higher in lower frequency bands and vice versa (Lankinen et al., 2014; Thiede et al., 2020). Although in this study we do not aim to evaluate the effect of narrativity levels on the ISC of different frequency bands, it is worth exploring it. For this, we filtered the EEG data into different frequency bins using the abovementioned Butterworth filter. The frequency bins comprise 1–2 Hz, 2–3 Hz, 3–4 Hz, 4–5 Hz, 5–6 Hz, 6–7 Hz, alpha (8–12 Hz), low-beta (13–20 Hz), and high-beta (21–40 Hz). We employed a high resolution in the lower frequency bands to verify whether our results aligned with previous findings in these frequency bands (Lankinen et al., 2014). After filtering the signal, we repeated the same procedure of ISC calculation as mentioned above and obtained the ISC for the frequency bands.

To explore whether the ISC differences are derived by either change in engagement scores or a change in narrativity levels, we implemented two correlation analyses: one for ISC vs. engagement scores and one for ISC vs. narrativity levels. To do this, we used a partial correlation approach considering subjects and engagement/level of narrativity as covariates depending on the correlation (e.g., for the correlation of ISC vs. engagement, the level of narrativity is considered as a covariate).

2.5. ISC based on a model trained only by HL data (ISC-HL)

We also tested whether we could assign a single participant to the HL or LL group based on participant's neural brain activity. In other words, we tested whether we could predict the level of narrativity the participant was exposed based on EEG activity when participants attended to HL and LL videos. To do so, we calculated the projections v_i using only the data from the HL group. We then calculated the ISC of each participant (from both the HL and LL groups) based on the projections calculated only from the HL group and called the result "ISC-HL." Therefore, in each group, each participant had an ISC-HL value showing how similar their neural activity was to the activity of all participants in the HL group. The difference between ISC-HL and ISC is that, while calculating ISC, we explored how similar the activity of a participant was to their own group (whether HL or LL). When calculating ISC-HL, however, we calculated how similar the activity of participants in both groups (HL and LL) was to the activity of only the participants in the HL group. We hypothesized that the brain activity of participants from the HL group would be more similar to the overall activity of all participants in the HL group (meaning higher ISC-HL) and that the brain activity of participants from the LL group would be less similar to the overall activity of all participants in the HL group (meaning lower ISC-HL). To avoid bias in this procedure, we excluded the test subject from the HL group while computing the projections v_i and then calculated the ISC-HL for the corresponding subject. As described in the previous section, we summed the two first stronger components of ISC-HL and also summed the six stimuli for each condition. Additionally, we assessed classification performance for both groups (HL vs. LL) using the area under the curve (AUC) characteristics of a fitted support vector machine model with leave-one-out approach. To determine the chance-level AUC, we randomly shuffled the labels 1,000 times. In each iteration, AUC was calculated through an identical process, starting from extracting the correlated components of HL-labeled group and following all subsequent steps described above. We calculated p-values using (1 + number of null AUC values $\geq$ actual AUC) / (1 + number of permutations) to see whether the observed AUC is above chance level. We repeated all the aforementioned steps separately for the PC and VR conditions.

2.6. Statistical analysis

To test statistical differences across narrativity levels, we conducted a two-way ANOVA. Based on the study design, our independent variables were narrativity level (HL vs. LL) and medium (PC vs. VR). The dependent variables were the scores from the two questionnaires (i.e., attention and immersion) and calculated neural reliability (as expressed by ISC). The output of such a comparison includes the main effects of "narrativity level" and "medium" and the interaction effect of "narrativity level × medium." The same procedure and dependent variables, but with narrativity level (HL vs. LL) and viewing order (first vs. second) as independent variables, were used to test statistical differences across viewing order. Participants watched each video twice: once in one medium and once in the other medium. The medium that was used first was counterbalanced across participants. For the analysis, we created a group for the first viewing and a group for the second viewing, regardless of the medium used. In both of these groups, participants watched the same videos, including both HL and LL video ads. In the present study, we focused on the effect of narrativity level and viewing order on brain activity patterns by comparing the calculated features in HL vs. LL conditions, and first vs. second viewing. Therefore, we do not report results for the main effect of medium. Nevertheless, we included medium type as a factor in the statistical analyses to control for its effect. Additionally, we performed *post hoc* analysis when we identified interaction effects but also when we

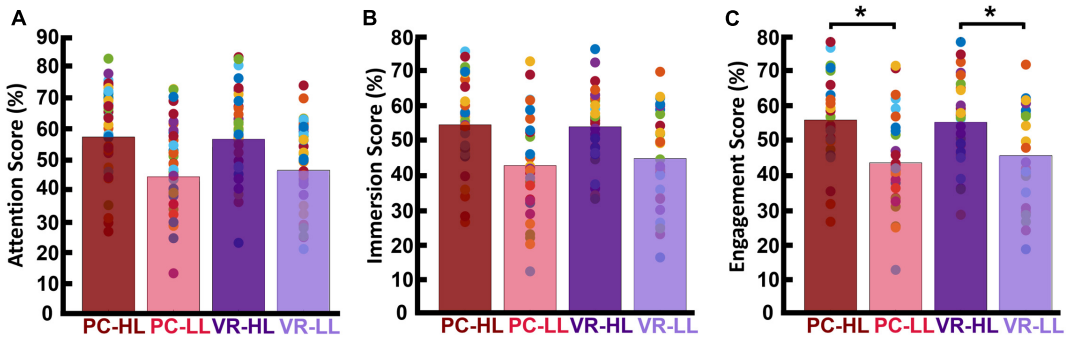


FIGURE 4
Participants' self-reported scores. Each bar is the average of the scores across participants and across corresponding video ads. The colored dots represent the score for each participant averaged across the corresponding videos. **(A)** Self-reported attention scores. **(B)** Self-reported immersion scores. **(C)** Engagement scores (average of attention and immersion scores). This shows that the HL engagement scores were significantly higher than the LL engagement scores in both PC ($p < 0.001$) and VR ($p < 0.001$) conditions. *$p < 0.05$.

judged it appropriate to report simple effects. We corrected the *p*-values of the *post-hoc* tests using the Bonferroni method. Finally, to avoid multiple comparison error of multiple statistical tests (such as comparing HL and LL within two extracted components; see results), we corrected the *p*-values based on the number of statistical test repetition errors (FDR correction) using the Benjamini-Hochberg method (Gerstung et al., 2014).

3. Results

3.1. Online validation test

Of the 156 participants who completed the online test, we considered 124 responses to be valid as the remaining participants answered too fast or failed to answer the attention question correctly. Results of the online test confirmed that the general public also perceived ads in the HL category as having a high narrativity level (M_{HL} = 70.95, SD_{HL} = 10.79) and adds in the LL category as having a low narrativity level (M_{LL} = 47.71, SD_{LL} = 13.52; see **Figure 1E** in the section "2. Materials and methods"). All six of the HL ads received higher scores than any of the six LL ads. An independent sample *t*-test confirmed that the difference between the two categories was statistically significant [$t(61) = 7.554$, $p < 0.001$].

3.2. Narrativity level modulates self-reported engagement

We first report our descriptive analysis of the attention, immersion, and engagement scores for each of the four conditions (HL–PC, HL–VR, LL–PC, and LL–VR; see section "2. Materials and methods" for abbreviations), then report results of the statistical analysis conducted on engagement scores. Note that our engagement metric is the average of the attention and immersion scores. Attention scores for the PC conditions were as follows: M_{HL-PC} was 57.03 (SD_{HL-PC} = 14.76), and M_{LL-PC} was 44.16 (SD_{LL-PC} = 13.92). The attention scores for the VR conditions were as follows: M_{HL-VR} was 56.26 (SD_{HL-VR} = 14.19), and M_{LL-VR} was 46.28 (SD_{LL-VR} = 14.41). **Figure 4A** shows the average attention scores in each of the four conditions. Immersion scores for the PC conditions were as follows: M_{HL-PC} was 54.26 (SD_{HL-PC} = 13.28), and M_{LL-PC} was 42.53 (SD_{LL-PC} = 15.09). Immersion scores for the VR conditions were as follows: M_{HL-VR} was 53.67 (SD_{HL-VR} = 10.81), and M_{LL-VR} was 44.56 (SD_{LL-VR} = 14.15). **Figure 4B** shows the average immersion scores in each of the four conditions. Engagement scores for the PC conditions were as follows: M_{HL-PC} was 55.65 (SD_{HL-PC} = 13.54), and M_{LL-PC} was 43.35 (SD_{LL-PC} = 14.04). Engagement scores for the VR conditions were as follows: M_{HL-VR} was 54.97 (SD_{HL-VR} = 12.16), and M_{LL-VR} was 45.42 (SD_{LL-VR} = 14.00). Results of the 2 × 2 repeated measures ANOVA showed a main effect of narrativity level on perceived engagement [$F(1,28) = 19.779$, $p < 0.001$]. *Post hoc* analysis showed that participants engaged more with HL ads than with LL ads in both PC [$F(28,1) = 17.03$, $p < 0.001$] and VR [$F(28,1) = 20.320$, $p < 0.001$] conditions. See **Figure 4C**.

3.3. Narrativity level modulates posterior inter-subject correlation

First, we tested whether the actual ISC values are significantly above the chance level and the results indicated that all the calculated ISCs are significantly higher than the generated null distribution (see **Supplementary Figure 1** for detailed information). Next, we applied the same statistical analysis used for the self-reported data to assess whether our neurological metric (ISC) changed according to narrativity level. To do so, we used summed ISC over the first two strongest components (see section "2. Materials and methods") as a representative of neural reliability. Results of the 2 × 2 repeated measures ANOVA showed a main effect of narrativity level, where ISC was higher when participants watched the HL ads than when they watched the LL ads [$F(1,28) = 4.467$, $p = 0.044$]. Although the interaction effect was not significant, we then conducted *post hoc* analysis to determine whether this difference occurred in the PC or VR condition. Results of the *post hoc* analysis showed that ISC was significantly higher for the HL ads (M_{HL} = 0.031, SD_{HL} = 0.019)

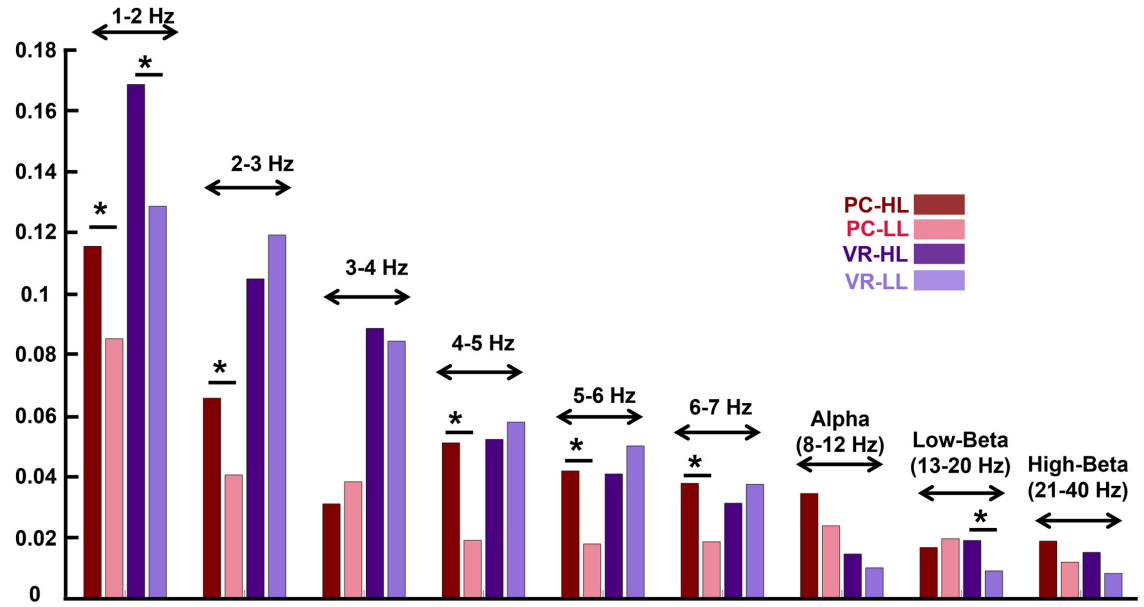


FIGURE 5
Inter-subject correlation of different frequency bins for the four conditions (see the figure legend). The stars above the bars indicate a significant difference between the two conditions ($p < 0.05$). The frequency bands are indicated above the bar charts of the four conditions. The figure indicates that in the lower frequency bands, the ISC values are higher compared to the ISC of higher frequency bands. In addition, the effect of the narrativity level on ISC values in the frequency bands is the same as its effect in wide-band EEG: In the PC condition, the ISC of HL is higher than the ISC of LL [in all frequency bins except for 3–4 Hz and low-beta; in alpha and high-beta, it is marginally significant ($p = 0.052$; $p = 0.056$, respectively)]. *$p < 0.05$.

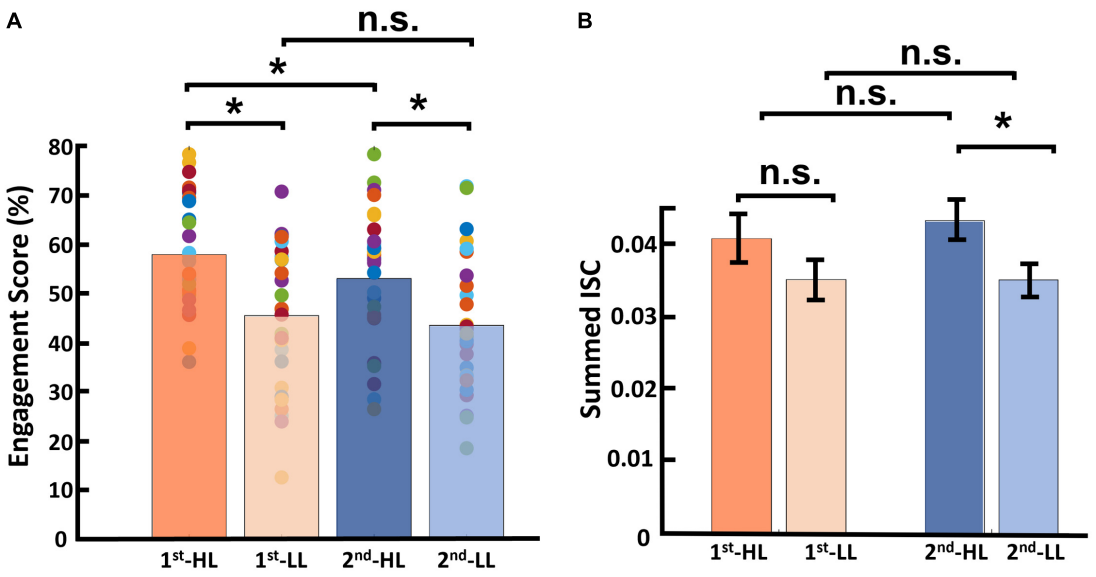


FIGURE 6
Results of viewing order (first and second) considering narrativity level (HL and LL). **(A)** Results of engagement scores, where for videos in the HL condition, engagement scores upon first viewing were significantly higher than those upon second viewing. Moreover, HL videos received significantly higher engagement scores than LL videos upon both first and second viewings. **(B)** Results for calculated ISC of viewing order, where the only significant difference was between the second viewing of HL videos and that of LL videos. *$p < 0.05$.

than for the LL ads ($M_{LL} = 0.020$, $SD_{LL} = 0.016$) in the PC condition [$F(28,1) = 5.266$, $p = 0.017$]. In the VR condition, the ISC of HL ads ($M_{HL} = 0.052$, $SD_{HL} = 0.027$) was still higher than that of LL ads ($M_{LL} = 0.050$, $SD_{LL} = 0.023$), but this difference was not significant [$F(1,28) = 0.226$, $p = 0.638$]. See **Figure 3A**.

Furthermore, to investigate which factors affect neural reliability, we compared the ISC of HL video ads to that of LL video ads separately for each component. We corrected the p-values because we repeated our calculations twice for the two components, and results can be seen in the top part of **Figure 3B**. Results showed

that, for the first component, the main effect of narrativity is not significant [$F(1,28) = 1.145$, *corrected* $p = 0.294$]. However, for the second component, the main effect of narrativity was significantly higher in HL video ads than in LL video ads [$F(1,28) = 5.412$, *corrected* $p = 0.045$]. *Post hoc* analysis of the second component revealed that the ISC of HL ads ($M_{HL} = 0.015$, $SD_{HL} = 0.010$) was significantly higher than that of LL ads ($M_{LL} = 0.005$, $SD_{LL} = 0.009$) in the PC condition [$F(1,28) = 13.676$, *corrected* $p = 0.002$]; however, there was no significant difference between HL and LL ads in the VR condition [$F(1,28) = 1.056$, *corrected* $p = 0.313$]. In addition, different correlated components had different patterns of scalp activities (**Figure 3B**, bottom part).

The correlation analysis revealed that for the correlation between ISC and engagement, neither on the PC ($r = -0.101$, *corrected-p* = 0.421) nor on the VR ($r = 0.047$, corrected-$p = 0.722$) conditions there was a statistically significant correlation (**Supplementary Figure 2A**). However, for the correlation between ISC and narrativity level, the PC condition showed a significant correlation ($r = 0.314$, *corrected-p* = 0.017), and the VR condition did not show a significant correlation ($r = 0.055$, *corrected-p* = 0.678), as shown in **Supplementary Figure 2B**.

3.4. Narrativity level modulates ISC of different frequency bands

The results for the ISC in different frequency bands indicate similar effects of narrativity level on ISC of different frequencies as of on a wide EEG band. They are illustrated in **Figure 5**.

In the PC condition, the ISC of some frequency bands are significantly higher in the HL compared to the LL condition as follows: 2−3 Hz ($M_{HL} = 0.066$, $SD_{HL} = 0.054$, $M_{LL} = 0.041$, $SD_{LL} = 0.045$, $F(1,28) = 2.207$, *corrected* $p = 0.031$), 4–5 Hz [$M_{HL} = 0.051$, $SD_{HL} = 0.458$, $M_{LL} = 0.198$, $SD_{LL} = 0.33$, $F(1,28) = 3.846$, *corrected* $p < 0.001$], 5–6 Hz [$M_{HL} = 0.042$, $SD_{HL} = 0.035$, $M_{LL} = 0.018$, $SD_{LL} = 0.037$, $F(1,28) = 2.478$, *corrected* $p = 0.016$], and 6–7 Hz [$M_{HL} = 0.038$, $SD_{HL} = 0.029$, $M_{LL} = 0.019$, $SD_{LL} = 0.035$, $F(1,28) = 2.232$, *corrected* $p = 0.029$]. However, some of the frequency bands did not show significant differences between HL and LL conditions as follows: 1–2 Hz ($M_{HL} = 0.116$, $SD_{HL} = 0.069$, $M_{LL} = 0.085$, $SD_{HL} = 0.071$, $F(1,28) = 1.639$, *corrected* $p = 0.106$), 3–4 Hz [$M_{HL} = 0.031$, $SD_{HL} = 0.048$, $M_{LL} = 0.039$, $SD_{LL} = 0.043$, $F(1,28) = -0.694$, *corrected* $p = 0.490$], alpha ($M_{HL} = 0.035$, $SD_{HL} = 0.022$, $M_{LL} = 0.024$, $SD_{LL} = 0.020$, $F(1,28) = 2.011$, *corrected* $p = 0.052$), low-beta ($M_{HL} = 0.017$, $SD_{HL} = 0.019$, $M_{LL} = 0.020$, $SD_{LL} = 0.021$, $F(1,28) = -0.539$, *corrected* $p = 0.592$), and high-beta ($M_{HL} = 0.019$, $SD_{HL} = 0.013$, $M_{LL} = 0.012$, $SD_{LL} = 0.013$, $F(1,28) = 1.946$, *corrected* $p = 0.056$).

In the VR condition, in ISC of 1–2 Hz ($M_{HL} = 0.169$, $SD_{HL} = 0.097$, $M_{LL} = 0.129$, $SD_{LL} = 0.060$, $F(1,28) = 2.264$, *corrected* $p = 0.027$) and low-beta ($M_{HL} = = 0.019$, $SD_{HL} = 0.017$, $M_{LL} = 0.009$, $SD_{LL} = 0.016$, $F(1,28) = 2.261$, *corrected* $p = 0.027$) there is a significant difference between HL and LL conditions. However, the other frequency bands did not show any significant difference between H and LL conditions as follows: 2–3 Hz [$M_{HL} = 0.169$, $SD_{HL} = 0.097$, $M_{LL} = 0.129$, $SD_{LL} = 0.060$, $F(1,28) = -1.111$, *corrected* $p = 0.271$], 3–4 Hz [$M_{HL} = 0.089$, $SD_{HL} = 0.036$, $M_{LL} = 0.085$, $SD_{LL} = 0.050$, $F(1,28) = 0.444$, *corrected* $p = 0.658$], 4–5 Hz [$M_{HL} = 0.052$, $SD_{HL} = 0.045$, $M_{LL} = 0.058$, $SD_{LL} = 0.041$, $F(1,28) = -0.538$, *corrected* $p = 0.592$], 5–6 Hz [$M_{HL} = 0.041$, $SD_{HL} = 0.037$, $M_{LL} = 0.050$, $SD_{LL} = 0.034$, $F(1,28) = -0.993$, *corrected* $p = 0.324$], 6–7 Hz [$M_{HL} = 0.032$, $SD_{HL} = 0.028$, $M_{LL} = 0.038$, $SD_{LL} = 0.029$, $F(1,28) = -0.889$, *corrected* $p = 0.377$], alpha [$M_{HL} = 0.153$, $SD_{HL} = 0.026$, $M_{LL} = 0.100$, $SD_{LL} = 0.022$, $F(1,28) = 0.723$, *corrected* $p = 0.472$], and high-beta [$M_{HL} = 0.015$, $SD_{HL} = 0.014$, $M_{LL} = 0.009$, $SD_{LL} = 0.013$, $F(1,28) = 1.955$, *corrected* $p = 0.055$].

3.5. Viewing order modulates self-reported engagement but not inter-subject correlation

To evaluate whether viewing order affects engagement scores and ISC, we implemented the abovementioned statistical procedure with narrativity level and viewing order as independent variables. We first report the resulting engagement scores and ISC scores.

Regarding engagement scores, the average and standard deviation for the first viewing group were as follows: M_{HL} was 57.71 ($SD_{HL} = 2.10$), and M_{LL} was 45.42 ($SD_{LL} = 2.65$). For the second viewing group, these figures were as follows: M_{HL} was 52.90 ($SD_{HL} = 2.57$), and M_{LL} was 43.36 ($SD_{LL} = 2.56$; see **Figure 6A**). Results of this statistical analysis revealed a significant main effect of viewing order. Further *post hoc* analysis indicated that this main effect was driven by the HL category, for which videos viewed first received significantly higher engagement scores than those viewed second [$F(1,28) = 12.784$, $p = 0.001$]. However, in the LL condition, there was no significant difference between the first and second viewing groups [$F(1,28) = 1.686$, $p = 0.205$]. Moreover, there was a main effect of narrativity level: of videos in the first viewing group, HL videos received significantly higher engagement scores than LL videos [$F(1,28) = 23.047$, $p < 0.001$]; this was also the case for videos in the second viewing group [$F(1,28) = 13.391$, $p = 0.001$].

Regarding ISC, the average and standard deviation for videos in the first viewing group were as follows: $M_{HL} = 0.040$ ($SD_{HL} = 0.025$), and M_{LL} was 0.035 ($SD_{LL} = 0.027$). For the second viewing group, these figures were as follows: M_{HL} was 0.043 ($SD_{HL} = 0.027$), and M_{LL} was 0.035 ($SD_{LL} = 0.023$; see **Figure 6B**). Results of this statistical analysis showed no significant main effect of viewing order. Further *post hoc* analysis showed that the HL video ads received higher ISC scores than LL video ads upon second viewing [$F(1,28) = 4.411$, $p = 0.045$], but not upon first viewing [$F(1,28) = 1.144$, $p = 0.294$].

3.6. ISC based on a model trained only by HL data (ISC-HL)

We calculated ISC-HL to test if it was possible to determine whether a single subject was watching an HL video ad or an LL video ad based on their neural activity. **Figures 7A, B** display the calculated ISC-HL in both VR and PC conditions. As expected, the neural responses of 23 (76.6%) participants were much more similar to those of the HL group (had higher ISC-HL) when they attended to HL videos than when they attended to LL videos in the PC condition (**Figure 7A**). In the VR condition, the

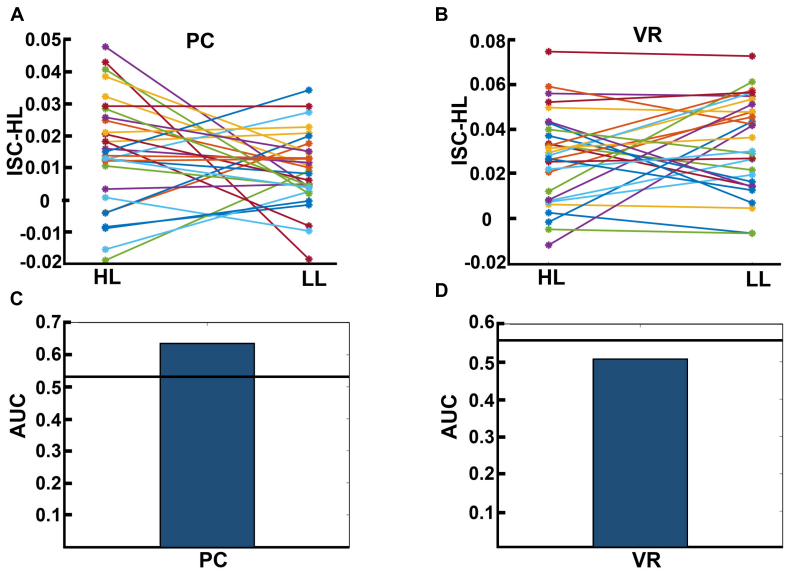


FIGURE 7
(A,B) Results for calculated ISC-HL for each participant in the PC and VR conditions. Lines and markers of each color show the ISC-HL of one participant in the HL and LL conditions. **(A)** This shows that 76.6% of participants had higher ISC-HL when they attended to HL video ads than when they attended to LL video ads. This shows that most of the participants who were exposed to HL video ads showed higher similarity to the trained model (which is derived from data of HL condition) than those who are exposed to LL video ads. **(B)** This relates to the same evaluation as panel **(A)**, but in the VR condition, where 50% of participants had higher ISC-HL. **(C,D)** Area under the curve (AUC) of classification performance for predicting exposure to HL and LL video ads in PC and VR conditions, respectively. The black lines show the average over randomized AUCs. **(C)** Shows that the AUC of classification based on participants' neural activity in the PC condition is higher than chance level, while panel **(D)** shows that the AUC of classification based on participants' neural activity in VR condition is below chance level.

neural responses of 15 (50%) participants showed higher ISC-HL while attending to HL videos than while attending to LL videos (**Figure 7B**). Given ISC-HL as the predictor of narrativity level, the classifier shows above-chance performance in the PC condition but not in the VR condition (**Figures 7C, D**). In the PC condition, the actual AUC was 0.637, while the chance-level AUC was 0.548 ($p < 0.001$). In the VR condition, the actual AUC was 0.507, while the chance-level AUC was 0.543 ($p = 0.245$).

4. Discussion

In this study, we tested whether different levels of narrativity (HL vs. LL) lead to differences in information processing reliability (represented by ISC), specifically while watching video ads. Furthermore, we evaluated whether different levels of narrativity cause differences in self-reported engagement ratings. To this aim, we presented HL and LL video ads to 32 participants while collecting their EEG signals. In addition, for each video ad, participants self-reported their levels of attention and immersion, which we considered two core factors of engagement. We calculated the ISC of each participant by calculating the similarity of their correlated components to those of the participant pool. One advantage of ISC analysis is that it does not require stimulus repetition; this is advantageous because such repetition causes decreased attention and engagement (Dmochowski et al., 2012; Ki et al., 2016). As expected, our results showed that both calculated ISC and engagement scores of HL video ads were significantly higher than those for LL video ads, suggesting that narrativity level modulates ISC and engagement. However, the modulation of ISC was not correlated with the degree of engagement with the narrative.

Previous studies have evaluated the relationship between narrative engagement—or its core component, attention—and ISC. Cohen et al. (2017) measured self-reported engagement scores and ISC while participants were exposed to naturalistic videos. They found that more engaging videos were processed uniformly in participants' brains, leading to higher ISC. Consistent with their results, Poulsen et al. (2017) reported that a lack of engagement manifests an unreliable neural response (meaning lower ISC), and they introduced EEG-ISC as a marker of engagement. Song et al. (2021a) investigated whether engagement ratings modulate ISC by evaluating the relationship between continuous self-reported engagement ratings and ISC using continuous naturalistic stimuli. They reported higher ISC during highly narrative engaging moments and concluded that ISC reflects engagement levels. In line with it, Schmälzle and Grall (2020) demonstrated that dynamic ISC aligns with reported levels of suspense (i.e., a proxy for engagement levels) of a narrative, and Grady et al. (2022) showed that ISC is higher during intrinsic engaging moments of a narrative. In another study, Dmochowski et al. (2012) used short video clips to evaluate attention and emotion using ISC. They found a close correspondence between expected engagement and neural correlation, suggesting that extracting maximally correlated components (ISC) reflects cortical processing of attention or emotion. Finally, Ki et al. (2016) investigated whether attentional states modulate ISC for audio and audiovisual narratives, and they concluded that higher attention leads to higher neural reliability across subjects. Inspired by previous studies (Busselle

and Bilandzic, 2009; Dmochowski et al., 2012; Lim et al., 2019), we measured narrative engagement by averaging two of its important components, attention and immersion.

Our results indicated that, using ISC (in the PC condition), we were able to significantly discriminate levels of narrativity: the ISC levels of HL video ads were significantly higher than those of LL video ads (**Figure 3A**). Consistent with our narrativity level discrimination, self-reported engagement ratings were significantly higher for HL video ads than for LL ones (**Figure 4C**). In addition, in the PC condition, the ISC-HL was able to predict the narrativity level based on neural responses (measured via EEG) to the stimuli. This means that the model used significantly predicted whether the participants were attending to HL or LL narrativity. The level of significance was not as strong as the previously reported accuracy for predicting attentional states (Ki et al., 2016); however, this could be due to explicit differences between experimental conditions (participants were made to count backward during the task to diminish their attentional state). This finding is the first step toward predicting exposure to different levels of narrativity based solely on the ISC of neural activity evoked by stimuli, which leads to a better comprehension of the brain mechanism that processes narrativity levels. Further studies with different designs should be conducted with a more specific focus on predicting narrativity level.

In contrast with other studies, we did not find a relationship between engagement levels and ISC, as the correlation analysis suggested. A plausible explanation for this finding is the type of stimulus employed. Several studies compared ISC scores of interrupted, scrambled, or reversed narratives with those of non-modified narratives. Dmochowski et al. (2012) reported higher ISC for non-scrambled narratives than for disrupted or scrambled narratives. Ki et al. (2016), Poulsen et al. (2017), and Grall et al. (2021) found higher ISC scores for a cohesive narrative than for a meaningless, scrambled or reversed narrative. These findings demonstrate that meaningless narratives are linked to lower ISC scores. Song et al. (2021b) used scrambled video clips to identify the moments in which narrative comprehension occurs, and they reported that story comprehension occurs when events are causally related to each other. Moreover, they showed that, in such moments, the underlying brain states were mostly correlated across subjects. The stimuli manipulations applied by previous studies (e.g., scrambling the narrative or presenting narratives that lacked causality) created greater differences between experimental conditions than those in our study. While we shared with these studies their interest in the subjects' engagement and other factors related to the reception of the narrative, we were also interested in investigating the possibility of classifying and discriminating structural characteristics of narrative artifacts. Therefore, we defined the narrativity levels (HL and LL) based on narrative structural elements and properties such as the presence of defined characters in an identifiable context, clear causal links between events, and closure resulting from the intertwining of these events. Our conceptualization of narrativity levels was inspired by Ryan (2007).

From this perspective, our results align with those of previous studies, indicating that information processing is highly consistent across subjects when participants are exposed to HL narrativity. In other words, the fact that the ISC of HL video ads is significantly higher than that of LL video ads in the PC condition (**Figure 3A**) is a promising indication that EEG-ISC could potentially be used to explore levels of narrativity (and correlative engagement) in different kinds of media artifacts. However, unlike previous studies using stimuli that were "either-or" regarding narrativity possession, our ISC was not mediated by engagement levels. This finding suggests that ISC can represent or capture other cognitive processes beyond engagement. The evaluation of the activity of the first two components separately (**Figure 3B**) provides some insights in this regard. For the first component, activity was distributed throughout the anterior and posterior regions, while in the second component, activity was concentrated in the posterior region. Overall, the second component drove the narrativity level modulation in response to the HL narrative video ads. While previous studies relating ISC to engagement found more widespread ISC (e.g., Dmochowski et al., 2012, 2014; Cohen et al., 2017; Song et al., 2021a), strong posterior ISC was linked to shared psychological perspective (Lahnakoski et al., 2014) and shared understanding of narrative (Nguyen et al., 2019). Lahnakoski et al. (2014) asked participants to take one perspective or another to interpret the events of a movie. When participants watched the movie and adopted the same perspective, posterior ISC was stronger than when they adopted different perspectives. The authors posited that ISC represented a shared understanding of the environment. This was supported by the study by Nguyen et al. (2019), which showed that participants with similar recalls of a narrative had stronger ISC in the posterior medial cortex and angular gyrus compared to those with dissimilar recalls. In our case, high levels of narrativity better immersed participants in the story world compared to low levels. This might have eased participants to take the perspective of the character(s) in HL videos. In addition, videos with low narrativity levels might not have been so successful in leading to similar perspective taking and shared understanding because the story was more fragmented than in HL video ads. Because ISC appears to be related to both engagement and shared understanding or perspective-taking, our findings suggest that ISC seems to be more sensitive to the latter than to the former (see Dini et al., 2022c). However, further investigation is necessary to elucidate why narrativity level did not appear to affect ISC in the same way in the VR condition. Chang et al. (2015) reported that ISC might change due to fatigue effects for participants in an fMRI scanner. Thus, one possible explanation for the non-significant effect of narrativity level on ISC in the VR condition is fatigue. Wearing the VR headset and the EEG cap for almost 25 min might have caused fatigue and discomfort (e.g., related to posture, weight, or itching), and such an effect might have masked the effects of narrativity level on ISC. Another plausible explanation is also related to the VR feature. As VR is an increasingly popular immersive technology, participants' expectations about the VR modality may have affected their level of attention. Participants may have been disappointed to find a 2D stimulus that perhaps failed to meet their expectations and therefore diminished their active attention.

In the narratology field, there is a substantial, ongoing debate about whether narrative should be considered an either-or property or a scalar property (i.e., a matter of more or less) (Ryan, 2007; Abbott, 2008). Therefore, in the last two decades, some narratologists have introduced the notion that different artifacts may have different levels of narrativity (Ryan, 2007; Oatley, 2011). It is widely accepted that, even though events in a story do not have to be chronologically ordered, the sequence of the events must nonetheless follow a narrative logic if closure is to be

achieved (this is something that is commonly encapsulated in the distinction between "story" and "discourse"). Therefore, a "scrambled" narrative ceases to be a narrative; instead, it is a series of unrelated events, lacking an internal temporal logic. If there is no discernable, recoverable chronological order of connected events, the sequence could hardly be considered a narrative. Although the previous studies presented in our literature review provided the methodological bases for our study, they were not centered on narrative qualities. In our study, we focus on investigating the plausibility of testing a particular expressive artifact for levels of narrativity. Therefore, our stimuli presented two different degrees of narrativity (LL and HL). We showed that small differences in narrativity levels (i.e., between HL and LL) have effects on ISC and engagement similar to the effects of more evident differences between such levels (e.g., scrambled vs. non-scrambled). However, the underlying reason for differences in ISC between narratives with some degree of narrativity level did not reflect differences in perceived engagement; rather, it seemed related to perspective taking and shared understanding. These findings are a step toward narrative comprehension, especially when considering narrative as a scalar property (Bruni et al., 2021).

Previous studies have reported that ISC values are sensitive to the frequency information of the input signal (Lankinen et al., 2014; Thiede et al., 2020), meaning that ISC values are higher in lower frequency bands and vice versa. In this study, we used a wide EEG band (1–40 Hz) for the analysis, and the results revealed lower ISC values compared to other studies (e.g., Ki et al., 2016). Therefore, we evaluated whether the ISC values change according to different frequency bands: 1–2 Hz, 2–3 Hz, 3–4 Hz, 4–5 Hz, 5–6 Hz, 6–7 Hz, alpha (8–12 Hz), low-beta (13–20 Hz), and high-beta (21–40 Hz). In line with previous studies (Lankinen et al., 2014), our results showed higher ISC values in lower frequency bands and lower ISC values in higher frequency bands (**Figure 5**). The possible explanation for this could be that in high frequencies, even small timing variations can substantially decrease the correlation across signals, leading to a decrease in ISC values. In addition, it is also possible that these differences derive from phase variations in higher frequencies across subjects. Although this study does not aim to elaborate on the effect of narrativity levels on ISC of different frequency bands, our results point to an interesting finding. Aligned with the results obtained from the wide-band analysis, we found that the ISC of HL is significantly higher than LL in different frequency bins. Even though some frequency bins did not show this effect, in most of them there is a tendency of higher ISC for the HL than the LL condition. In summary, the analysis of ISC for different frequency bins supported the findings of the wide-band ISC in terms of the effect of narrativity on shared neural responses across participants.

To test whether viewing order modulates engagement and ISC, we separated the data into two groups—first and second viewing—and conducted statistical analyses considering narrativity level and viewing order to be independent variables. Dmochowski et al. (2012) found that attentional engagement decreases when participants watch a stimulus for the second time compared to the first time. Moreover, they reported significantly lower ISC for the second viewing. Ki et al. (2016) confirmed and extended their results by declaring that neural responses become less reliable upon second viewing of a stimulus, showing significantly lower ISC. In an fMRI study, Song et al. (2021b) investigated the effect of viewing order on neural activity, watching scrambled videos twice. Their results replicated the findings of the aforementioned study and other studies (Poulsen et al., 2017; Imhof et al., 2020) by showing that neural states across participants were less synchronized when watching the videos, including the scrambled videos, for the second time. However, Chang et al. (2015) conducted a combined EEG-magnetoencephalography (MEG) study and reported increased ISC during the second viewing of the stimuli. They declared that participants' prediction of the story structure increased during the second viewing, resulting in a more similar EEG-MEG activity across participants, and that this contradiction with previous studies might be due to the different time window selected for ISC calculation. Our results showed that engagement level dropped during the second viewing, supporting previous studies (Dmochowski et al., 2012). For the HL videos ads, engagement scores for the first viewing were significantly higher than those for the second viewing, although this was not the case for the LL ads (**Figure 6A**). The decrease in engagement scores for HL ads but not LL ads might be due to differences intrinsic to the classification of our stimuli. While videos in the HL category included all or almost all narrativity elements (Ryan, 2007), those in the LL category included only a few. Thus, we could say that the HL video ads were more storytelling-based than the LL video ads. Therefore, it could be that factors such as suspense, expectation, and suspension of disbelief, which naturally occur when watching stories, are attenuated when the story is viewed for the second time, lowering engagement levels in the case of HL video ads. These same factors are not present or are present to a lesser extent when watching LL video ads, and therefore, engagement levels were not harmed. However, there was no significant difference in ISC between the first and second viewing in either HL or LL video ads, suggesting that viewing order does not modulate ISC. Though the current dataset cannot provide definite answers, these differences in findings might be explained by our study design. First, our participants were exposed to 12 short video ads, a greater number of stimuli than previous studies employed when investigating viewing order (Dmochowski et al., 2012; Chang et al., 2015; Ki et al., 2016; Poulsen et al., 2017; Song et al., 2021b). Watching 12 videos in sequence might have reduced participants' memories of details of the stories. Hence, when watching the videos for the second time (although they were perceived as less engaging) there remained a substantial amount of information to be processed, which could have been reflected in the ISC of the second viewing. Another possible factor that might have hindered participants' short-term memories of the videos and affected ISC levels is the time span and the tasks performed between the two viewings. In our study, the second exposure to the video ads was separated from the first by about 20 min. During those 20 min, participants performed two different tasks for a separate study. Therefore, the current dataset was not able to capture the neural underpinnings of viewing order. Considering the limitations of this study, future studies could be conducted to capture this effect.

5. Conclusion

This study investigated whether high or low narrativity levels in video advertisement would significantly affect

self-reported engagement and shared neural responses across individuals, measured through EEG-ISC, with the ads. The findings demonstrated that a higher narrativity level led to an increase in engagement with the ad and an increase in ISC of the posterior part of the brain, also during a second viewing. Interestingly, the results suggest that ISC may be more sensitive to shared perspective-taking—which could also indicate a shared understanding of the narrative—than engagement levels. Moreover, the findings imply that narratives with higher narrativity levels evoke similar interpretations in their audience compared to narratives with lower narrativity levels. This study advances the elucidation of the viewers' way of processing and understanding a given communication artifact as a function of the narrative qualities expressed by the level of narrativity.

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 Local Ethics Committee (Technical Faculty of IT and Design, Aalborg University) and performed in accordance with the Danish Code of Conduct for Research and the European Code of Conduct for Research Integrity. The patients/participants provided their written informed consent to participate in this study.

Author contributions

AS, HD, LB, and EB designed the experiment. AS and HD collected the data and wrote the main manuscript. HD analyzed the neuro data and prepared the figures. AS analyzed the self-reported data. LB and EB contributed to writing the manuscript. All authors revised and approved the manuscript.

Funding

This work was supported by the Rhumbo (European Union's Horizon 2020 Research and Innovation Program under the Marie Skłodowska-Curie Grant Agreement No. 813234).

Acknowledgments

We are grateful to Tirdad Seifi Ala for helping with the technical aspects of the data analysis and revising an early version of the manuscript, and Thomas Anthony Pedersen for providing technical support and help during the early stages of the study. This study is available as a preprint (Dini et al., 2022b).

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.2023.1160981/full#supplementary-material

References

Abbott, H. P. (2008). *The Cambridge introduction to narrative*, 2nd Edn. Cambridge: Cambridge University Press, doi: 10.1017/CBO9780511816932

Blankertz, B., Lemm, S., Treder, M., Haufe, S., and Müller, K.-R. (2011). Single-trial analysis and classification of ERP components—a tutorial. *Neuroimage* 56, 814–825. doi: 10.1016/j.neuroimage.2010.06.048

Bruni, L. E., Dini, H., and Simonetti, A. (2021). "Narrative cognition in mixed reality systems: Towards an empirical framework," in *Virtual, augmented and mixed reality*, eds J. Y. C. Chen and G. Fragomeni (Cham: Springer International Publishing), 3–17. doi: 10.1007/978-3-030-77599-5_1

Busselle, R., and Bilandzic, H. (2009). Measuring narrative engagement. *Media Psychol.* 12, 321–347. doi: 10.1080/15213260903287259

Chang, W.-T., Jääskeläinen, I. P., Belliveau, J. W., Huang, S., Hung, A.-Y., Rossi, S., et al. (2015). Combined MEG and EEG show reliable patterns of electromagnetic brain activity during natural viewing. *Neuroimage* 114, 49–56. doi: 10.1016/j.neuroimage.2015.03.066

Chun, M. M., Golomb, J. D., and Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annu. Rev. Psychol.* 62, 73–101. doi: 10.1146/annurev.psych.093008.100427

Cohen, S. S., Henin, S., and Parra, L. C. (2017). Engaging narratives evoke similar neural activity and lead to similar time perception. *Sci. Rep.* 7:4578. doi: 10.1038/s41598-017-04402-4

Cohen, S. S., and Parra, L. C. (2016). Memorable audiovisual narratives synchronize sensory and supramodal neural responses. *eNEURO* 3, ENEURO.0203–16.2016. doi: 10.1523/ENEURO.0203-16.2016

Coker, K. K., Flight, R. L., and Baima, D. M. (2021). Video storytelling ads vs argumentative ads: How hooking viewers enhances consumer engagement. *J. Res. Interact. Mark.* 15, 607–622. doi: 10.1108/JRIM-05-2020-0115

de Cheveigné, A., and Parra, L. C. (2014). Joint decorrelation, a versatile tool for multichannel data analysis. *Neuroimage* 98, 487–505. doi: 10.1016/j.neuroimage.2014.05.068

Dini, H., Simonetti, A., Bigne, E., and Bruni, L. E. (2022a). EEG theta and N400 responses to congruent versus incongruent brand logos. *Sci. Rep.* 12:4490. doi: 10.1038/s41598-022-08363-1

Dini, H., Simonetti, A., and Bruni, L. E. (2022c). Exploring the neural processes behind narrative engagement: An EEG study. *bioRxiv* [Preprint]. doi: 10.1101/2022.11.28.518174

Dini, H., Simonetti, A., Bigne, E., and Bruni, L. E. (2022b). Higher levels of narrativity lead to similar patterns of posterior EEG activity across individuals. *bioRxiv* [Preprint]. doi: 10.1101/2022.09.23.509168

Dmochowski, J. P., Bezdek, M. A., Abelson, B. P., Johnson, J. S., Schumacher, E. H., and Parra, L. C. (2014). Audience preferences are predicted by temporal reliability of neural processing. *Nat. Commun.* 5:4567. doi: 10.1038/ncomms5567

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

Escalas, J. E. (1998). "Advertising narratives: What are they and how do they work?," in *Representing consumers*, ed. B. Stern (Oxfordshire: Routledge), 267–289. doi: 10.4324/9780203380260-19

Escalas, J. E., Moore, M. C., and Britton, J. E. (2004). Fishing for Feelings? Hooking viewers helps!. *J. Consum. Psychol.* 14, 105–114. doi: 10.1207/s15327663jcp1401\&2_12

Gerstung, M., Papaemmanuil, E., and Campbell, P. J. (2014). Subclonal variant calling with multiple samples and prior knowledge. *Bioinformatics* 30, 1198–1204. doi: 10.1093/bioinformatics/btt750

Grady, S. M., Schmälzle, R., and Baldwin, J. (2022). Examining the relationship between story structure and audience response. *Projections* 16, 1–28. doi: 10.3167/proj.2022.160301

Grall, C., Tamborini, R., Weber, R., and Schmälzle, R. (2021). Stories collectively engage listeners' brains: Enhanced intersubject correlations during reception of personal narratives. *J. Commun.* 71, 332–355. doi: 10.1093/joc/jqab004

Green, M. C., and Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *J. Pers. Soc. Psychol.* 79, 701–721. doi: 10.1037/0022-3514.79.5.701

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

Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J.-D., Blankertz, B., et al. (2014). On the interpretation of weight vectors of linear models in multivariate neuroimaging. *Neuroimage* 87, 96–110.

Hogan, P. C. (2011). *Affective narratology: The emotional structure of stories*. Nebraska: University of Nebraska Press.

Houghton, D. M. (2021). Story elements, narrative transportation, and schema incongruity: A framework for enhancing brand storytelling effectiveness. *J. Strateg. Mark* 1–16. doi: 10.1080/0965254X.2021.1916570

Imhof, M. A., Schmälzle, R., Renner, B., and Schupp, H. T. (2020). Strong health messages increase audience brain coupling. *Neuroimage* 216:116527. doi: 10.1016/j.neuroimage.2020.116527

Jajdelska, E., Anderson, M., Butler, C., Fabb, N., Finnigan, E., Garwood, I., et al. (2019). Picture this: A review of research relating to narrative processing by moving image versus language. *Front. Psychol.* 10:1161. doi: 10.3389/fpsyg.2019.01161

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

Kim, E., Ratneshwar, S., and Thorson, E. (2017). Why narrative ads work: An integrated process explanation. *J. Advert.* 46, 283–296. doi: 10.1080/00913367.2016.1268984

Lahnakoski, J. M., Glerean, E., Jääskeläinen, 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

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

Lim, S., Yeo, M., and Yoon, G. (2019). Comparison between concentration and immersion based on EEG analysis. *Sensors* 19:1669. doi: 10.3390/s19071669

Nastase, S. A., Gazzola, V., Hasson, U., and Keysers, C. (2019). Measuring shared responses across subjects using intersubject correlation. *Soc. Cogn. Affect. Neurosci.* 14, 669–687. doi: 10.1093/scan/nsz037

Nguyen, M., Vanderwal, T., and Hasson, U. (2019). Shared understanding of narratives is correlated with shared neural responses. *Neuroimage* 184, 161–170. doi: 10.1016/j.neuroimage.2018.09.010

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. doi: 10.1073/pnas.1206095109

Oatley, K. (2011). *Such stuff as dreams: The psychology of fiction*. Hoboken, NJ: John Wiley & Sons.

Parra, L., and Sajda, P. (2003). Blind source separation via generalized eigenvalue decomposition. *J. Mach. Learn. Res.* 4, 1261–1269.

Parra, L. C., Spence, C. D., Gerson, A. D., and Sajda, P. (2005). Recipes for the linear analysis of EEG. *Neuroimage* 28, 326–341.

Petroni, A., Cohen, S. S., Ai, L., Langer, N., Henin, S., Vanderwal, T., et al. (2018). The variability of neural responses to naturalistic videos change with age and sex. *eNEURO* 5, ENEURO.0244–17.2017. doi: 10.1523/ENEURO.0244-17.2017

Poulsen, A. T., Kamronn, S., Dmochowski, J., Parra, L. C., and Hansen, L. K. (2017). EEG in the classroom: Synchronised neural recordings during video presentation. *Sci. Rep.* 7:43916. doi: 10.1038/srep43916

Redcay, E., and Moraczewski, D. (2020). Social cognition in context: A naturalistic imaging approach. *Neuroimage* 216:116392. doi: 10.1016/j.neuroimage.2019.116392

Regev, M., Simony, E., Lee, K., Tan, K. M., Chen, J., and Hasson, U. (2019). Propagation of information along the cortical hierarchy as a function of attention while reading and listening to stories. *Cereb. Cortex* 29, 4017–4034. doi: 10.1093/cercor/bhy282

Richardson, D. C., Griffin, N. K., Zaki, L., Stephenson, A., Yan, J., Curry, T., et al. (2020). Engagement in video and audio narratives: Contrasting self-report and physiological measures. *Sci. Rep.* 10:11298. doi: 10.1038/s41598-020-68253-2

Ryan, M.-L. (2007). "Toward a definition of narrative," in *The Cambridge companion to narrative*, ed. D. Herman (Cambridge: Cambridge University Press), 22–36. doi: 10.2196/24588

Schmälzle, R. (2022). Theory and method for studying how media messages prompt shared brain responses along the sensation-to-cognition continuum. *Commun. Theory* 32, 450–460. doi: 10.1093/ct/qtac009

Schmälzle, R., and Grall, C. (2020). The coupled brains of captivated audiences. *J. Media Psychol.* 32, 187–199. doi: 10.1027/1864-1105/a000271

Simony, E., Honey, C. J., Chen, J., Lositsky, O., Yeshurun, Y., Wiesel, A., et al. (2016). Dynamic reconfiguration of the default mode network during narrative comprehension. *Nat. Commun.* 7:12141. doi: 10.1038/ncomms12141

Song, H., Finn, E. S., and Rosenberg, M. D. (2021a). Neural signatures of attentional engagement during narratives and its consequences for event memory. *Proc. Natl. Acad. Sci. U.S.A.* 118:e2021905118. doi: 10.1073/pnas.2021905118

Song, H., Park, B., Park, H., and Shim, W. M. (2021b). Cognitive and neural state dynamics of narrative comprehension. *J. Neurosci.* 41, 8972–8990. doi: 10.1523/JNEUROSCI.0037-21.2021

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

Tchernev, J. M., Collier, J., and Wang, Z. (2021). There and back again? Exploring the real-time cognitive journey of narrative transportation. *Commun. Res.* 50, 312–337. doi: 10.1177/00936502211018577

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., and Farmer, J. D. (1992). Testing for nonlinearity in time series: The method of surrogate data. *Phys. D Nonlinear Phenom.* 58, 77–94.

Thiede, A., Glerean, E., Kujala, T., and Parkkonen, L. (2020). Atypical MEG inter-subject correlation during listening to continuous natural speech in dyslexia. *Neuroimage* 216:116799. doi: 10.1016/j.neuroimage.2020.116799

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

Wang, R. W. Y., Chang, Y.-C., and Chuang, S.-W. (2016). EEG spectral dynamics of video commercials: Impact of the narrative on the branding product preference. *Sci. Rep.* 6:36487. doi: 10.1038/srep36487

frontiers | Frontiers in Human Neuroscience

TYPE Original Research
PUBLISHED 25 May 2023
DOI 10.3389/fnhum.2023.1168108

OPEN ACCESS

EDITED BY
Hidehiko Okamoto,
International University of Health and Welfare (IUHW), Japan

REVIEWED BY
Pietro Aricò,
Sapienza University of Rome, Italy
Assaf Harel,
Wright State University, United States

*CORRESPONDENCE
Daniel E. Callan
✉ dcallan@atr.jp

RECEIVED 17 February 2023
ACCEPTED 27 April 2023
PUBLISHED 25 May 2023

CITATION
Callan DE, Fukada T, Dehais F and Ishii S (2023) The role of brain-localized gamma and alpha oscillations in inattentional deafness: implications for understanding human attention. *Front. Hum. Neurosci.* 17:1168108. doi: 10.3389/fnhum.2023.1168108

The role of brain-localized gamma and alpha oscillations in inattentional deafness: implications for understanding human attention

Daniel E. Callan[1,2]*, Takashi Fukada[1,3], Frédéric Dehais[2] and Shin Ishii[1,3]

[1]Brain Information Communication Research Laboratory, Advanced Telecommunications Research Institute International, Kyoto, Japan, [2]Institut Supérieur de l'Aéronautique et de l'Espace, University of Toulouse, Toulouse, France, [3]Graduate School of Informatics, Kyoto University, Kyoto, Japan

Introduction: The processes involved in how the attention system selectively focuses on perceptual and motor aspects related to a specific task, while suppressing features of other tasks and/or objects in the environment, are of considerable interest for cognitive neuroscience. The goal of this experiment was to investigate neural processes involved in selective attention and performance under multi-task situations. Several studies have suggested that attention-related gamma-band activity facilitates processing in task-specific modalities, while alpha-band activity inhibits processing in non-task-related modalities. However, investigations into the phenomenon of inattentional deafness/blindness (inability to observe stimuli in non-dominant task when primary task is demanding) have yet to observe gamma-band activity.

Methods: This EEG experiment utilizes an engaging whole-body perceptual motor task while carrying out a secondary auditory detection task to investigate neural correlates of inattentional deafness in natural immersive high workload conditions. Differences between hits and misses on the auditory detection task in the gamma (30–50 Hz) and alpha frequency (8–12 Hz) range were carried out at the cortical source level using LORETA.

Results: Participant auditory task performance correlated with an increase in gamma-band activity for hits over misses pre- and post-stimulus in left auditory processing regions. Alpha-band activity was greater for misses relative to hits in right auditory processing regions pre- and post-stimulus onset. These results are consistent with the facilitatory/inhibitory role of gamma/alpha-band activity for neural processing. Additional gamma- and alpha-band activity was found in frontal and parietal brain regions which are thought to reflect various attentional monitoring, selection, and switching processes.

Discussion: The results of this study help to elucidate the role of gamma and alpha frequency bands in frontal and modality-specific regions involved with selective attention in multi-task immersive situations.

KEYWORDS

inattentional deafness, EEG, gamma, alpha, natural cognition, neuroergonomics, attention

Introduction

When carrying out multiple tasks at the same time, it is often the case that performance on one or multiple tasks may degrade. This is especially true during scenarios with high task demands. Indeed, attentional mechanisms are implemented to selectively enhance relevant neural processes based on current behavioral goals. Attentional selection has been proposed to be necessary because of the limited processing capacity of the brain (Buschman and Kastner, 2015) or alternatively as a means of supporting potential action (Allport, 1987; Neumann, 1987; Edelman, 1989). For example, one may not hear someone speaking to them when they are immersed in some tasks that requires considerable attention (Cherry, 1953), such as operating a vehicle, using a smartphone, and/or playing video games. There have been several studies that have investigated the phenomenon of "inattentional deafness", which is the inability to consciously perceive and respond to audible sounds resulting from attention being directed elsewhere (Macdonald and Lavie, 2011; Dalton and Fraenkel, 2012; Koreimann et al., 2014; Raveh and Lavie, 2015; Kreitz et al., 2016; Scheer et al., 2018). The extent to which "inattentional deafness" occurs is thought to be dependent on processes related to selective and divided attention (Lavie, 2005; Callan et al., 2018; Dehais et al., 2019a,b). The goal of this study was to elucidate brain-localized neural correlates that predict successful and unsuccessful auditory perception during a dual-task situation both before and after presentation of the auditory stimulus. Our objective was also to further identify brain-localized neural correlates that predict why some individuals perform better than others.

Cortical oscillations especially, in the gamma (>30 Hz) and alpha frequency range (8 to 14 Hz), are thought to play a large role in the underlying brain processes mediating attention (Clayton et al., 2015). It has been proposed that gamma-band activity promotes task-relevant activity in modality relevant perceptual processing regions (Clayton et al., 2015). Greater gamma-band activity means greater facilitation for attended stimuli (Golumbic et al., 2013), whereas alpha-band activity is proposed to be involved with inhibition of task-irrelevant processes (Clayton et al., 2015). Greater alpha activity usually occurs in the non-dominant task modality brain processing regions and is thought to be suppressive in nature. There have been several EEG and MEG studies that have identified source-localized activity that is related to attention and performance (Fries et al., 2001; Ergenoglu et al., 2004; Hanslmayr et al., 2007; Womelsdorf and Fries, 2007; van Dijk et al., 2008; Tallon-Baudry, 2009; Wyart and Tallon-Baudry, 2009; Rieder et al., 2011; Clayton et al., 2015; Yuan et al., 2016; Wittenberg et al., 2018; Zhou et al., 2021). Most of these studies involved visual rather than auditory processing tasks. Greater post-stimulus gamma (and to a lesser extent pre-stimulus gamma) is associated with facilitation of modality-specific attention and performance (see Rieder et al., 2011 for review; Wyart and Tallon-Baudry, 2009; Yuan et al., 2016), whereas greater pre- and post-stimulus alpha activity in modality-specific brain regions is associated with degraded attention and performance (van Dijk et al., 2008; Clayton et al., 2015). Conversely, reduction in pre- and post-stimulus alpha activity in modality-specific brain regions is associated with facilitation of attention and performance (Ergenoglu et al., 2004; Hanslmayr et al., 2007; van Dijk et al., 2008; Zhou et al., 2021).

There have been a number of studies that have investigated neural correlates of inattentional deafness (Giraudet et al., 2015; Molloy et al., 2015; Durantin et al., 2017; Callan et al., 2018; Scheer et al., 2018; Dehais et al., 2019a,b; Schlossmacher et al., 2021; Somon et al., 2022). Brain imaging studies using fMRI have identified various brain regions and networks involved with attention under dual- and multi-task situations (Dux et al., 2006; Tombu et al., 2011; Szameitat et al., 2016). One region that has been cited as being active during situations of attentional overload, often referred to as attentional bottleneck, is the superior medial frontal cortex including pre-supplementary motor area (pre-SMA). This region has been implicated with processes related to inattentional deafness (Durantin et al., 2017). The results of this fMRI study indicate that the right inferior frontal gyrus (IFG) becomes active during episodes of inattentional deafness and further reveal the presence of suppressive connectivity between this region and auditory processing regions in the right superior temporal gyrus (Durantin et al., 2017). There have also been a considerable number of electroencephalography EEG and magnetoencephalography MEG studies that have found significant differences in various event-related potentials associated with inattentional deafness (Giraudet et al., 2015; Molloy et al., 2015; Dehais et al., 2019a,b; Schlossmacher et al., 2021). Molloy et al. (2015) conducted a MEG study in which participants performed a visual search task as the dominant task. The study revealed that auditory event-related responses, which were localized in the auditory processing brain regions (including the superior temporal sulcus and posterior middle temporal gyrus), were significantly greater when the concurrent visual task had a low workload, as compared to when it had a high workload. In addition, the P300 potential, thought to reflect conscious awareness, was only present for the auditory event-related responses during the low visual workload search task. The visual event-related potentials, localized to visual processing brain regions, were shown to be differentially greater when the visual task was of high workload compared to low workload (Molloy et al., 2015). These results are consistent with facilitation of brain regions involved with the task dominant modality and inhibition of brain regions involved with the task non-dominant modality. As of yet, there have been no studies that have reported source-localized brain activity related to pre- and post-stimulus gamma- and/or alpha-band activity related to inattentional deafness.

For the related phenomena of inattentional blindness, there have also been a considerable number of experiments investigating underlying brain activity (see Hutchinson, 2019 for review). Similar to inattentional deafness, inattentional blindness is the inability to consciously perceive a fully visible object as a result of attention being directed to another task, event, or object (Mack and Rock, 1998; Simmons and Chabris, 1999). EEG and MEG studies have identified inattentional blindness-related differences in event-related potentials localized to visual processing brain regions as well as those involved with attention processing in parietal and frontal brain regions (Schubo et al., 2001; Ruz et al., 2005; Guzzon and Casco, 2011; Pitts et al., 2012; Schelonka et al., 2017; Hutchinson, 2019). Studies (Harris et al., 2020; Hutchinson et al., 2021) have shown that pre- and post-stimulus alpha activity in

parietal-occipital (visual processing) regions predicts inattentional blindness. Additional support that alpha activity may be related to inattentional blindness comes from a study showing that alpha-band transcranial alternating current stimulation tACS over occipital regions induces inattentional blindness (Hutchinson et al., 2020). As was pointed out in the review by Hutchinson (2019) and investigated by Pitts et al. (2014), there have been no studies implicating gamma-band activity for inattentional blindness. This is surprising given theories implicating gamma-band activity in task-specific facilitation (Clayton et al., 2015), as well as studies implicating it in attention and performance (see Wyart and Tallon-Baudry, 2009; Rieder et al., 2011 for review; Yuan et al., 2016).

It is unclear why gamma-band activity has not been found in previous studies in relation to inattentional blindness/deafness. One possibility is that this lack of a finding may be related to the type of tasks and stimuli used in these experiments. For the most part, the stimuli and tasks used to investigate brain activity underlying inattentional blindness/deafness are often oversimplified and not representative of real-world conditions (see Hutchinson, 2019 for review). It may be the case that these artificial tasks, employed in many of these experiments, do not engage the attention system that evolved to act in more natural situations. Indeed, if one considers the hypothesis (Allport, 1987; Edelman, 1989) that attention evolved to selectively enhance neural processes based on behavioral goals directed toward specific actions—whether executed explicitly or implicitly implied—then the "naturalness" of the task appears to play a critical role in engaging the attention system. This is because it is the choice of value-dependent action that determines which modalities to enhance and which modalities to inhibit. The appearance of "limited capacity" is a natural consequence of such a selectional value-dependent-based attentional system (Neumann, 1987; Edelman, 1989). It has further been stated that in order to understand the distributed brain processes underlying natural human behavior ("natural cognition"), it is important to investigate within the context of real-world situations (Makeig et al., 2009; Gramann et al., 2014, 2021). This is a key goal of neuroergonomics (Parasuraman and Rizzo, 2008; Dehais et al., 2020). Indeed, a relevant example of this neuroergonomic approach comes from a study (Callan et al., 2018) that involved an auditory detection task in pilots in flight. This experiment, which was conducted in a real-world environment, induced a high rate of inattentional deafness. It was reported that disruption in neural phase synchrony in theta and alpha frequency bands is associated with performance decrements resulting from inattentional deafness (Callan et al., 2018). It may be the case that running experiments under natural real-world like conditions is the only way in which one can investigate the underlying neural processes that truly engage natural human attention.

This study seeks to determine brain-localized neural oscillations in gamma and alpha frequency ranges that predict auditory perceptual performance using EEG on a more engaging "natural" dual-task paradigm to induce inattentional deafness. The primary task in the experiment was to play the Nintendo Wii Skateboard Arena game that uses the Wii Balance Board for control. The virtual skateboarding task was selected for this experiment because it represents a real-world situation that is engaging and likely to induce considerable inattentional deafness. The secondary task involved an auditory stimulus difference detection task (which was in effect a 1-back task). In this task, participants had to press a button when one of two stimuli was different from the one previously presented. This task was selected as the non-dominant task to investigate inattentional deafness, rather than a simple audio detection task, because it requires utilization of echoic sensory maintenance processes thought to require greater attentional processing demands. It is predicted that pre- and post-stimulus gamma and alpha-band activity related to auditory task performance will be in accordance with theories of the role of brain oscillations for attention (Clayton et al., 2015).

The experimental tasks and predictions are outlined as follows: The Wii game is composed of six levels in which various skateboarding skills are performed. The goal for each level is to complete the tasks given in a continuous manner in the briefest amount of time. Between each level, there is a transition period of relative non-activity in which the participant pushes a button to continue on to the next level. The auditory task occurs continuously throughout the experiment. The goal of the task is to push a button every time there is a stimulus change from the previous one (auditory stimuli are presented every 2–3 s). The primary hypotheses consist of the following: It is predicted that during the skateboarding task, performance on the auditory task will be degraded as a result of the high workload of the dual-task demands imposed, initiating the phenomena of inattentional deafness. Based on theories of brain oscillations for attention (Clayton et al., 2015) outlined above, it is predicted that gamma brain activity will be greater for auditory misses than hits in brain regions known to be involved with inattentional deafness (Durantin et al., 2017) including the superior medial frontal cortex and the inferior frontal gyrus. It is further predicted that gamma-band activity will be degraded/enhanced in auditory cortical processing regions with respect to misses and hits reflecting periods of inattentional deafness and successful dual-task attentional processing. In contrast, alpha activity, that is considered to be suppressive in nature (Clayton et al., 2015), is predicted to be greater in auditory processing regions for misses over that of hits. Conversely, alpha activity is predicted to be reduced in visual and motor processing regions involved with the skateboarding task. Based on the attention literature reviewed above, and on the continuous nature of both the skateboarding and auditory tasks, we maintain that these predictions will hold for activity both pre- and post-auditory stimulus onset. To explicitly test changes in oscillatory brain activity that arise from induced and or evoked properties of the auditory stimulus that are different from ongoing processes that are present before stimulus onset, an event-related spectral perturbation ERSP analysis was conducted. This analysis takes into account baseline activity prior to stimulus onset on a single trial basis. This can have a profound effect on the pattern of brain activity from analysis of post-stimulus activity without baseline removal (Basar, 1998; Makeig et al., 2004). For example, differences in brain activity that are present both pre- and post-stimulus presentation between auditory hits and misses will likely be removed if baseline correction is applied. It is predicted that brain regions involved with attentional salience (ventral attention network), including the inferior frontal gyrus known to be involved

with attentional switching (Doeller et al., 2003; Perianez et al., 2004; Tamber-Rosenau et al., 2018), will show greater ERSP for hits over that of misses. Additionally, it is predicted that primary and especially secondary auditory processing regions, involved with complex processing of the acoustic features of the stimulus, will show greater ERSP for hits over misses. This pattern of differential gamma- and alpha-band activity pre- and post-stimulus onset, as well as ERSP related activity (given above), is predicted to be a signature of auditory task performance across participants. If there is a tradeoff between performing well in auditory tasks and skateboarding tasks, we can predict a negative correlation between brain activity related to auditory hits and misses. This is especially expected for the ERSP analysis, which reflects activity underlying processes related to attentional saliency and switching.

Above, we mentioned that one method used to evaluate inattentional deafness/blindness involves manipulating workload. This approach has been employed in this study to assess the phenomenon. To confirm that the observed differences in oscillatory brain activity between hits and misses are indeed related to inattentional deafness and not merely differences in perceptual performance, we compared the experimental contrasts of auditory hits relative to misses for the dual task during Wii skateboarding (high dual-task workload) with the transition period between levels of the skateboarding task (low dual-task workload). It is predicted that the same pattern of differential activity as discussed above that occur during the skateboarding task (high dual-task workload) will be maintained when contrasted with activity between hits and misses during the transition period between levels (low dual-task workload).

FIGURE 1
Experimental setup for dual-task virtual skateboard task and auditory difference detection task. The participant is wearing the CGX Quick-32 dry-wireless EEG and Bluetooth insert earphones. The Wii Balance Board is used to control the virtual skateboard game. Video for the skateboard game is presented on a 55-inch LCD about 2 m from the participant. Responses on the auditory task are made by pressing the B button on the Wii mote held in the right hand.

Materials and methods

Participants

This study included 14 participants (six female participants) aged 20–52 years (mean = 23.8, SE = 2.27). A modified version of the Edinburgh handedness questionnaire (that also included questions related to "hand in which chopsticks are used" and "what foot you are better at kicking with") revealed that 13 of the participants were completely right-handed, whereas one participant used both right and left hands depending on the task. All participants reported normal hearing and normal or corrected to normal vision. All but one of the participants had prior experience playing Wii Balance Board games. Participants that did not have experience with the Wii Fit Plus Skateboard Arena game received training on a separate day from the experiment for approximately 1 to 2 h until they could consistently reach the beginner level 6. Participants that could not perform the Wii skateboard game up to level 6 were excluded from the later experiment. There were no other exclusion criteria. Originally, there were 19 participants that were recruited for the EEG experiment but five participants were excluded for reasons including the following: extremely noisy EEG data (three participants) and machine errors related to data acquisition (two participants). The experimental procedures were approved by the ATR Human Subject Review Committee (ethics approval number 158) and were carried out in accordance with the principles expressed in the WMA Declaration of Helsinki. The confidentiality rights of all participants were observed.

Experimental tasks and procedures

This experiment consisted of two concurrent tasks that included the following: (1) performing the Wii Skateboard Arena game; (2) performing an auditory difference detection task (see Figure 1 for a picture of the setup of the experiment with a participant doing the dual tasks). The participants were instructed to do the best they could on the Wii Skateboard Arena game by focusing their attention to it while at the same time trying to also carry out the secondary auditory difference detection task. The Wii Skateboard Arena game uses the Balance Board to detect changes in forces at the four corners of the board caused by body movement. The Wii Balance Board was placed approximately 2 m from the 55-inch LCD display upon which the video of the game was presented. No background audio sound from the game was presented. The game consists of six levels. Levels 1 to 5 focus on specific tasks [(1) maneuvering over various targets on the ground,

(2) maneuvering over ramps and doing tricks, (3) doing tricks on a half-pipe, (4) jumping over and grinding on rails, and (5) jumping onto and grinding on a platform]. Level 6 is a combination of all the tasks from levels 1 to 5 with the addition of cones that must be avoided. Between each of the levels is a brief transition period in which the participant is required to press a button to proceed to the next level. Maneuvering and performing tricks in the game on the Wii Balance Board is accomplished by shifting one's weight, differentially standing on toes/heals, and by knee extensions. The participants were instructed to start the next level by pressing the A button on the hand-held Wii mote controller. At the end of each game, when level 6 was finished, a final score is displayed to the participant. The participants are instructed to restart the game after level 6 until the audio task is finished. The Wii game should be continued to be played when the audio task finishes until the end of level 6. Then, the experiment is over. The button-press responses and the Wii Balance Board force sensor data were captured by lab streaming layer (LSL, UCSD, SCCN) for synchronization with other data streams.

The auditory difference detection task required participants to push the B button on the bottom of the hand-held Wii mote controller when the current audio stimuli being played was different from the previous one played. There were two audio stimuli consisting of an upward chirp from 2 to 4 kHz (100 msec in duration) and a downward chirp from 4 to 2 kHz (100 msec in duration). A total of 400 of each type of stimuli were presented in random order with an interstimulus interval randomly determined between 2 and 3 sec. The auditory difference detection task experiment was approximately 33 min long. MATLAB was used to present the audio stimuli and send triggers to LSL identifying which stimuli are presented. The low latency EPOS GTW 270 Bluetooth earbuds were used together with a low latency aptX Bluetooth transmitter to present the audio stimuli to the participants. Two cables were routed from the audio out of the computer. One cable was to the low latency aptX Bluetooth transmitter, and the other was to the Cognionics Trigger Box. The threshold on the trigger box was set to identify the onset of the audio stimuli. The audio level was set manually for each participant prior to the experiment to be as loud as they thought they could comfortably tolerate for the 30- to 40-min experiment. The audio stimuli were played continuously throughout the experiment and therefore were played during the various Wii Skateboard Arena levels as well as between the levels. The analyses investigated event-related spectral power differences between hits relative to misses for auditory change events approximately 1 s pre-stimulus onset (−1000 to −50 msec), one half sec post-stimulus onset (0 to 500 msec), as well as an analysis in which the average spectral power from −250 to −50 msec pre-stimulus onset was used as a baseline for post-stimulus activity from 0 to 500 msec on a single trial basis [event-related spectral perturbation analyses ERSP (Makeig, 1993)]. The time segment ranges used for pre- and post-stimulus conditions were selected for the following reasons: The relatively long pre-stimulus range of approximately 1 s and the post-stimulus range of 0.5 s were selected to attempt to extract sustained attentional processes involved with suppression and enhancement related to dual-task processing thought to underlying inattentional deafness. Both the primary skateboarding task and the secondary auditory task were continuous in nature. The auditory task requires attentional maintenance of the previous stimulus in echoic memory in relation to the next stimuli that is presented every 2 to 3 s. Longer time segment ranges for pre- and post-periods were not selected to avoid potential overlap of pre- and post-brain activity. The beginning of the pre-stimulus range ensured at least 1 to 2 s of time from the previous stimulus onset. The end of the post-stimulus onset range was at least 0.5 to 1.5 s prior to the beginning of the pre-stimulus range for the next trial. The same post-stimulus time range was used for the ERSP analysis. The time period between −250 and −50 msec prior to stimulus onset was selected as a reasonable amount of time to account for baseline activity for the ERSP analysis. The reason why analyses were not carried out across multiple time points both pre- and post-stimulus onset to achieve better temporal resolution was because of the added cost of the need to correct for multiple comparisons for statistical significance.

Physiological recording and analysis

The EEG data were measured using the Cognionics CGX Quick-32r dry-wireless EEG system (Cognionics, Inc., San Diego). The sampling rate was 500 Hz with 24-bit analog-to-digital conversion. We used 29 electroencephalography EEG channels on the headset, with ground and reference electrodes located on the temporal bone behind the left and right ear respectively. The system utilizes active electrodes to minimize external noise pickup and artifacts. In addition to EEG data, the CGX Quick-32r headset also has accelerometer data in three axes. The EEG and accelerometer data were acquired wirelessly and streamed into LSL for recording and synchronization.

The EEG data were processed using the EEGLAB toolbox (Delorme and Makeig, 2004) using a similar pipeline as given in Bigdely-Shamlo et al. (2015), Callan et al. (2018), and Sasaki et al. (2019). These preprocessing steps are used to improve signal quality and remove artifacts to be able to extract brain activity.

- The raw EEG data were band-pass filtered from 3 to 100 Hz using a Hamming windowed Sinc FIR filter.
- To extract considerable head movement artifacts from the EEG, as a result of playing the Wii Skateboard Arena game, the three accelerometer axis data in the CGX Quick-32r headset were regressed out of the EEG data (with 0 delay setting) using the CWRegrTool in EEGLAB.
- Line noise (60 Hz) was removed using the Cleanline EEGLAB toolbox (default settings).
- Automatic channel rejection was based on poor correlation to robust estimate based on other channels (0.8).
- The rejected channels were interpolated.
- Common average referencing of channels was conducted after interpolation of missing channels (An additional channel with all zeros was added so as to not lose 1 rank as a result of average referencing).
- Artifact subspace reconstruction (ASR) (see Mullen et al., 2013 and Chang et al., 2018) (Euclidian distance) was used to remove non-stationary high-variance signals from the

EEG (standard deviation cutoff for removal of bursts = 20; windowed criterion = 0.25). Two analyses were conducted: one in which the time windows were removed for which ASR did not repair completely; and another in which no time windows were removed.

- Common average referencing of channels was conducted on the two datasets (time windows removed and time windows not removed) (An additional channel with all zeros was added so as to not lose 1 rank as a result of average referencing).
- ICA using PCA reduction was used on the dataset in which the time windows were removed after ASR cleaning. The number of rejected channels determined the rank reduction by PCA.
- The weights of the ICA were then applied to the ASR results without the time windows removed.
- Dipole fitting of the source for each independent component (IC) using Dipfit was conducted.
- ICLabel (version 1.1) was used to identify independent components (ICs) that are brain-related and artifact-related. ICLabel is a toolbox that allows for automated classification of ICs into seven different categories: Brain, Muscle, Eye, Heart, Line Noise, Channel Noise, and Other (Pion-Tonachini et al., 2019) at expert level of performance. ICLabel uses IC topomaps, power spectral density from 3 to 100 Hz, and equivalent current dipole to categorize each IC (Pion-Tonachini et al., 2019). The criteria of selecting "Brain" ICs in our study were based on the percentage of "brain" categorization over 50%.
- Brain-related ICs were retained, and all other ICs were removed from the dataset (only brain-related ICs were projected to the EEG electrodes).
- The Bluetooth audio presentation system latency of approximately 45 msec was corrected. The audio events were extracted from −2 s before until 2 s after presentation of the audio stimulus based on the trigger from the audio output cable.
- Source localization of the brain-related activity was carried out using LORETA Key software (Fuchs et al., 2002; Pascual-Marqui, 2002; Jurcak et al., 2007). See below for details.

Source localization and statistical analysis

LORETA Key software employing sLORETA (Fuchs et al., 2002; Pascual-Marqui, 2002; Jurcak et al., 2007) was used to determine source localization of brain-related activity on the surface of the cortex. The position of the electrodes of the CGX Quick-Cap 32r on the head was determined by reference to the 10–10 system within the LORETA Key software. Using the LORETA Key software, the event EEG data (−2 to 2 sec) were transformed from EEG to time–frequency cross-spectrum in the gamma frequency band (30.5 to 50 Hz) and the alpha frequency band (8.5 to 12 Hz), with a window width of 250 samples (continuous Gaussian window), and a delta T between running windows equal to 25 (50msec). The time-varying cross-spectra files were then converted to sLORETA files that localizes the activity across 6,239 voxels covering the cortex with 5-mm resolution.

For each participant, the sLORETA-converted individual trials were submitted to statistical analyses using a two-sample unequal variances *t*-test for the following contrasts: Auditory change hits vs. misses for stimuli presented during levels 1 to 6 as well as during the transition period between levels. The resultant sLORETA files for these contrasts for each participant were used in random effects group level analyses using SnPM permutation analysis (5,000 randomizations) within the LORETA Key software. The random effects level analyses for data on levels 1 to 6 included single group analysis that the mean was not equal to zero as well as single group regression with the variable hit rate of levels 1 to 6 for the contrast of hits minus misses and the variable maximum Wii score across all games played during the experiment. Hit rate was used as the correlation metric rather than D-prime because the false alarm rate was low and only hit and miss trials were used in the analyses of brain activity. The maximum Wii score was selected rather than other potential metrics such as mean score as it is commonly what is used to assess who is better in video game performance. To further assess the extent to which the results can be concluded to be both a product of performance and workload (thought to underlying processes related to inattentional deafness), random effects paired analyses (using SnPM permutation analysis (5,000 randomizations) within the LORETA Key software) were carried out for these same contrasts for skateboarding levels 1 to 6 (high dual-task workload) compared to the transition period between levels (low workload). Corrected critical thresholds for multiple comparisons were used in accordance with SnPM (Nichols and Holmes, 2001). Because of our theoretical interest in the roll that the auditory cortex plays in relation to successful and impaired perception as a result of inattentional deafness, region of interest (ROI) analyses were carried out for all contrasts in both left (MNI−55,−25, 10) and right (MNI 55,−25, 10) auditory cortex based on the centroid coordinate of Brodmann area 41 and 42. Only the significant results using SnPM will be reported in the results section.

Results

Behavioral results

This experiment consisted of two tasks: (1) the Wii Skateboard Arena game task; (2) the auditory difference detection task. The two tasks were done concurrently. Because a great deal of attention had to be given to the Wii Skateboard Arena task, there was low performance on the auditory difference detection task even though the stimuli were played at a high audio level and the room was quiet with no background sound from the game being presented. The mean total number of change trials across participants was 401.1 (SE = 4.2): mean hit rate = 0.30 (SE = 0.046), mean false alarm rate = 0.04 (SE = 0.009), mean d' = 1.06 (SE = 0.11). The breakdown of the mean and SE hit rates (HRs) and false alarm rates (FARs) by Wii game level segment is as follows (see Figure 2): Level 1: HR = 0.252 (SE = 0.048) FAR = 0.070 (SE = 0.020), Level 2: HR = 0.263 (SE = 0.051) FAR = 0.038 (SE = 0.010), Level 3: HR = 0.182 (SE = 0.048) FAR = 0.047 (SE = 0.016), Level 4: HR = 0.207 (SE = 0.057) FAR = 0.037 (SE = 0.011), Level 5: HR = 0.119 (SE = 0.038) FAR = 0.035 (SE = 0.012), Level 6: HR = 0.162 (SE = 0.042) FAR = 0.037 (SE = 0.011), transition period between levels: HR =

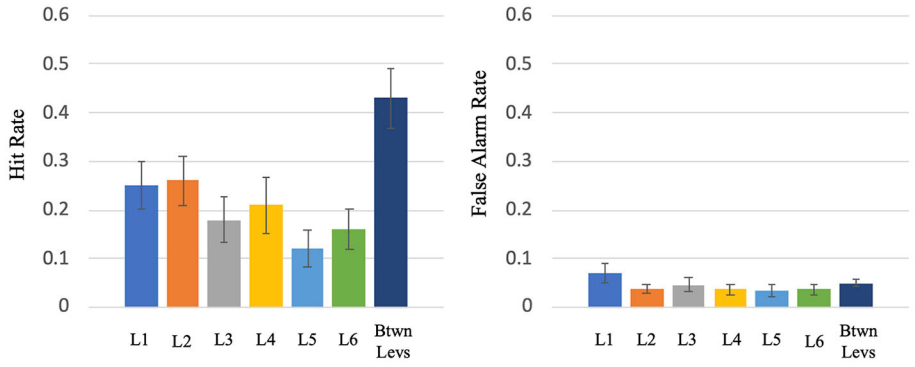


FIGURE 2
Hit rate and false alarm rate for the various levels (L1 to L6) and the transition period between levels of the virtual skateboard game.

0.431 (SE = 0.061) FAR = 0.050 (SE = 0.008). For levels 1 to 6, the mean HR = 0.189 (SE = 0.044) and FAR = 0.040 (SE = 0.010). A repeated measures ANOVA test on levels 1 to 6 indicated that there is a significant difference in auditory task hit rate between the different levels, F (5,65) = 8.5, $p < 0.001$. A paired t-test indicated a significant difference in the mean performance of Wii levels 1 to 6 (mean HR = 0.189) compared to that of the transition period between Wii levels (mean HR = 0.431) ($T = -7.387$, $p < 0.0001$, df = 13).

There was no significant overall linear trend in Wii skateboarding game performance over the course of the experiment [Using Wilcoxon signed rank test ($p > 0.05$) that the median r score was greater than zero]. The correlation coefficient across participants ranged from $r = 0.67$ to $r = -0.45$ with the median being $r = 0.11$. The total number of games varied across participants depending on how fast they could get through the games in the same amount of time. The number of games across participants ranged from 8 to 12 with 9 being the median number of games. There were no participants that showed a significant correlation between Wii game performance and game number. Furthermore, there was no significant overall correlation between the auditory task performance (measured by hit rate) with that of Wii game performance over the course of the separate games composing the experiment [Wilcoxon signed rank test ($p > 0.05$)]. The correlation coefficient across participants ranged from $r = 0.56$ to $r = -0.81$ with the median being $r = -0.10$. There was only one participant that showed a significant correlation ($p < 0.05$, r = −0.81) between auditory task performance and Wii performance across the course of the separate games composing the experiment.

EEG sLORETA results

Gamma band

Random Effects level analyses were conducted for the contrast of auditory task hits vs. misses for stimuli presented during levels 1 to 6 of the Wii Skateboard Arena game. For the random effects analysis, that sLORETA gamma-band activity was not equal to zero for the contrast of hits relative to misses of average pre-stimulus power differences within the time segment from −1,000 to −50 msec, there was a significant decrease ($p < 0.05$ two-tail corrected for multiple comparisons; SnPM using 5,000 randomizations, T threshold = 3.39) in gamma-band spectral power in the superior and middle frontal gyrus (Brodmann area BA 8) extending into medial frontal gyrus (BA9) and pre-SMA (BA6), as well as the cingulate gyrus (BA32) (see Figure 3A and Table 1). Significant activity was present in the pre-SMA determined by the atlas defined by Sallet et al. (2013). No significant power differences ($p > 0.05$ one-tail corrected for multiple comparisons; SnPM using 5,000 iterations) for the contrast of hits relative to misses were present for the average of the time segment from 0 to 500 msec post-stimulus onset.

Random effects single group regression analyses were used to determine how participant level performance on the auditory task related to corresponding differences in brain activity. For the sLORETA contrast of hits relative to misses on the auditory task, using each individuals hit rate as the regression variable, statistically significant positive correlation ($p < 0.05$ two-tail corrected for multiple comparisons; SnPM using 5,000 iterations; R threshold = 0.742) was present for averaged pre- stimulus activity (−1,000 to −50 msec) in left inferior parietal lobule (BA40), supramarginal gyrus (BA40), angular gyrus (BA39), and auditory processing areas in the left temporal lobe including auditory cortex (BA41, 42), superior temporal gyrus (BA22), and middle temporal gyrus (BA21) (see Figure 3B and Table 1). For averaged post-stimulus activity from 0 to 500 msec, a significant positive correlation between brain activity and hit rate ($p < 0.05$ two-tail corrected for multiple comparisons; SnPM using 5,000 iterations; R threshold = 0.740) was present in left auditory processing areas including the primary auditory cortex (BA42), the superior temporal gyrus STG (BA22), and the middle temporal gyrus MTG (BA21) (see Figure 3C and Table 1).

It should be noted that there was no statistically significant correlation ($p > 0.05$ uncorrected) between brain activity (sLORETA) for hits alone and individual participant hit rate on the auditory task nor between brain activity for misses alone and individual participant hit rate. The lack of a significant correlation of brain activity, for the unitary variables of hits or misses alone,

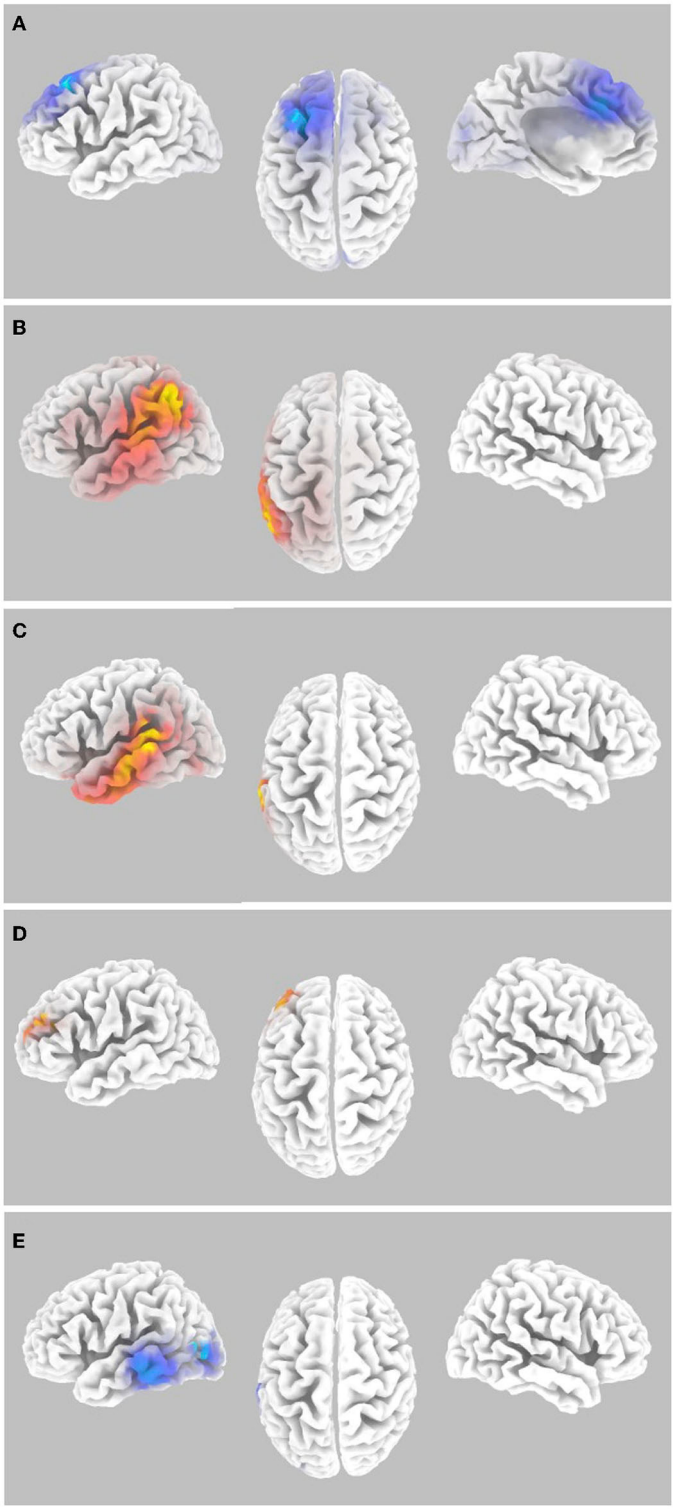


FIGURE 3
Gamma-band differences for the high workload dual-task condition occurring during the Wii skateboarding task. **(A)** Pre-stimulus onset: auditory hits relative to misses. Threshold for multiple comparisons ($p < 0.05$) rendered on the brain is $T = -3.00$ one-tail (darker blue). **(B)** Pre-stimulus onset: correlation of individual auditory hit rate with contrast of auditory hits relative to misses. Threshold for multiple comparisons ($p < 0.05$) rendered on the brain is $R = 0.674$ one-tail (red). **(C)** Post-stimulus onset: correlation of individual auditory hit rate with contrast of auditory hits relative to misses. Threshold for multiple comparisons ($p < 0.05$) rendered on the brain is $R = 0.682$ one-tail (red). **(D)** ERSP: auditory hits relative to misses. Threshold for multiple comparisons ($p < 0.05$) rendered on the brain is $T = 4.54$ 1-tail (yellow). **(E)** ERSP: correlation of individual skateboard game performance with contrast of auditory hits relative to misses. No activity was found to be significant using corrected thresholds. The threshold used for displaying the figure was set to $R = -0.57$ (blue). A region of interest (ROI) analysis in auditory cortex (MNI −55, −25, 10) did reveal significant negatively correlated activity $R = -0.655$ (ROI threshold $p < 0.05 = 0.537$, two-tailed).

TABLE 1 Gamma band, High Workload: Hits relative to Misses.

Contrast	Brain region	MNI coordinate x, y, z	RorT
Pre	SFG BA8	−30, 25, 55	−3.48
	MFG BA8	−20, 20, 45	−3.46
	CG BA32	−10, 20,40	−3.40
	SMFC Pre-SMA BA6,8,9	−10, 25, 40	−3.39
Pre RGR HR	IPL BA40	−55, −55, 45	0.818
	AG BA39	−55, −65, 35	0.808
	SMG BA40	−60, −55, −35	0.804
	SPL BA7	−40, −60, 50	0.766
	AC BA42	−65, −35, 20	0.802
	AC BA41	−55, −30, 10	0.790
	STG BA22	−60, −40, 20	0.796
	STG BA22	−65, −25, 0	0.777
	MTG BA21	−65, −15, −10	0.768
	ITG BA20	−65, −20, −20	0.755
	Insula BA13	−45, −35, 20	0.762
Post RGR HR	AudCor BA42,22	−65, −35, 20	0.755
	STG BA22	−61, −33, 7	0.732
	MTG BA21	−60, −19, −11	0.748
ERSP	SFG, DLPFC BA10,46	−35, 45, 25	4.59

BA, Brodmann area; SFG, superior frontal gyrus; MFG, middle frontal gyrus; CG, cingulate gyrus; SMFC, superior medial frontal cortex [pre-SMA as determined by the atlas defined by Sallet et al. (2013)]; IPL, inferior parietal lobule; AG, angular gyrus; SMG, supramarginal gyrus; SPL, superior parietal lobule; AC, auditory cortex; STG, superior temporal gyrus; MTG, middle temporal gyrus; ITG, inferior temporal gyrus; SFG, superior frontal gyrus; DLPFC, dorsolateral prefrontal cortex. BA46 was determined by atlas of Sallet et al. (2013).

with participant performance (hit rate) strongly suggests that the significant correlation in brain activity found for the contrast of hits minus misses (Figures 3B, C and Table 1) is actually related to the difference rather than any particular variable characteristic of one of the unitary variables of hits or misses alone (e.g., number of hit or miss trials). It should also be noted that randomly selecting trials such that the number of hits and misses were equal produced vary similar results as are shown in Figures 3B, C, suggesting that the results are not due to unequal numbers of trials between hits and misses.

An ERSP analysis that takes into account trial level baseline activity just before stimulus presentation (−250 to−50 msec) was conducted to determine changes in spectral power that are a result of the auditory event. The random effects single group analysis of the sLORETA contrast for hits relative to misses for the average time segment from 0 to 500 msec post-stimulus onset showed statistically significant ($p < 0.05$ one-tail corrected for multiple comparisons; SnPM using 5,000 iterations, T threshold = 4.54) increased gamma-band power in regions along the inferior frontal junction including the left superior frontal gyrus and middle frontal gyrus (BA10), as well as the left dorsolateral prefrontal cortex DLPFC (BA46), determined using the atlas of Sallet et al. (2013) (see Figure 3D and Table 1).

The random effects single group regression analyses between participant level performance on the auditory task and the differential ERSP of hits relative to misses in the gamma frequency range did not show any statistically significant differences ($p >$ 0.05 one-tail corrected for multiple comparisons; SnPM using 5,000 iterations). However, the regression analysis between participant level maximum performance on the Wii skateboard game and differential ERSP sLORETA power differences for hits relative to misses did show a significant negative correlation in left auditory processing regions using a region of interest analysis centered in the auditory cortex (BA41, 42) at MNI coordinate (−55, −25, 10) ($R =$ −0.65; $p < 0.05$ two-tail, SnPM using 5,000 iterations; R threshold = 0.537) (see Figure 3E).

To assess the contribution of both workload and auditory task performance as underlying factors of inattentional deafness for gamma-band spectral power, SnPM paired analyses of the contrasts reported above for auditory hits relative to misses during Wii skateboarding levels 1 to 6 (high workload) were compared with the same contrasts for the transition period between Wii skateboarding levels (low workload). Only the following contrasts were found to show significant differential activity. The paired group regression analyses for auditory performance for pre-stimulus brain activity showed significant differences for levels 1 to 6 relative to the period between levels in the following brain regions: precentral (BA4) and postcentral gyrus (BA3 and 5) in the most superior and medial regions corresponding to foot, leg, and trunk representation, as well as the superior temporal gyrus (BA22) (see Figure 4A and Table 2) ($p < 0.05$ one-tail corrected for multiple comparisons; SnPM using 5,000 iterations; R threshold = 0.712). The ROI analysis in left auditory cortex (MNI −55, −25, 10) was also significant ($R = 0.643$; $p < 0.05$ one-tail, SnPM using 5,000 iterations; R threshold = 0.462). For the post-stimulus contrast, no significant differential brain activity was found when correcting for multiple comparisons (see Figure 4B). However, the ROI analysis did reveal significant correlation in left auditory cortex (MNI −55, −25, 10) ($R = 0.467$; $p < 0.05$ one-tail, SnPM using 5,000 iterations; R threshold = 0.441). Both the pre- and post-stimulus analyses showed significant differential activity for the correlation with hit rate across participants in auditory cortical areas for high workload relative to low workload conditions.

Significant differential activity for the paired group analysis for levels 1 to 6 relative to the period between levels was also present for the ERSP contrast of hits relative to misses. The brain regions found to show significant differential activity consisted of the inferior frontal gyrus IFG (BA 45,44,47), middle frontal gyrus (BA10 and 11), as well as the left dorsolateral prefrontal cortex DLPFC (BA46) and the superior temporal gyrus (BA38) ($p < 0.05$ one-tail corrected for multiple comparisons; SnPM using 5,000 iterations, T threshold = 4.306) (Figure 4C and Table 2). Activity in regions of the IFG, MTF, and DLPFC overlapped with those found in the ERSP analysis of Wii levels 1 to 6 reported above (Figure 3D and Table 1). The paired group regression analyses for Wii performance for ERSP showed significant differences for levels 1 to 6 relative to the period between levels in the following brain regions: IFG (BA47 and 45) and STG (BA22) ($p < 0.05$ one-tail corrected for multiple comparisons; SnPM using 5,000 iterations; R

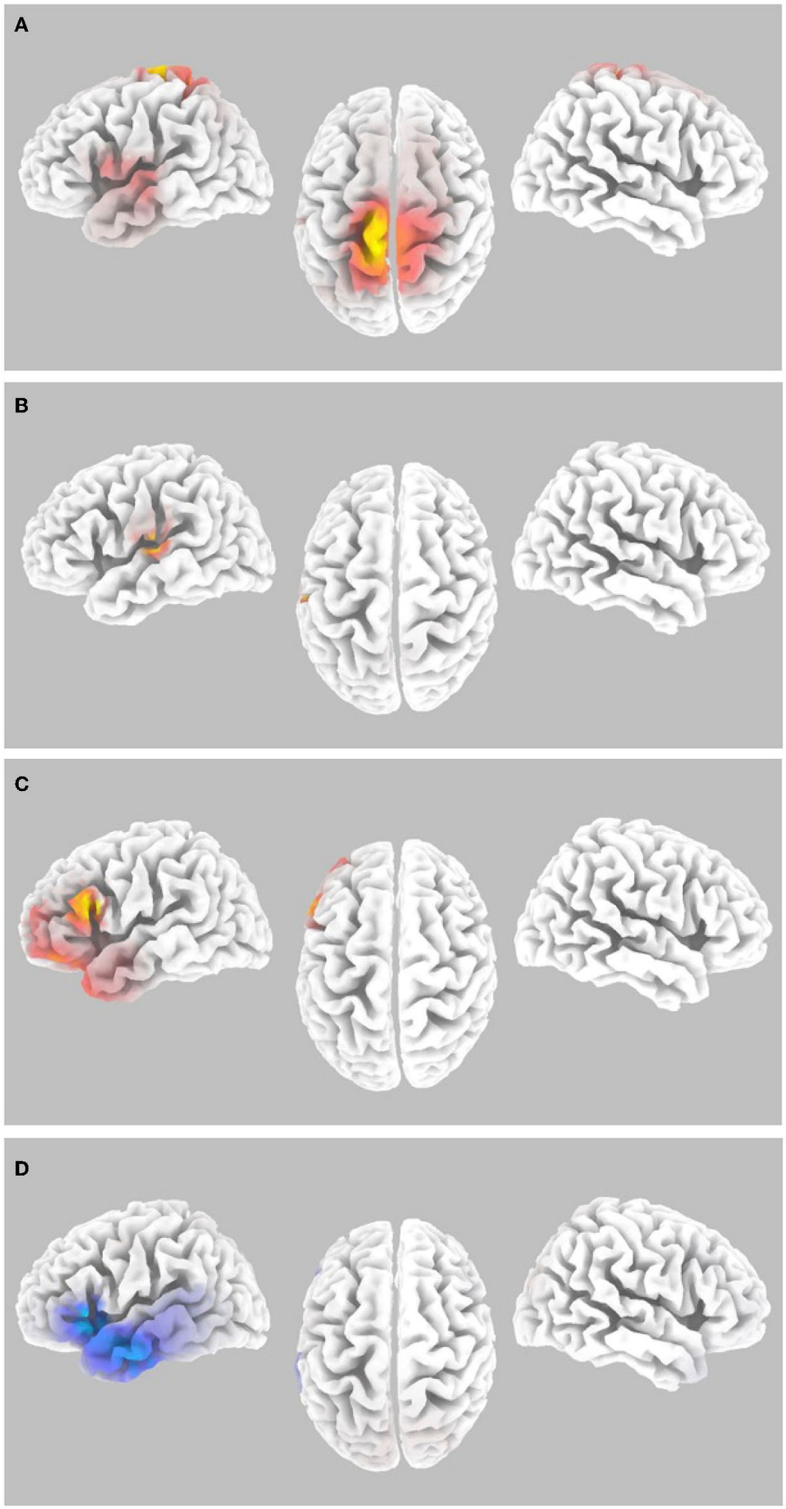


FIGURE 4
Gamma-band differences for the high workload dual-task condition occurring during the Wii skateboarding task relative to the low workload condition occurring in the transition period between Wii levels. **(A)** Pre-stimulus onset: correlation of individual auditory hit rate with contrast of auditory hits relative to misses. Threshold for multiple comparisons ($p < 0.05$) rendered on the brain is $R = 0.712$ 1-tail (red). **(B)** Post-stimulus onset: correlation of individual auditory hit rate with contrast of auditory hits relative to misses. No activity was found to be significant using corrected thresholds. The threshold used for displaying the figure was set to $R = 0.441$ (blue). A region of interest ROI analysis in auditory cortex (MNI−55,−25, 10) did reveal significant negatively correlated activity $R = -0.467$ (ROI threshold $p < 0.05 = 0.537$, one-tailed). **(C)** ERSP: auditory hits relative to misses. Threshold for multiple comparisons ($p < 0.05$) rendered on the brain is $T = 4.31$ one-tail (red). **(D)** ERSP: correlation of individual skateboard game performance with contrast of auditory hits relative to misses. Threshold for multiple comparisons ($p < 0.05$) rendered on the brain is $R = -0.799$ 1-tail (blue).

TABLE 2 Gamma band, High Workload relative to Low Workload: Hits relative to Misses.

Contrast	Brain region	MNI coordinate x, y, z	RorT
Pre RGR HR	PreCG BA4	−10, −35, 70	0.758
	PreCG BA4	10, −40, 70	0.719
	PostCG BA5	−5, −45, 70	0.744
	PostCG BA3	−20, −40, 65	0.725
	STG BA22	−65, −15, 5	0.715
ERSP	IFG BA45,44	−60, 15, 20	5.46
	IFG BA47	−45, 40, −15	5.11
	MFG BA11	−35, 35, −15	5.08
	SFG BA10, 46	−35, 55, 20	4.59
	STG BA38	−35, 20, −35	4.80
	MTG BA21	−45, 10, −40	4.55
ERSP RGR Wii	IFG BA47, 45	−50, 20, 0	−0.817
	STG BA22	−50, 15, −5	−0.811

BA, Brodmann area; PreCG, precentral gyrus; PostCG, postcentral gyrus; STG, superior temporal gyrus; IFG, inferior frontal gyrus; MFG, middle frontal gyrus; SFG, superior frontal gyrus; MTG, middle temporal gyrus.

threshold = −0.799). The ROI analysis in left auditory cortex (MNI −55, −25, 10) was also significant ($R = -0.513$; $p < 0.05$ one-tail, SnPM using 5,000 iterations; R threshold = −0.456). Activity in the ROI analysis of the left auditory cortex was also found in the ERSP analysis of Wii levels 1 to 6 reported above (Figure 3E).

Alpha band

The same Random Effects level analyses were conducted for alpha band as for the gamma band above. Only the analyses reported below were significant when correcting for multiple comparisons SnPM using 5,000 randomizations. All other analyses failed to reach significance when correcting for multiple comparisons using a one-tail threshold.

Within the alpha frequency band, the random effects analysis over the sLORETA data for auditory task hits relative to misses of average pre-stimulus power differences within the time segment from −1,000 to −50 msec showed a significant decrease ($p <$ 0.05 two-tail corrected for multiple comparisons; SnPM using 5,000 randomizations, T threshold = 2.981) in spectral power in predominantly the right insula spreading into the STG and the IFG (see Figure 5A and Table 3). The MFG and STG bilaterally also showed a significant decrease in alpha-band power (see Figure 5A and Table 3). No significant increases in alpha-band power were present when correcting for multiple comparisons (SnPM using 5,000 randomizations, one-tailed threshold).

For average post-stimulus power differences (0 to 500 msec) in the alpha frequency band between hits and misses, a significant decrease was present in the right auditory cortex, STG, and MTG ($p < 0.05$ two-tail corrected for multiple comparisons; SnPM using 5,000 randomizations, T threshold = 2.863) (see Figure 5B and Table 3).

Random effects single group regression analyses were used to determine how participant level performance on the auditory change detection task (using each individual hit rate as the regression variable) related to corresponding differences in brain activity for alpha-band sLORETA hits relative to misses. For average pre-stimulus activity (−1,000 to −50 msec), a statistically significant positive correlation ($p < 0.05$ two-tail corrected for multiple comparisons; SnPM using 5,000 iterations; R threshold = 0.696) was present in the IPL extending into the somatosensory cortex in the pre- and postcentral gyrus as well as the PMC (see Figure 5C and Table 3). No negative correlations were found to be significant when correcting for multiple comparisons (SnPM using 5,000 randomizations, one-tailed threshold).

Paired analyses (SnPM) of the contrasts reported above for auditory hits relative to misses during Wii skateboarding levels 1 to 6 (high workload) were compared with the same contrasts for the transition period between Wii skateboarding levels (low workload) to assess the contribution of both workload and auditory task performance as underlying factors of inattentional deafness for alpha-band spectral power. No significant differential activity was found for any of the contrasts or for any of the ROI analyses in the left and right auditory cortex.

Discussion

The aim of this study was to investigate whether EEG could reveal brain oscillations in the gamma and alpha frequency range, both pre- and post-stimulus presentation, that are indicative of inattentional deafness. The experiment's findings revealed that significant differences in brain activity between misses and hits (during the auditory stimulus difference detection task) occurred under the high workload skateboarding dual-task condition (presumably due to inattentional deafness) in some of the contrasts analyzed pre- and post-stimulus onset for the gamma and alpha frequency bands. Individual participant performance on the auditory stimulus difference detection task was also found to be correlated with differential brain activity for hits relative to misses on the auditory task. Furthermore, individual participant performance on the Wii skateboarding task was also found to be correlated with differential brain activity for hits relative to misses on the auditory task. Together, the results of this experiment investigating neural correlates of inattentional deafness are in accordance with the role of gamma- and alpha-band brain oscillations for attention and performance (Clayton et al., 2015).

To ensure that these results are a product of attentional processes thought to underlying inattentional deafness, a comparison of the relationship between auditory task hits and misses on the high workload dual-task condition occurring during levels 1 to 6 of the Wii skateboarding task was compared to a lower workload condition occurring during the transition period between Wii skateboarding levels. While there is some overlap in this differential workload-based analyses and the original high workload analyses in the gamma frequency range (Figures 3, 4 and Tables 1, 2), no significant activity was found for the differential workload-based analyses for the alpha frequency range. Even though this is the case, we still maintain that the results in the high workload dual-task condition are indeed a product of attentional

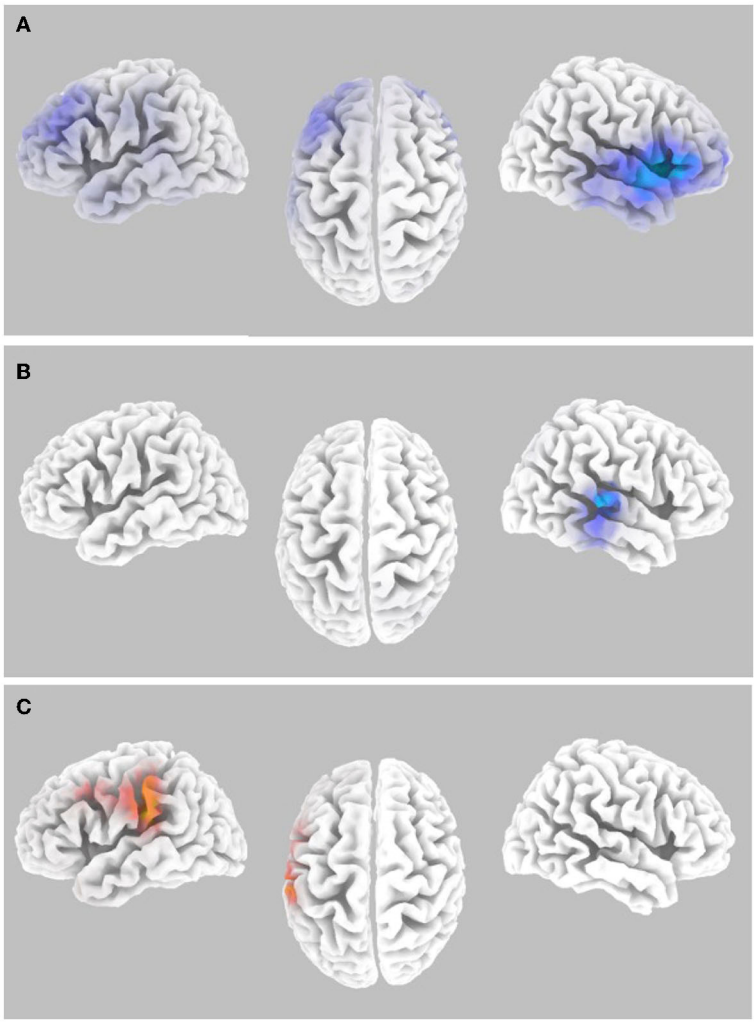


FIGURE 5
Alpha-band differences for the high workload dual-task condition occurring during the Wii skateboarding task. **(A)** Pre-stimulus onset: auditory hits relative to misses. Threshold for multiple comparisons ($p < 0.05$) rendered on the brain is $T = 2.98$ two-tail (blue). **(B)** Post-stimulus onset: auditory hits relative to misses. Threshold for multiple comparisons ($p < 0.05$) rendered on the brain is $T = 2.86$ 2-tail (blue). **(C)** Pre-stimulus onset: correlation of individual auditory hit rate with contrast of auditory hits relative to misses. Threshold for multiple comparisons ($p < 0.05$) rendered on the brain is $R = 0.696$ two-tail (red).

processes thought to underlying inattentional deafness. Indeed, this is especially true for pre-stimulus activity in which it is difficult to imagine what other process than that of attention could account for prediction of future perceptual performance. We discuss limitations of the low workload condition below and why the results of the high workload condition alone do not overlap more with the contrast of high workload relative to the low workload condition.

While previous EEG and MEG studies have not found neural correlates of inattentional deafness/blindness in the gamma frequency range (Pitts et al., 2014; Hutchinson, 2019) our study, which used a more ecologically valid task, revealed gamma-band activity consistent with the attentional and performance-related functions previously reported in the literature (see Rieder et al., 2011 and Clayton et al., 2015, for reviews). For the contrast of auditory misses > hits for gamma-band activity averaged from −1000 to −50 msec pre-stimulus onset, significant differential activity was present over a single cluster spreading across the SFG, MFG, CG, and the SMFC (including pre-SMA) (see Figure 3A and Table 1). The location of this differential brain activity is consistent with many previous fMRI studies that have identified activity related to conditions indicative of an attentional bottleneck under high workload conditions (attentional overload) (Dux et al., 2006; Tombu et al., 2011; Szameitat et al., 2016). Of particular relevance is the fMRI study conducted by Durantin et al. (2017), which investigated auditory alarm perception in the context of a flight simulation piloting task and identified activity in the same brain region associated with episodes of inattentional deafness, implicating its involvement in processes of selective and divided attention. It has been suggested that these regions in the prefrontal cortex are involved with monitoring and evaluation of performance according to current task goals such that the attention system can continuously excite task-relevant processes and inhibit task-irrelevant processes (Clayton et al., 2015). Greater activation of

TABLE 3 Alpha band, High Workload: Hits relative to Misses.

Contrast	Brain region	MNI coordinate x, y, z	RorT
Pre	Insula BA13	40, 10, 0	−4.15
	STG BA22	60, 0, −5	−3.99
	IFG BA47	55, 20, 5	−4.03
	MFG BA10	30, 55, −10	−3.53
	MFG BA9	−30, 25, 40	−3.56
Post	Aud Cortex BA42	65, −20, 10	−3.36
	STG BA22	65, −25, 0	−3.21
	MTG BA21	60, −30, −10	−3.02
Pre RGR HR	IPL/SMG BA40	−55, −30, 25	0.735
	PostCG BA3	−60, −25, 40	0.725
	PreCG BA4	−50, −15, 40	0.699
	PMC BA6	−50, −5, 35	0.703

BA, Brodmann area; STG, superior temporal gyrus; IFG, inferior frontal gyrus; MFG, middle frontal gyrus; MTG, middle temporal gyrus; IPL, inferior parietal lobule; SMG, supramarginal gyrus; PostCG, postcentral gyrus; PreCG, precentral gyrus; PMC, premotor cortex.

this region under high workload conditions may reflect a greater need to inhibit task-irrelevant processes. In our case, in which participants are performing virtual skateboarding as the primary task, this would be auditory processes related to the secondary task.

The significantly greater level of differential alpha activity for auditory misses over hits in right hemisphere auditory processing brain regions both pre- (Figure 5A and Table 3) and post-stimulus (Figure 5B and Table 3) onset is consistent with the role of alpha oscillations as inhibiting task-irrelevant processes (Clayton et al., 2015). In the case of this experiment, the auditory task is considered the secondary task to be inhibited when the action demands of the primary task of virtual skateboarding are high.

Given the established role of cortical oscillations in attentional processes (Clayton et al., 2015), it was hypothesized that successful dual-task performance on the auditory task, indicated by an auditory hit, would elicit greater gamma-band activity in the relevant auditory processing regions of the brain. However, contrary to this prediction there was no significant differential gamma-band activity for pre- or post-stimulus contrasts of auditory hits greater than misses for this study. An additional way to investigate performance-related gamma activity is by performing a correlation analysis between overall individual auditory task performance and gamma-band brain activity for the contrast of auditory task hits relative to misses. The results for these correlation analyses for both pre- (Figure 3B and Table 1) and post-stimulus (Figure 3C and Table 1) analyses show significant gamma-band activity in left hemisphere brain regions involved with auditory processing including the primary auditory cortex and superior temporal gyrus/sulcus. This was especially the case for post-stimulus onset gamma-band performance correlated activity in which the focal point was localized to the left primary auditory cortex (Figure 3C and Table 1). However, for the pre-stimulus onset gamma-band performance correlated activity, the focal point was localized to the left inferior parietal cortex, including the temporal parietal junction (Figure 3B and Table 1). This result is interesting in that the IPL and TPJ are part of the ventral attention network VAN that is thought to be involved with processing of saliency (for a review see Vossel et al., 2014; Dehais et al., 2019a,b). However, in this case, the activity in these VAN brain regions is prior to the onset of the stimulus upon which saliency occurs. Given that there is no stimulus present for which saliency of the VAN to respond to, it may be the case that the VAN is being primed to respond to the presentation of a future stimulus. This is consistent with a previous finding (Marois et al., 2004; Todd et al., 2005) reporting that a demanding task can suppress TPJ activity and in return prevent the processing of incoming stimuli. It could also be the case that pre-stimulus onset gamma-band performance correlated activity centered in the IPL may correspond to processes related to auditory maintenance utilizing verbal rehearsal (Henson et al., 2000).

A correlation analysis between overall individual auditory task performance and alpha-band brain activity for the contrast of auditory task hits relative to misses was also conducted. No significant correlation with performance was present pre-stimulus onset with brain activity. However, post-stimulus brain activity (for the contrast of auditory hits relative to misses) did show a significant positive correlation with individual auditory task performance centered in the SMG region of the IPL and extending into the post- and precentral gyrus as well as the premotor cortex (Figure 5C and Table 3). One of the functions that the IPL is thought to take part in is acting as the somatosensory association cortex involved with perception of limb and body location in space (Callan et al., 2012; Limanowski and Blankenburg, 2016; Ruotolo et al., 2019). In accordance with theories of the roles of neural oscillations (Clayton et al., 2015), the greater alpha activity in the IPL may be associated with greater inhibition of task-related processes that are involved with representation of body control important for the virtual skateboarding task. The results are consistent with the hypothesis that brain regions important for executing the virtual skateboarding task are inhibited to a greater extent for individuals that do better on the auditory task. It should be noted, however, that there was no negative correlation between individual auditory task performance and virtual skateboarding task performance.

To investigate potential ongoing changes in gamma- and alpha-band activity as a result of stimulus presentation, an event-related spectral perturbation (ERSP) analysis was conducted. This analysis takes into account baseline gamma (alpha) band activity prior to stimulus onset relative to gamma (alpha) band activity after stimulus onset. The results of the ERSP analyses revealed significantly greater gamma-band activity for auditory hits relative to misses in the SFG and DLPFC (Figure 3D and Table 1) (No significant ERSP was found for the alpha band). These brain regions (SFG and DLPFC) within the inferior frontal junction have been implicated in the integration of the dorsal attention network (DAN involved with executive processing) and the VAN (involved with saliency of sensory stimuli) during task switches and/or attention shifts (Tamber-Rosenau et al., 2018). Activity in this region, in our study, is consistent with successful dual-task performance in which integration and switching between the DAN and VAN in response to an auditory stimulus would enhance auditory task performance during the concurrent virtual skateboarding task.

Although the event-related spectral perturbation (ERSP) analysis did not reveal any significant differential gamma-band activity in the auditory processing regions for hits relative to misses (when correcting for multiple comparisons), a subsequent region of interest (ROI) analysis conducted in the left auditory cortex showed a significant negative correlation between participants' performance in the skateboarding game and the brain activity associated with auditory hits compared to misses (Figure 3E). The findings align with Clayton et al. (2015) proposition that oscillations play both excitatory and inhibitory roles. Specifically, individuals who exhibit higher levels of selective attention toward the skateboarding task demonstrate reduced facilitatory gamma-band enhancement in the auditory processing regions when performing the auditory task. It should be pointed out that although auditory brain activity is negatively correlated with skateboard performance, there is no behavioral correlation between skateboard performance and the auditory task. One potential explanation for the lack of a negative correlation in behavior between the two tasks could be that the auditory task was not sensitive enough to detect threshold differences resulting from reduction in gamma-band facilitation. In this experiment, the auditory stimuli were presented at a level that was deemed to be loud but comfortably tolerable for the 30- to 40-min duration of the experiment. Therefore, the stimuli were always at a level that was quite audible. Additionally, it is possible that for individuals that are really good at the virtual skateboard game, they may be utilizing more automatic processing typical of expert skill that is somewhat independent from executive attentional processes which may therefore allow for these individuals to do relatively better at both the skateboarding and the auditory tasks. Further research is needed to explore the hypothesis of the interaction of controlled and automated processes and the effects of attention on multi-task performance.

One interesting finding in our study was the apparent lateralization of gamma-band activity to the left auditory cortex and alpha-band activity to the right auditory cortex. One potential reason for the left hemisphere auditory processing region laterality for gamma-band activity is related to the facilitation of specific acoustic features for perception. According to Ivry and Robertson (1998), the left hemisphere is specialized for processing high frequency patterns, e.g., the transitions in speech. The 100 msec duration chirp sounds with a frequency sweep from 2 to 4 kHz is characteristic of transitions found in speech sounds. Alternatively, it is also possible that this activity may represent verbal short-term memory rehearsal that is known to have a left hemisphere dominance (Henson et al., 2000). While post-stimulus gamma-band activity was lateralized to the left hemisphere auditory processing areas (Figure 3C and Table 1), post-stimulus alpha-band activity was lateralized to right hemisphere auditory processing areas (Figure 5B and Table 3). Alpha-band activity is thought to be related to inhibition of task-irrelevant processes (Clayton et al., 2015). Therefore, if the primary task is skateboarding, one would expect greater inhibition of brain activity in brain regions not involved with this primary task which include general auditory detection ("awareness") in this case localized to the right hemisphere. Consistent with this hypothesis, in our study we see greater alpha-band activity for misses over hits in right hemisphere auditory brain regions signifying greater inhibition. Gamma reflects successful dual-task performance that varies across individuals and is facilitative (for the auditory task) in nature whereas alpha reflects the ability to suppress task-irrelevant processes (auditory) that may interfere with the focus of attention on the primary skateboarding task and is present across participants.

There are many potential reasons why we were able to find gamma-band activity in this study investigating inattentional deafness that were not present in other studies investigating inattentional deafness/blindness. (1) Most of basic MEG, fMRI, and EEG studies fail to induce high rate of inattentional deafness as they mainly report the effect of task workload on auditory processing (it is necessary to investigate misses as they are the potential occurrence of inattentional deafness/blindness). (2) Our study uses a challenging real-world immersive task that is likely to require greater engagement of the attention system, that in part utilizes gamma oscillations for facilitation of performance. (3) In many studies investigating inattentional deafness/blindness, there is only a single task upon which the participants are to act on and there is no task upon which the unexpected stimulus is acted on. In this case, it is primarily the saliency VAN that may be dominating the presence or absence of inattentional deafness/blindness. However, in our study, a dual-task paradigm was used; therefore, presence or absence of inattentional deafness relies on shifting of attention and successful divided attention in which both the VAN and the DAN are involved, thus utilizing greater gamma-band activity. (4) While most studies employing a dual secondary task utilize a very simple detection paradigm, we utilized an auditory task similar to that of a 1-back task in which the echoic memory of the previous stimulus needs to be maintained, thus engaging additional attention and related gamma-band activity. (5) Many previous studies simply did not do analyses to investigate source-localized gamma-band activity. (6) Our choice of the preprocessing pipeline to remove artifacts and extract brain activity and to conduct statistical analysis computed at the source level on the cortex may be more conducive to finding gamma-band activity differences between conditions. (7) Both of our tasks are continuous, thus requiring maintained dual-task attentional processes that may not be as prevalent when the tasks are more discrete in time. These are only some of the potential reasons why we were able to find gamma-band activity related to inattentional deafness/blindness where previous studies have not (Pitts et al., 2014; Hutchinson, 2019), even though it is known that gamma-band activity is important for attention and performance (Rieder et al., 2011; Clayton et al., 2015). It is beyond the scope of this study to determine which of these potentially multiple reasons are responsible for our finding of auditory performance-related gamma-band activity. Although we surmise that engaging the natural human attention system using complex ecologically valid continuous immersive tasks may have something to do with it.

It is important to acknowledge that there are a number of limitations to this study. The first involves the nature of the auditory difference detection task employed. Although the auditory difference detection task involves greater involvement of attention, which was the motivation for its use, one cannot discriminate whether performance is a result of missing the previous stimulus (prime) or the target, or both. Given the relatively low false alarm rate we can be somewhat confident that hits actually represent true hits and not guesses, it is difficult to ascertain whether a

miss was due to an inability to sustain the previous stimulus in echoic memory, or an inability to perceive the current stimulus. Furthermore, we cannot discern whether misses were actually perceived, but there was an inability to be able to respond.

Another limitation is related to differences in the way inattentional deafness/blindness is defined by some as an unexpected event or object. In our task, as well as the tasks of many other studies investigating inattentional deafness, the participant is well aware that there are other stimuli not related to the primary task that are being presented as part of the experiment. In our study, since a dual-task paradigm is employed, the participants have explicit knowledge of the stimuli presented in the secondary task. Our experiment employed virtual skateboarding as the primary task and auditory difference detection as the secondary task. According to our definition of inattentional deafness/blindness, it is the process of selective attention being directed elsewhere that causes a lack of the ability to consciously perceive and act upon an object or event, that is of interest.

Although we utilized a comparison between the high workload condition (dual task during Wii skateboarding) and a low workload condition (transition period between Wii skateboarding levels) to assess the involvement of attentional load on performance above that of just perceptual processing, the results were not conclusive perhaps due to limitations of the low workload condition. Even during the low workload condition, performance on the auditory task was considerably low (hit rate equal to 0.431), suggesting that it may not actually be a low workload condition but rather also a moderately high workload condition that is significantly easier than the high workload condition occurring while doing Wii skateboarding (hit rate equal to 0.189). While the transition period between Wii levels certainly appears to have less dual-task attentional load, this period nevertheless involves a button press to continue to the next level that may direct attention away from the auditory task. The additional activity in IFG known to be involved with stimulus induced attentional switching (Doeller et al., 2003; Perianez et al., 2004; Tamber-Rosenau et al., 2018) in the high workload vs. low workload comparison that are not present in the high workload contrast alone (ERSP analyses Figures 3, 4 and Tables 1, 2), may reflect to some extent these differences in task-related attentional switching demands. Another potential reason why the low workload condition did not have the expected high performance as predicted, may be due to the continuous nature of the tasks. It may be the case that the suppressive attentional mechanisms active during the high workload dual-task condition (during Wii skateboarding) are maintained during the brief transition period between levels.

The apparent low auditory performance level for both the high workload and the low workload conditions represents a potential floor effect that could have important implications with respect to the results found and lack of results predicted to be found in this study. When performance is so low, it is difficult to discern whether participants were actually doing the auditory task or not. The degree of dual-task attentional engagement is considerably different in these two cases. Given the significant differential activity between misses and hits in theoretically relevant regions involved with attentional processing and inattentional deafness (Durantin et al., 2017), we do not believe this to be the case in our study. In addition, the variability in auditory task performance across participants was enough to reveal significant correlation with brain activity for hits relative to misses in theoretically relevant brain regions (mainly the auditory processing regions, see Figures 3B, C and Table 1). However, as discussed above (with respect to the high workload vs. low workload comparison), the apparent floor effect in performance may be responsible for the lack of findings that were predicted. It was predicted that there would be a tradeoff between Wii game performance and auditory task performance. However, this relationship was not found at the individual participant level across the course of the Wii games composing the experiment or across participants (See behavioral results). Although no significant correlation was found between the participants' performance in the Wii game and their performance in the auditory task, an interesting discovery was made. There was a significant negative correlation between the maximum Wii score achieved by the participants and the gamma-band ERSP activity in the left auditory cortex (see Figures 3E, 4D). This suggests that individuals who are more proficient in the Wii game exhibit lower neural processing of auditory stimuli in this particular brain region. This could represent suppressive processes resulting from inattentional deafness and directed attention to the Wii skateboarding task over the auditory task. The floor effect in auditory task performance may be hiding the behavioral relationship to support this conclusion.

Another potential limitation of our study is that the artifact removal and brain activity extraction procedures utilizing regression of head movement correlated activity, such as ASR, ICA, and IC label, are not able to completely separate artifact from brain activity. It is likely that some relevant task-related activity has been removed from the data along with the artifacts. Therefore, the lack of finding of a difference in brain activity between conditions does not mean that it is not actually present. It is also possible that differential activity found to be present between conditions is the result of differing degree of artifacts. A major concern when reporting gamma-band activity differences for EEG and MEG is the correspondence in the same frequency region as that of muscle activity (Yuval-Greenberg et al., 2008; Muthukumaraswamy, 2013). Muscle activity is in the range from approximately 20 to 300 Hz and can be picked up by EEG electrodes especially for muscles involved with eye movement including saccades as well as muscles on the face, neck, and shoulders (Yuval-Greenberg et al., 2008; Muthukumaraswamy, 2013). Given the high movement demands and the visual motor nature of the Wii skateboarding task, these muscle-related artifacts are of considerable concern. As pointed out by Muthukumaraswamy (2013), one way to address potential muscle artifacts in EEG is to remove them using ICA (a procedure used in our preprocessing pipeline). One limitation Muthukumaraswamy (2013) points out of using ICA to remove muscle artifacts is that it requires an expert to identify which components correspond to brain and artifact components. The development of ICLabel (Pion-Tonachini et al., 2019), that was used in our study, which is known to have expert level performance at selection of brain and artifact components, to some degree accounts for this apparent limitation. While it is known that movement artifacts related to eye movement and muscle activity can contaminate EEG data, we took several steps to attempt to remove potential muscle-related artifacts from the EEG

data. The primary method used was ICA combined with ICLabel that identifies whether a component is mostly brain, muscle, eye movement, and other artifacts. All subjects had components for eye movement and other muscle-related artifacts that were removed from the data. In addition, ASR was used to remove non-stationary artifacts that may arise from movement. Nevertheless, it is possible that there is greater muscle activity during misses than hits as a result of greater involvement in the Wii skateboarding game during these time periods that could confound our results. One reason we do not believe this to be the case is that the muscle activity for eye movement and saccades should be source-localized to regions on the edges of the brain in the orbital frontal cortex near the eyes or in the inferior temporal cortical regions near the neck muscles (this was not the case, see Figures 3–5). The localization of brain activity in our study in theoretically relevant regions does suggest that they are not merely the result of contamination by muscle artifact. Another potential confound is the presence of a button press for hits but not for misses. While this is not an issue for pre-stimulus contrasts and those involving correlation across participants, it may be so for post-stimulus contrasts including the ERSP analysis. Given the lack of brain activity in motor planning or execution related areas involved with pressing a button, we do not believe that this was a confounding issue for our results.

Conclusion

The results of our experiment suggest some of the potential underlying neural attentional processes that are responsible for inattentional deafness. Under conditions of high workload, brain regions in the prefrontal cortex (SMFC and pre-SMA) are engaged to a greater extent as a result of greater processing related to monitoring and evaluation of which task-relevant brain processes to facilitate and which task-irrelevant brain processes to inhibit dependent on action demands. The result of this process is selective inhibition of task-irrelevant processes reflected by greater alpha activity (inhibition) in auditory processing regions evident for misses over hits. Additionally, in some individuals, efficient dual-task performance is facilitated by excitation of task-relevant processes in auditory processing regions reflected by greater gamma-band activity for hits over misses correlated with overall participant level auditory task performance. Gamma brain activity in regions in the inferior frontal junction (e.g., IFG and DLPFC) is important in integrating DAN (executive) and VAN (salience) during task switches and attention shifts in multi-task situations reflected by greater ERSP activity for hits over misses in our study. Although our study reveals some of the potential processes involved with occurrence of inattentional deafness, further research is needed using different tasks that better modulate workload and avoid a potential floor effect to see how well they generalize. Additionally, further experiments investigating the functional connectivity between frontal and perceptual processing areas are needed to better understand the mechanisms involved with selective excitation of task-relevant processes and selective inhibition of task-irrelevant processes as well as the processes involved with attentional shifts during multi-task situations.

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 Advanced Telecommunications Research Institute International Human Subject Review Committee. 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

DC: conceptualization, investigation, methodology, software, formal analysis, visualization, and writing—original draft. TF: investigation. FD: writing—original draft and writing—review and editing. SI: writing—review and editing and funding acquisition. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by a project commissioned by the New Energy and Industrial Technology Development Organization (NEDO), Grant Number: JPNP20006. The funder played no roll in the design, data collection, analysis, and writing of this research.

Acknowledgments

We would like to specially thank the support staff at ATR including Yuki Shiojiri, Yoko Matsumoto, Kaori Nakamura, and Noriko Masui for making this research possible.

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

References

Allport, A. (1987). Selection for action: Some behavioral and neurophysiological considerations of attention and action. In: Heuer, H., and Sandears, A. F., editors. *Perspectives on Perception and Action*. Hillsdal, NJ: Lawrence Erlbaum Associates. p. 395–419.

Basar, B. (1998). *Brain Function and Oscillations I: Brain Oscillations, Principles and Approaches*. Berlin: Springer.

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K., and Robbins, K. (2015). The PREP pipeline: standardized preprocessing for large-scale EEG analysis. *Front. Neuroinform.* 9, 1–20. doi: 10.3389/fninf.2015.00016

Buschman, T. J., and Kastner, S. (2015). From behavior to neural dynamics: an integrated theory of attention. *Neuron* 88, 127–144. doi: 10.1016/j.neuron.2015.09.017

Callan, D., Gamez, M., Cassel, D., Terzibas, C., Callan, A., Kawato, M., et al. (2012). Dynamic visuomotor transformation involved with remote flying of a plane utilizes the 'Mirror Neuron' system. *PLoS ONE* 7, 1–14. doi: 10.1371/journal.pone.0033873

Callan, D. E., Gateau, T., Durantin, G., Gonthier, N., and Dehais, F. (2018). Disruption in neural phase synchrony is related to identification of inattentional deafness in real-world setting. *Hum. Brain Map.* 39, 2596–2608. doi: 10.1002/hbm.24026

Chang, C., Hsu, S., Pion-Tonachini, L., and Jung, T. P. (2018). Evaluation of Artifact subspace reconstruction for automatic EEG artifact removal. In: *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Manhattan, NY: IEEE.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *J. Acoustical Soc. Am.* 25, 975–979. doi: 10.1121/1.1907229

Clayton, M., Yeung, N., and Kadosh, R. C. (2015). The roles of cortical oscillations in sustained attention. *Trends Cogn. Sci.* 18, 188–195. doi: 10.1016/j.tics.2015.02.004

Dalton, P., and Fraenkel, N. (2012). Gorillas we have missed: sustained inattentional deafness for dynamic events. *Cognition* 124, 367–372. doi: 10.1016/j.cognition.2012.05.012

Dehais, F., Hodgetts, H. M., Causse, M., Behrend, J., Durantin, G., and Tremblay, S. (2019a). Momentary lapse of control: a cognitive continuum approach to understanding and mitigating perseveration in human error. *Neurosci. Biobehav. Rev.* 100, 252–262. doi: 10.1016/j.neubiorev.2019.03.006

Dehais, F., Karwowski, W., and Ayaz, H. (2020). Brain at work and in everyday life as the next frontier: grand field challenges for neuroergonomics. *Front. Neuroergon.* 1, 583733. doi: 10.3389/fnrgo.2020.583733

Dehais, F., Roy, R., and Scannella, S. (2019b). Inattentional deafness to auditory alarms: Inter-individual differences, electrophysiological signature and single trial classification. *Behav. Brain Res.* 360, 51–59. doi: 10.1016/j.bbr.2018.11.045

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Method.* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009

Doeller, C. F., Opitz, B., Mecklinger, A., Krick, C., Reith, W., and Schroger, E. (2003). Prefrontal cortex involvement in preattentive auditory deviance detection; neuroimaging and electrophysiological evidence. *NeuroImage* 20, 1270–1282. doi: 10.1016/S1053-8119(03)00389-6

Durantin, G., Dehais, F., Gonthier, N., Terzibas, C., and Callan, D. E. (2017). Neural signature of inattentional deafness. *Hum. Brain Map.* 38, 5440–5455. doi: 10.1002/hbm.23735

Dux, P. E., Ivanoff, J., Asplund, C. L., and Marois, R. (2006). Isolation of a central bottleneck of information processing with time-resolved fMRI. *Neuron* 52, 1109–1120. doi: 10.1016/j.neuron.2006.11.009

Edelman, G. (1989). *The Remembered Present: A Biological Theory of Consciousness*. New York, NY: Basic Books.

Ergenoglu, T., Demiralp, T., Bayraktarogul, Z., Ergen, M., Beydagi, H., and Uresin, Y. (2004). Alpha rhythm of the EEG modulates visual detection performance in humans. *Cogn. Brain Res.* 20, 376–383. doi: 10.1016/j.cogbrainres.2004.03.009

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

Fuchs, M., Kastner, J., Wagner, M., Hawes, S., and Ebersole, J. (2002). A standardized boundary element method volume conductor model. *Clin. Neurophysiol.* 113, 702–712. doi: 10.1016/S1388-2457(02)00030-5

Giraudet, L., St-Louis, M.-E., Scannella, S., and Causse, M. (2015). P300 event-related potential as an indicator of inattentional deafness? *PLoS ONE 10*: e0118556. doi: 10.1371/journal.pone.0118556

Golumbic, E. M. Z., Ding, N., Bickel, S., Lakatos, P., Schevon, C. A., McKhann, G. M., et al. (2013). Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party". *Neuron* 77, 980–991. doi: 10.1016/j.neuron.2012.12.037

Gramann, K., Ferris, D., Gwin, J., and Makeig, S. (2014). Imaging natural cognition in action. *Int J Psychophysiol.* 91, 22–29. doi: 10.1016/j.ijpsycho.2013.09.003

Gramann, K., McKendrick, R., Baldwin, C., Roy, R. N., Jeunet, C., Mehta, R. K., et al. (2021). Grand field challenges for cognitive neuroergonomics in the coming decade. *Front. Neuroergon.* 2, 643969. doi: 10.3389/fnrgo.2021.643969

Guzzon, D., and Casco, C. (2011). The effect of visual experience on texture segmentation without awareness. *Vision Res.* 51, 2509–2516. doi: 10.1016/j.visres.2011.10.006

Hanslmayr, S., Aslan, A., Staudigl, T., Klimesch, W., Herrmann, C. S., and Bauml, K. (2007). Prestimulus oscillations predict visual perception performance between and within subjects. *NeuroImage* 37, 1465–1473. doi: 10.1016/j.neuroimage.2007.07.011

Harris, A., Dux, P., and Mattingley, J. (2020). Awareness is related to reduced post-stimulus alpha power: a no-report inattentional blindness study. *Eur J Neurosci.* 52, 4411–4422. doi: 10.1111/ejn.13947

Henson, R. N. A., Burgess, N., and Frith, C. D. (2000). Recoding, storage, rehearsal and grouping in verbal short-term memory: an fMRI study. *Neuropsychologia* 38, 426–440. doi: 10.1016/S0028-3932(99)00098-6

Hutchinson, B. (2019). Toward a theory of consciousness: a review of the neural correlates of inattentional blindness. *Neurosci. Biobehav. Rev.* 104, 87–99. doi: 10.1016/j.neubiorev.2019.06.003

Hutchinson, B., Pammer, K., and Bandara, K. (2020). tACS stimulation at alpha frequency selectively induces inattentional blindness. *Brain Topogr.* 33, 317–326. doi: 10.1007/s10548-020-00762-9

Hutchinson, B., Pammer, K., and Jack, B. (2021). Pre-stimus alpha predicts inattentional blindness. *Conscious. Cogn.* 87, 103034. doi: 10.1016/j.concog.2020.103034

Ivry, R., and Robertson, L. (1998). *The Two Sides of Perception*. Cambridge, MA: MIT Press.

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

Koreimann, S., Gula, B., and Vitouch, O. (2014). Inattentional deafness in music. *Psychol. Res.* 78, 304–312. doi: 10.1007/s00426-014-0552-x

Kreitz, C., Furley, P., Simons, D. J., and Memmert, D. (2016). Does working memory capacity predict cross-modally induced failures of awareness? *Conscious. Cogn.* 39, 18–27. doi: 10.1016/j.concog.2015.11.010

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *TRENDS Cogn. Sci.* 9, 75–82. doi: 10.1016/j.tics.2004.12.004

Limanowski, J., and Blankenburg, F. (2016). Integration of visual and proprioceptive limb position information in human posterior parietal, premotor, and extrastriate cortex. *J. Neurosci.* 36, 2582–2589. doi: 10.1523/JNEUROSCI.3987-15.2016

Macdonald, J. S., and Lavie, N. (2011). Visual perceptual load induces inattentional deafness. *Atten. Percep. Psychophy.* 73, 1780–1789. doi: 10.3758/s13414-011-0144-4

Mack, A., and Rock, I. (1998). *Inattentional Blindness*. Cambridge, MA: MIT press.

Makeig, S. (1993). Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. *Electroencephalography and Clin. Neurophysiol.* 86, 283–293. doi: 10.1016/0013-4694(93)90110-H

Makeig, S., Delorme, A., Westerfield, M., Jung, T. P., Townsend, J., Courchesne, E., et al. (2004). Electroencephalographic brain dynamics following manually responded visual targets. *PLoS Biol.* 2, 747–762. doi: 10.1371/journal.pbio.0020176

Makeig, S., Gramann, K., Jung, T. P., Sejnowski, T. J., and Poizner, H. (2009). Linking brain, mind and behavior. *Int. J. Psychophysiol.* 73, 95–100. doi: 10.1016/j.ijpsycho.2008.11.008

Marois, R., Yi, D. J., and Chun, M. M. (2004). The neural fate of consciously perceived and missed events in the attentional blink. *Neuron* 41, 465–472. doi: 10.1016/S0896-6273(04)00012-1

Molloy, K., Griffiths, T. D., Chait, M., and Lavie, N. (2015). Inattentional deafness: Visual load leads to time-specific suppression of auditory evoked responses. *J. Neurosci.* 35, 16046–16054. doi: 10.1523/JNEUROSCI.2931-15.2015

Mullen, T., Kothe, C., Chi, Y. M., Ojeda, A., Kerth, T., Makeig, S., et al. (2013). Real-time modeling and 3D visualization of source dynamics and connectivity using wearable EEG. *Conf. Proc. IEEEE Eng. Med. Biol. Soc.* 2013, 2184–2187. doi: 10.1109/EMBC.2013.6609968

Muthukumaraswamy, S. (2013). High-frequency brain activity and muscle artifacts in MEG/EEG: a reviw and recommendations. *Front. Hum. Neurosci.* 7, 1–11. doi: 10.3389/fnhum.2013.00138

Neumann, O. (1987). Beyond capacity: A functional view of attention. In: Heuer, H., and Sanders, A. F., Hillsdale, N. J., editors. *Perspectives on Perception and Action*. Mahwah, NJ: Erlbaum Associates 361–394.

Nichols, T. E., and Holmes, A. P. (2001). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum. Brain Map.* 15, 1–25. doi: 10.1002/hbm.1058

Parasuraman, R., and Rizzo, M. (2008). *Neuroergonomics: The Brain at Work*. New York, NY: Oxford University Press.

Pascual-Marqui, R. D. (2002). Standardized low- resolution brain electromagnetic tomography (sloreta): technical details. *Methods Find. Exp. Clin. Pharmacol.* 24 (Suppl. D), 5–12.

Perianez, J. A., Maestu, F., Barcelo, F., Fernandez, A., Amo, C., and Alonso, T. O. (2004). Spatiotemporal brain dynamics during preparatory set shifting; MEG evidence. *NeuroImage* 21, 687–695. doi: 10.1016/j.neuroimage.2003.10.008

Pion-Tonachini, L., Kreutz-Delgado, K., and Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage* 198, 181–197. doi: 10.1016/j.neuroimage.2019.05.026

Pitts, M. A., Martínez, A., and Hillyard, S. A. (2012). Visual processing of contour patterns under conditions of inattentional blindness. *J Cogn Neurosci* 24, 287–303. doi: 10.1162/jocn_a_00111

Pitts, M. A., Padwal, J., Fennelly, D., Martinez, A., and Hillyard, S. A. (2014). Gamma band activity and the p3 reflect post-perceptual processes, not visual awareness. *NeuroImage* 101, 337–350. doi: 10.1016/j.neuroimage.2014.07.024

Raveh, D., and Lavie, N. (2015). Load-induced inattentional deafness. *Atten. Percept. Psychophys.* 77, 483–492. doi: 10.3758/s13414-014-0776-2

Rieder, M., Rahm, B., Williams, J., and Kaiser, J. (2011). Human gamma-band activity and behavior. *Int. J. Psychophysiol.* 79, 39–48. doi: 10.1016/j.ijpsycho.2010.08.010

Ruotolo, F., Ruggiero, G., Raemaekers, M., Iachini, T., van Der Ham, I. J. M., et al. (2019). Neural correlates of egocentric and allocentric frames of reference combined with metric and non-metric spatial relations. *Neuroscience* 409, 235–252. doi: 10.1016/j.neuroscience.2019.04.021

Ruz, M., Worden, M. S., Tudela, P., and McCandliss, B. D. (2005). Inattentional amnesia to words in a high attentional load task. *J. Cogn. Neurosci.* 17, 768–776 doi: 10.1162/0898929053747685

Sallet, J., Mars, R., Noonan, M, Neubert, F., Jbabdi, S., O'Reilly, J., et al. (2013). The organization of dorsal frontal cortex in humans and macaques. *J. Neurosci.* 33, 12255–12274. doi: 10.1523/JNEUROSCI.5108-12.2013

Sasaki, M., Iversen, J., and Callan, D. E. (2019). Music improvisation is characterized by increase EEG spectral power in prefrontal and perceptual motor cortical sources and can be reliably classified from non-improvisatory performance. *Front. Hum. Neurosci.* 13, 435. doi: 10.3389/fnhum.2019.00435

Scheer, M., Bülthoff, H., and Chuang, L. (2018). Auditory task irrelevance: a basis for inattentional deafness. *Hum. Factors* 60, 428–440. doi: 10.1177/0018720818760919

Schelonka, K., Graulty, C., Canseco-Gonzalez, E., and Pitts, M. A. (2017). ERP signatures of conscious and unconscious word and letter perception in an inattentional blindness paradigm. *Conscious. Cogn.* 54, 56–71. doi: 10.1016/j.concog.2017.04.009

Schlossmacher, I., Dellert, T., Bruchmann, M., and Straube, T. (2021). Dissociating neural correlates of consciousness and task relevance during auditory processing. *Neuroimage*, *228*, 117712. doi: 10.1016/j.neuroimage.2020.117712

Schubo, A., Meinecke, C., and Schroger, E. (2001). Automaticity and attention: investigating automatic processing in texture segmentation with event-related brain potentials. *Cogn. Brain Res.* 11, 341–361. doi: 10.1016/S0926-6410(01)00008-8

Simmons, D., and Chabris, C. (1999). Gorillas in our midst: sustained inattentional blindness for dynamic events. *Perception* 28, 1059–1074. doi: 10.1068/p281059

Somon, B., Roy, R., Simonetti, I., and Dehais, F. (2022). Ecological measures of cognitive impairments in aeronautics: theory and application. In: Fairclough, S., Zander, T., editors. *Current Research in Neuroadaptive Technology*. London: Academic Press. p. 117–138.

Szameitat, A., Vanloo, A., and Muller, H. (2016). Central as well as peripheral attentional bottle necks in dual-task performance activate lateral prefrontal cortices. *Front. Hum. Neurosci.* 10, 119. doi: 10.3389/fnhum.2016.00119

Tallon-Baudry, C. (2009). The roles of gamma-band oscillatory synchrony in human visual cognition. *Front. Biosci.* 14, 321–332. doi: 10.2741/3246

Tamber-Rosenau, B. J., Asplund, C. L., and Marois, R. (2018). Functional dissociation of the inferior frontal junction from the dorsal attention network in top-down attentional control. *J. Neurophysiol.* 120, 2498–2512. doi: 10.1152/jn.00506.2018

Todd, J. J., Fougnie, D., and Marois, R. (2005). Visual short-term memory load suppresses temporo-parietal junction activity and induces inattentional blindness. *Psychol. Sci.* 16, 965–972. doi: 10.1111/j.1467-9280.2005.01645.x

Tombu, M. N., Asplund, C. L., Dux, P. E., Godwin, D., Martin, J. W., and Marois, R. (2011). A unified attentional bottleneck in the human brain. *Proc. Natl. Acad. Sci.* 108, 13426–13431. doi: 10.1073/pnas.1103583108

van Dijk, H., Schoffelen, J. M., Oostenveld, R., and Jensen, O. (2008). Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *J. Neurosci.* 28, 1816–1823. doi: 10.1523/JNEUROSCI.1853-07.2008

Vossel, S., Geng, J. J., and Fink, G. R. (2014). Dorsal and ventral attention systems: distinct neural circuits but collaborative roles. *Neuroscientist* 20, 150–159. doi: 10.1177/1073858413494269

Wittenberg, M., Baumgarten, T., Schnitzler, A., and Lange, J. (2018). U-shaped relation between prestimulus alpha and poststimulus gamma-band power in temporal tactile perception in the human somatosensory cortex. *J. Cogn. Neurosci.* 30, 552–564. doi: 10.1162/jocn_a_01219

Womelsdorf, T., and Fries, P. (2007). The role of neuronal synchronization in selective attention. *Current Opin. Neurobiol.* 17, 154–160. doi: 10.1016/j.conb.2007.02.002

Wyart, V., and Tallon-Baudry, C. (2009). How ongoing fluctuations in human visual cortex predict perceptual awareness: baseline shift versus decision bias. *J. Neurosci.* 29, 8715–8725. doi: 10.1523/JNEUROSCI.0962-09.2009

Yuan, X., Li, H., Liu, P., Yuan, H., and Huang, X. (2016). Prestimulus beta and gamma oscillatory power predicts perceived audiovisual simultaneity. *Int. J. Psychophysiol.* 107, 29–36. doi: 10.1016/j.ijpsycho.2016.06.017

Yuval-Greenberg, S., Tomer, O., Keren, A. S., Nelken, I., and Deouell, L. Y. (2008). Transient induced gamma-band response in EEG as a manifestation of miniature saccades. *Neuron* 58, 429–411. doi: 10.1016/j.neuron.2008.03.027

Zhou, Y., Iemi, L., Schoffelen, J., de Lange, F., and Haegens, S. (2021). Alpha oscillations shape sensory representation and perceptual sensitivity. *J. Neurosci.* 41, 9581–9592. doi: 10.1523/JNEUROSCI.1114-21.2021

frontiers | Frontiers in Human Neuroscience

TYPE Brief Research Report
PUBLISHED 22 June 2023
DOI 10.3389/fnhum.2023.1209555

OPEN ACCESS

EDITED BY
Tetsuo Kida,
Institute for Developmental Research, Japan

REVIEWED BY
John J. McDonald,
Simon Fraser University, Canada
Bettina Forster,
City, University of London, United Kingdom

*CORRESPONDENCE
Elena Gherri
✉ elena.gherri@unibo.it

RECEIVED 20 April 2023
ACCEPTED 08 June 2023
PUBLISHED 22 June 2023

CITATION
Gherri E, Fiorino FR, Iani C and Rubichi S (2023) Searching for a tactile target: the impact of set-size on the N140cc. *Front. Hum. Neurosci.* 17:1209555. doi: 10.3389/fnhum.2023.1209555

Searching for a tactile target: the impact of set-size on the N140cc

Elena Gherri[1]*, Fabiola Rosaria Fiorino[2], Cristina Iani[3,4] and Sandro Rubichi[2,4]

[1]Dipartimento di Filosofia e Comunicazione, University of Bologna, Bologna, Italy, [2]Department of Biomedical, Metabolic and Neural Science, University of Modena and Reggio Emilia, Modena, Italy, [3]Department of Surgery, Medicine, Dentistry and Morphological Sciences With Interest in Transplant, Oncology and Regenerative Medicine, University of Modena and Reggio Emilia, Modena, Italy, [4]Center of Neuroscience and Neurotechnology, University of Modena and Reggio Emilia, Modena, Italy

The time needed to find a visual target amongst distractors (search task) can increase as a function of the distractors' number (set-size) in the search-array (inefficient search). While the allocation of attention in search tasks has been extensively investigated and debated in the visual domain, little is known about these mechanisms in touch. Initial behavioral evidence shows inefficient search behavior when participants have to distinguish between target and distractors defined by their vibro-tactile frequencies. In the present study, to investigate the allocation of attention to items of the search-array we measured the N140cc during a tactile task in which the set-size was manipulated. The N140cc is a lateralized component of event-related brain potentials recently described as a psychophysiological marker of attentional allocation in tactile search tasks. Participants localized the target, a singleton frequency, while ignoring one, three or five homogeneous distractors. Results showed that error rates linearly increased as a function of set-size, while response times were not affected. Reliable N140cc components were observed for all set-sizes. Crucially, the N140cc amplitude decreased as the number of distractors increased. We argue that the presence of additional distractors hindered the preattentive analysis of the search array resulting in increased uncertainty about the target location (inefficient preattentive stage). This, in turn, increased the variability of the deployment of attention to the target, resulting in reduced N140cc amplitudes. Consistent with existing behavioral evidence, these findings highlight systematic differences between the visual and the tactile attentional systems.

KEYWORDS

touch, event-related potentials (ERP), selective attention, N140cc, tactile search, set-size

1. Introduction

The ability to select relevant information from a cluttered environment strongly depends on selective attention mechanisms. These mechanisms are typically investigated in search tasks in which participants have to identify the presence (or a feature) of a task relevant item (target) presented simultaneously with a number of irrelevant ones (distractors). Search efficiency is typically assessed by manipulating the number (set-size) and features of distractors presented with the target. Efficient search behavior is unaffected by set-size, as indicated by a flat slope function between performance indicators

(response times and/or accuracy) and set-size and is usually observed when target and distractors can be distinguished on the basis of a distinctive feature (e.g., Egeth et al., 1972; Treisman and Gelade, 1980). The preattentive processing of all search-array items in parallel guides attention to the target (parallel search) (e.g., Wolfe, 1994, 2021). By contrast, inefficient search is characterized by increasingly worst performance as a function of set-size (e.g., Treisman and Gelade, 1980; Wolfe, 1994) and is often observed when target and distractors are highly similar (e.g., Duncan and Humphreys, 1989) or when they differ in terms of a combination of features (e.g., Treisman, 1991). Inefficient behavior, in terms of increase of RTs and errors as a function of set-size, has been suggested to reflect the serial allocation of attention to single items in the array (serial search) until the target is identified (e. g., Treisman and Gelade, 1980). However, this pattern of behavior can also be explained by a delayed shift of attention to the target due to longer preattentive processing of the array (*inefficient preattentive stage*, e.g., Folk and Remington, 1998) or by a longer attentional processing of the target (*inefficient attentional processing*, e.g., Christie et al., 2015).

While most studies on search behavior involved visual search tasks, researchers have started to address analogous questions in the tactile domain. When participants were allowed to haptically explore the tactile stimuli presented to their fingers, results revealed that search efficiency depended on the specific dimension defining target and distractors (c.f. Lederman and Klatzky, 1997; Overvliet et al., 2007). Efficient search was observed with abrupt surface discontinuity (e.g., deep vs. shallow hole) and material (e.g., roughness), while inefficient search was reported when participants searched for specific contours (e.g., slant vs. curve surface) and relative orientation (vertical vs. horizontal) (c.f. Lederman and Klatzky, 1997; Overvliet et al., 2007).

Other researchers have exploited the ability of the somatosensory system to discriminate the frequencies of different vibrotactile stimuli (e.g., Johansson and Vallbo, 1979; Bark et al., 2008). Groen et al. (2008) presented the array to the participants' abdomen and manipulated both the set-size and the similarity between target and distractors. While search was inefficient on target present trials, no set-size effect was observed on target absent trials and no effect of target-distractor similarity was observed. Similar results were reported by Halfen et al. (2020) in a passive tactile search task in which vibrotactile stimuli were delivered to different parts of the body (hand, back, legs, etc.). Results showed set-size effects on RTs of both target-present and target-absent trials. Thus, in touch, search appears more inefficient on target present than on target absent trials, in contrast to evidence from the visual domain (e.g., Wolfe, 2021).

Event-related potential (ERP) studies have recently identified the electrophysiological correlate of target selection in tactile search tasks (e.g., Katus et al., 2015; Forster et al., 2016; Ambron et al., 2018; Katus and Eimer, 2019; Mena et al., 2020; Gherri et al., 2021, 2022). ERPs elicited by the search-array are typically more negative over the hemisphere contralateral to the target side compared to those observed over the ipsilateral hemisphere. This lateralized ERP component (labeled N140cc, N2cc, or CCN in different studies) is typically elicited over somatosensory areas from about 100–140 ms from the onset of the search-array (e.g., Forster et al., 2016). The presence of the N140cc in tactile search tasks was deemed particularly interesting due to the analogies with the N2pc component observed in visual search tasks (e.g., Luck and Hillyard, 1994; Eimer, 1996; Luck et al., 1997). The N2pc is considered an ERP marker of attentional allocation in the visual field and has been widely used to investigate the mechanisms underlying selective attention in visual search tasks over the last 30 years (for reviews see Luck, 2012; Eimer, 2014). Hence, recent studies have started to assess whether and to what extent the tactile N140cc can be considered the functional equivalent of the visual N2pc (Katus et al., 2015; Forster et al., 2016; Katus and Eimer, 2019; Mena et al., 2020; Gherri et al., 2021, 2022).

In a series of recent tactile search ERP studies, participants were asked to localize a target presented simultaneously with one salient but irrelevant singleton distractor and with several other homogeneous distractors (Mena et al., 2020; Gherri et al., 2021, 2022). On target-absent trials, an N140cc was elicited contralateral to the singleton distractor (Mena et al., 2020). On target-present trials, the amplitude of the target-elicited N140cc was reduced in the presence of a contralateral singleton distractor (Mena et al., 2020; Gherri et al., 2022). These observations suggest that the N140cc component reflects the allocation of tactile spatial attention to potentially task-relevant items in the array.

Notably, existing evidence has also shown potentially relevant differences between the mechanisms responsible for visual and tactile target selection. In the visual domain, the interference created by a singleton distractor increased as the distance between target and distractor decreases within the same hemifield, as revealed by worst performance and reduced N2pc components for contiguous singletons, due to the progressively overlapping neural resources assumed to represent these items (e.g., Hilimire et al., 2009, 2010; Gaspar and McDonald, 2014). In touch, however, performance improved (but the N140cc amplitude decreased) when the target and the singleton distractor were presented to contiguous fingers of the same hand (Gherri et al., 2021). This reveals systematic differences in the processes underlying target selection in the presence of a salient distractor between vision and touch.

The studies reviewed above can be considered as a first attempt to characterize the functional meaning of the N140cc component and the mechanisms responsible for the selection of relevant information in touch. However, basic questions related to search efficiency and its impact on the N140cc component remain completely unexplored. In the visual modality, different search tasks have yielded different pattern of N2pc modulations by set-size, with some studies showing a decreasing N2pc with increasing distractors (e.g., Mazza et al., 2009; Tay et al., 2022), while others showing set-sizes effects on the N2pc duration (e.g., Christie et al., 2015). To investigate whether, similarly to the visual modality, set-size modulates the N140cc component elicited during a tactile search task, we asked participants to localize a target (on the left or right hand) while ignoring simultaneous distractors. In line with existing ERP studies, target and distractors differed with respect to their vibrotactile frequencies. On different trials, one, three or five homogeneous distractors were presented with the target (see **Figure 1**). In line with existing evidence, if performance in the search task is affected by set-size (c.f. Groen et al., 2008; Halfen et al., 2020), we expect to observe also a modulation of the N140cc amplitude.

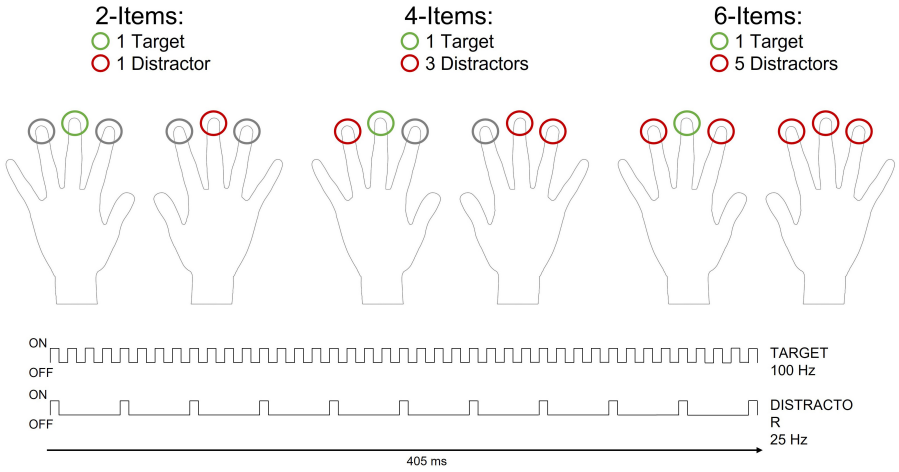


FIGURE 1
Schematic representation of the tactile search-array (405 ms duration) and the different types of set-sizes included in the present study. On each trial one target was presented to the index, middle or ring finger of the left or right hand (green circle). One, three, or five homogenous distractors were simultaneously presented with the target on different trials (red circles). The number of distractors varied randomly across trials. All search arrays included a symmetrical configuration of tactile stimuli over the left and right hand such that stimuli were always presented to mirror-symmetric (homologous) fingers of opposite hands. Target and distractors differed with respect to their vibrotactile frequencies. The target vibration (100 Hz) consisted of a sequence of pulses during which the rod was in contact with the skin for 5 ms ("ON"), followed by an inter-pulse interval of 5 ms ("OFF") in which no contact was being made. For the distractor vibration vibrations (25 Hz), ON pulses of 5 ms were interleaved by OFF periods of 35 ms.

2. Materials and methods

2.1. Participants

A total of thirty-seven volunteers were recruited via word-of-mouth at the University of Modena and Reggio Emilia. All of them had a normal or corrected-to-normal vision and no history of neurological disorders. Three participants were excluded from the analyses due to low accuracy levels in the behavioral task (overall accuracy levels across set-sizes below 60%). Following the ERP data processing procedure, six participants' data were excluded due to a low number of trials (less than 60% of trials left after rejection of ERP artifacts for at least one of the set-size conditions). A total of 28 participants remained in the sample (23 females and 5 males, M_{age} = 21 years, SD_{age} = 3.6 years; 27 right handed and 1 ambidextrous, Oldfield, 1971).

This project was approved by the Area Vasta Emilia Nord (AVEN) Ethics Committee and followed the Helsinki Declaration principles. All participants signed an informed consent before starting the experiment.

2.2. Stimuli and apparatus

Participants placed their hands on a table with their palms down. Tactile stimuli were presented using 12 V solenoids (Heijo Research Electronics, London, UK) driving a metal rod with a blunt conical tip. The tip of the tactile stimulators touched the skin whenever a current passed through the solenoid. Six tactile stimulators were used in total, each one was attached with adhesive medical tape to the inner side of the top phalanx of the left and right index, middle and ring fingers. The distance between the index fingers of the two hands was 10 cm. Once in the correct position, the hands were covered with a black cardboard, on top of which there was a white pin aligned with the body midline which served as fixation-point. To mask the sounds made by tactile stimulators, one speaker was positioned on the table close to the hands and presented white noise (65 dB SPL) throughout the experimental blocks. Two vertically arranged foot-pedals (top and bottom pedals) were positioned under the toes and heel of one of the participants' feet. Participants were asked to keep one foot on these response pedals during the task.

On each trial, a search-array was presented to mirror-symmetric (homologous) fingers of the left and right hand. There were three different types of search-arrays in which two, four or six vibro-tactile stimuli were presented simultaneously (2-Items, 4-Items and 6-Items arrays, respectively, see **Figure 1**). One target stimulus was presented on each trial, while the number of distractor(s) varied on different trials (1, 3 or 5 distractors). Vibro-tactile stimuli differed with respect to their vibration frequencies (25 Hz or 100 Hz). The target was the fastest vibration (100 Hz), while the distractors were the slowest vibrations (25 Hz). These frequencies consisted of a rapid sequence of pulses during which the rod was in contact with the skin for 5 ms, followed by a variable inter-pulse interval set at 35 ms and 5 ms for the slow and fast vibrations, respectively (**Figure 1**). All stimuli had a total duration of 405 ms. The tactile search-array started and ended with all the stimulators touching the skin simultaneously to prevent participants from using the offset of the stimuli to complete the task.

2.3. Procedure

Each trial started with a 300 ms empty interval which was followed by the simultaneous presentation of the tactile target

and of the distractor(s) (405 ms duration). The search-array presentation was followed by a 1,800 ms interval which was used to collect foot responses. Thus, the interstimulus interval between search-arrays was set at 2.505 ms.

Ten blocks of 72 trials each were completed by participants. The tactile target was presented randomly and with equal probability to the index, middle or ring finger of the left or right hand (12 trials for each target location in each block). Within each block of trials, 2-, 4-, and 6-items search-arrays were equally likely (each presented on 24 trials per block).

Participants' task was to identify the location of the target (left or right hand) by pressing the top or bottom pedal with their toes or heels. They were instructed to keep their eyes on the central fixation at all times and to press the pedals as fast and as accurately as possible. To eliminate lateralized motor activity in the grand averaged ERP waveforms, participants completed five successive blocks responding with their right foot, and the remaining five blocks with their left foot. The order of the responding foot was counterbalanced across participants.

Prior to the beginning of the experiment participants completed two 36-trials practice blocks (which was repeated whenever average accuracy fell below 60% during this training), following a familiarization procedure with the stimuli frequencies. At the end of each block participants received verbal feedback about their performance (average response time and accuracy).

2.4. Electroencephalography recording and data analysis

Electroencephalography (EEG) was recorded with a BrainAmp amplifier system (500 Hz sampling rate) from 64 active electrodes positioned according to 10–20 system. EEG data was analyzed using Brain Vision Analyser (version 2.0.4.368). EEG was digitally re-referenced to the average of the left and right earlobe and was digitally filtered offline (high-pass filter 0.53 Hz, low-pass filter 40 Hz and notch filter 50 Hz). The EEG was epoched into 450 ms intervals starting 100 ms before and ending 350 ms after the search-array onset. Trials with eye blinks, horizontal eye movements and other artifacts (voltage exceeding ± 80 μV at any electrode sites) were excluded from further analysis, as were trials with response errors. Participants with less than 60% of the trials in at least one of the set-size conditions were excluded from the analyses. This led to the exclusion of 6 participants. The average number of trials included in each condition for the remaining 28 participants was 185 for the 2-items array (77% of these trials), 167 for the 4-Items (70%) and 150 for the 6-Items (63%).

Event-related potentials elicited by the presentation of the search array on correct trials were averaged relative to a 100 ms pre-stimulus baseline separately for all combinations of set-sizes (2-Items vs. 4-Items vs. 6-Items search-array) and target side (left vs. right hand). The N140cc was quantified for each participant and for each set-size on the basis of ERP mean amplitudes obtained at lateral central electrodes C3/4 and C5/6 (where this component was maximal in the present study, in line with previous studies from our lab, Gherri et al., 2021) over the hemisphere contralateral and ipsilateral to the target side in the 110–250 ms post-array onset measurement window (see Gherri et al., 2021, 2022). To investigate whether set-size modulated the amplitude of the N140cc component, a repeated-measures analysis of variance (ANOVA) was conducted on the pooled mean amplitude values measured at electrodes pairs C3/4 and C5/6 for the factors set-sizes (2-Items vs. 4-Items vs. 6-Items search-array) and laterality (hemisphere contralateral vs. ipsilateral to the target side). In these analyses, the presence of reliable lateralized components is reflected by the main effect of the factor laterality, indicating significant differences between the contralateral and ipsilateral hemispheres to the target side. Following significant laterality x set-sizes interactions, separate analyses were carried out for each set-size, to determine the presence of reliable N140cc lateralized components. Then, the difference waveforms between contralateral and ipsilateral ERPs were calculated, separately for the different set-sizes. To determine whether the N140cc amplitude was modulated by set-size, we run contrasts (by means of *t*-tests) between these difference waveforms observed for the different set-sizes (Bonferroni corrections were applied to these multiple comparisons). Finally, to test whether the N140cc amplitude was linearly related to the set-size of the search-array, polynomial orthogonal contrasts were carried out.

While the method of choice to measure the onset and offset of a lateralized component is the fractional peak latency obtained from jackknife-averaged ERPs (e.g., Miller et al., 1998; Ulrich and Miller, 2001; Kiesel et al., 2008) within the time window of interest (110–250 ms), the lack of clear peaks in the N140cc components makes this approach prone to distortions. Thus, to examine the time course of the target-elicited N140cc in the different search arrays, the mean N140cc amplitudes were also analyzed in 20 ms consecutive time windows starting 90 ms and terminating 310 ms post-array onset, separately for each set-size (e.g., McDonald et al., 2013; Christie et al., 2015). Results are shown in **Table 1**. In these analyses, for any three consecutive 20 ms time windows at least two had to reveal significant effects to indicate the reliable presence of the N140cc components.

TABLE 1 Comparisons between ipsilateral and contralateral waveforms computed separately for the three set-sizes for consecutive 20 ms time-windows, from 90 to 310 ms post-array onset.

20 ms time-windows	2-Items t (*p*-value)	4-Items t (*p*-value)	6-Items t (*p*-value)
90–110	−2.402 (0.023)*	−0.620 (0.540)	−0.833 (0.412)
110–130	−4.894 (0.001)*	−1.979 (0.058)	−0.958 (0.347)
130–150	−6.714 (0.000)*	−2.943 (0.007)*	−0.124 (0.019)*
150–170	−8.380 (0.000)*	−3.788 (0.001)*	−1.905 (0.067)
170–190	−5.820 (0.000)*	−3.856 (0.001)*	−3.333 (0.003)*
190–210	−5.982 (0.000)*	−3.507 (0.002)*	−2.975 (0.006)*
210–230	−5.707 (0.000)*	−4.374 (0.000)*	−3.197 (0.004)*
230–250	−5.653 (0.000)*	−2.926 (0.007)*	−2.079 (0.047)*
250–270	−2.379 (0.025)*	−0.791 (0.436)	−1.541 (0.135)
270–290	−1.262 (0.218)	−0.770 (0.448)	−2.002 (0.055)
290–310	0.336 (0.740)	−0.998 (0.327)	−1.990 (0.057)

Each cell shows the values of this *t*-test and the corresponding *p*-value in brackets. Asterisks denote the presence of a significant lateralization ($p < 0.05$). In these analyses, for any three consecutive 20 ms time windows at least two had to show significant effects to indicate the reliable presence of the N140cc components (highlighted in gray).

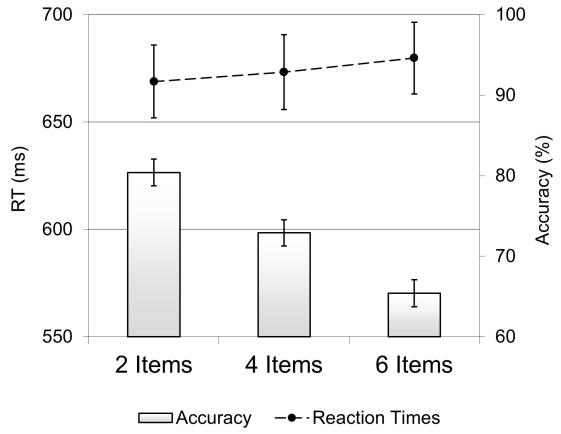


FIGURE 2
The effect of set-size on behavioral performance. Accuracy rates (bar graph) and mean RTs (line graph) in the 2, 4, and 6 items set-size trials. Error bars represent the standard errors of the means.

In addition to the planned analyses of the N140cc, an exploratory analysis was also performed on an earlier lateralized component emerged between 70 and 100 ms post-array onset. **Figures 3**, **4** clearly show this early lateralization overlapping with the N80 component, characterized by an amplitude modulation by set-size analogous to the one observed for the N140cc. To further quantify the statistical reliability of this early lateralization and its amplitude modulation by set-size, mean amplitude values obtained between 70 and 100 ms post-array onset were analyzed following the same analysis procedure adopted for the N140cc.

Mean RTs were calculated on correct response times trimmed to 2.5 standard deviations from the mean (calculated separately for each participant and each set-size). Mean reaction times (RTs) as well as accuracy rates were submitted to separate repeated measure ANOVAs with set-size (2 Items vs. 4 Items vs. 6 Items search-array) as within-subjects factor. Following set-sizes main effects, we run contrasts between the means observed for the different set-sizes (Bonferroni corrections were applied to these multiple comparisons). Similarly to the ERP analysis, polynomial orthogonal contrasts were also carried out to investigate whether mean RTs and accuracy rates were linearly dependent on set-size.

For all analyses, Greenhouse-Geisser adjustments to the degrees of freedom were applied where appropriate, and unadjusted *p*-values were reported.

3. Results

3.1. Behavioral results

For the accuracy data (**Figure 2**, bar graph), the ANOVA showed a main effect of set-size, $F(1.74, 46.9) = 95.9$, $p < 0.001$, $\eta_p^2 = 0.78$, with participants' performance being most accurate in the 2 Items array ($M = 80.4\%$, SE = 1.7), intermediate in the 4 Items array ($M = 73\%$, SE = 1.6) and worst in the 6 Item arrays ($M = 65\%$, SE = 1.7) (all Bonferroni adjusted contrasts, $p < 0.001$). Accuracy rates linearly decreased with increasing set-sizes, $F(1,27) = 144$, $p < 0.001$, $\eta_p^2 = 0.84$.

Results of the ANOVA carried out on RTs revealed no significant effect of set-size, $F(1.38,37.2) = 1.7$, $p = 0.189$, $\eta_p^2 = 0.06$, see **Figure 2** (line graph).

3.2. ERP results

Figure 3 shows ERP waveforms elicited at pooled electrodes C3/4 and C5/6 contralateral (solid lines) and ipsilateral (dashed lines) to the location of the tactile target, separately for each set-size. The corresponding difference waveforms (**Figure 4**) were obtained by subtracting ERPs elicited at electrodes ipsilateral to the target from contralateral ERPs. As can be seen from both **Figures 3**, **4**, the lateralized N140cc component was present approximately between 110 and 250 ms post array onset and appeared to be modulated by set-size. Interestingly, the N140cc was preceded by an earlier lateralization between 70 and 100 ms post array, similarly modulated by set-size. The scalp distribution of the lateralized components observed in the 110–250 ms post-array interval is shown in **Figure 4**.

3.2.1. Early lateralization (70–100 ms)

An exploratory analysis of the mean amplitude values measured between 70 and 100 ms post array onset showed that the main effect of set-size was not statistically significant, $F(1.9, 50.9) = 1.06$, $p = 0.35$, $\eta_p^2 = 0.038$, suggesting no differences between ERPs elicited by the different set-sizes. The main effect of the factor laterality, $F(1,27) = 18.2$, $p < 0.001$, $\eta_p^2 = 0.43$, revealed the presence of a reliable lateralized component between 70 and 100 ms regardless of set-size. This was further modulated by set-size (laterality × set-size interaction), $F(1.7,45) = 5.19$, $p < 0.013$, $\eta_p^2 = 0.16$. Pair-wise comparisons between ipsilateral and contralateral ERP amplitudes carried out separately for each set-size revealed the presence of reliable lateralizations for the 2-items array [$t(27) = 4.97$, $p < 0.001$, $d = 0.9$] and 4-items array [$t(27) = 2.95$, $p < 0.006$, $d = 0.6$] but not for the 6-items array [$t(27) = 1.47$, $p = 0.15$, $d = 0.3$]. The amplitude of this early lateralized component linearly decreased as a function of set-size [$F(1,27) = 8.1$, $p < 0.008$, $\eta_p^2 = 0.23$]. Bonferroni-adjusted contrasts carried out between the amplitudes observed for the different set-sizes showed a larger lateralized component for 2- Items arrays compared to 4- and 6-Items arrays (both $p < 0.033$). However, no statistical difference emerged between 4- and 6-Items arrays ($p = 0.9$).

3.2.2. N140cc (110–250 ms)

The planned analysis of mean amplitude values (measured in the 110–250 ms interval) revealed a significant main effect of set-size, $F(1.7,47.2) = 5.6$, $p < 0.006$, $\eta_p^2 = 0.17$, that was driven by the fact that ERP amplitudes increased with set-size regardless of laterality. The N140cc component was reliably present between 110 and 250 ms (main effect of the factor laterality, $F(1,27) = 40.02$, $p < 0.001$, $\eta_p^2 = 0.6$). Crucially, the interaction of interest between laterality and set-size was significant, $F(1.59,42.8) = 13.1$, $p < 0.001$, $\eta_p^2 = 0.33$. Pairwise comparisons between ipsilateral and contralateral ERP amplitudes carried out separately for each set-size revealed the presence of reliable lateralized N140cc components for each set-size (all $t(27) > 2.9$, $p < 0.007$, $d > 0.055$).

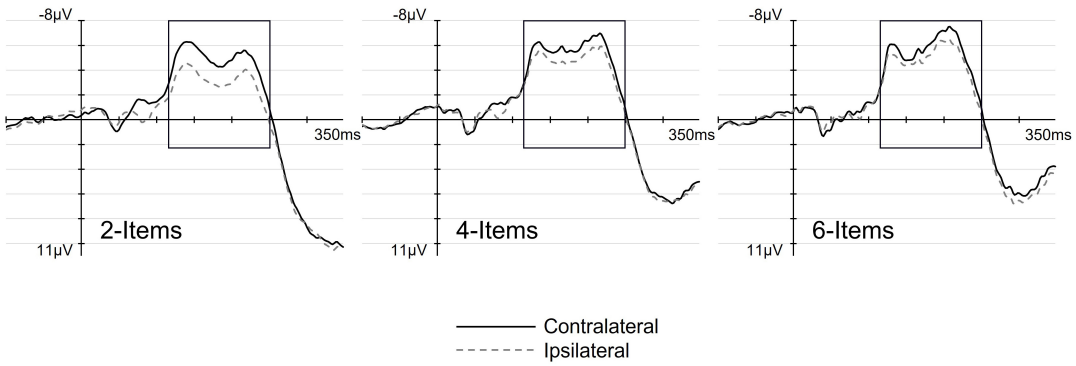


FIGURE 3
The effect of set-size on the N140cc lateralized component. Panels show ERPs elicited over pooled electrodes C3/4 and C5/6, contralateral (solid line) and ipsilateral (dashed line) to the target side, separately for the 2-, 4-, and 6-Items search-arrays.

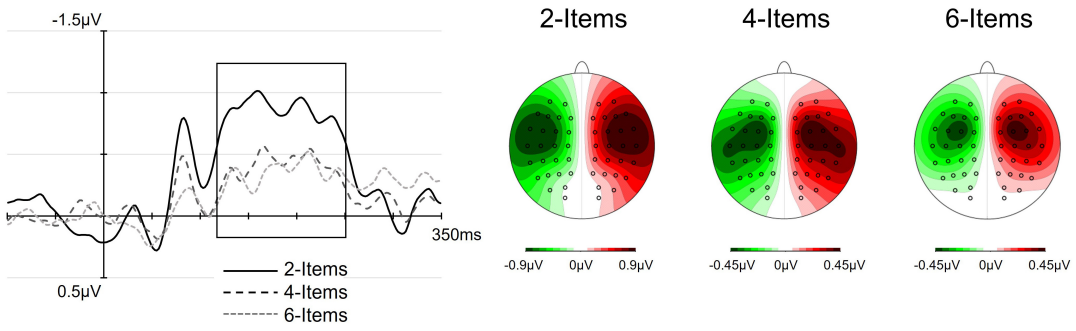


FIGURE 4
The effect of set-size on the N140cc lateralized component. **(Left)** panels show the difference waveforms (contralateral–ipsilateral waveforms) for pooled electrodes sites C3/4 and C5/6. In these figures, the black lines represent the N140cc amplitude in the 2-items array, the dark gray lines (long dash) represent the 4-items array and the light gray line (short dash) represent the 6-item array. The box shows the time window considered for statistical analyses (110–250 ms). **(Right)** panels show the scalp distribution of the lateralized N140cc component measured in the 110–250 ms interval, separately for each set-size.

The amplitude of the N140cc component linearly decreased as a function of set-size, $F(1,27) = 17.6$, $p < 0.001$, $\eta_p^2 = 0.39$. Bonferroni-adjusted contrasts carried out between the N140cc amplitude observed for the different set-sizes showed a larger N140cc for 2- Items arrays compared to 4 and 6 Items arrays (both $p < 0.001$). However, no statistical difference emerged between 4 and 6 Items arrays ($p = 0.6$).

To evaluate the time course of the N140cc we computed the mean amplitude values of the waveforms ipsilateral and contralateral to the target separately for each set-size in consecutive 20 ms intervals from 90 to 310 ms (see **Table 1**). Reliable N140cc emerged in the 2-Items array from 90 to 270 ms post-array. The N140cc was observed in the 4-Items array between 130 and 250 ms, while it was present between 170 and 250 ms in the 6-Items array.

4. Discussion

Within the more general purpose of understanding similarities and differences between the neural mechanisms responsible for the allocation of attention in the visual and tactile sensory domains, the aim of this study was to investigate whether the amplitude of the N140cc component elicited during a tactile search task is modulated by set-size. Participants showed inefficient search behavior, as indicated by the fact that accuracy rates linearly decreased as a function of distractors' number, although RTs were not affected by set-size. The N140cc component was reliably present contralateral to the target side regardless of the number of distractors. Crucially, the amplitude and time course of the N140cc component was modulated by set-size. The N140cc amplitude was maximal and emerged in earlier time-windows in the 2-items array and decreased linearly as the number of distractors increased.

According to Wolfe's (1994) Guided Search model, an activation map is created upon the presentation of the search-array. The location of the target is characterized by the largest activation in the map and attention is subsequently shifted toward this location. When the target pops out from the array, its identification is not hindered by increasing distractor numbers (efficient search) and does not require further attentional processing. By contrast, when the target-defining feature is relatively subtle, the quality of this pre-attentive processing can be affected by the distractor number and an in-depth attentional processing may be necessary for the target identification.

In the present study, the number of distractors reduced the accuracy of target localization showing that the target did not pop out from the array despite being more salient than the homogenous

distractors (c.f. Groen et al., 2008; Halfen et al., 2020). This is further supported by the presence of reliable N140cc for each set-size which confirms that attention was shifted to the target side for in-depth attentional processing. Because the N140cc is assumed to be time-locked to the shift of attention toward the target location (Mena et al., 2020), ERP differences between the different set-sizes can help to interpret the behavioral inefficiency observed in the present study. In search accuracy experiments, set-size effects are assumed to be directly related to capacity limits on the quality of processing (c.f. Shaw, 1984; Palmer et al., 1993). We argue that the manipulation of set-size already affected the preattentive analysis of the search array, leading to weaker activations of the target location on the map when additional distractors flanked the target. In turns, this reduced attentional guidance for larger set-sizes affected the attentional processing of the target as shown by reduced and delayed N140cc components. This increased uncertainty about the target location resulted in delayed or more variable times when attention was shifted and in an increased number of attention shifts to either side of the array. Due to the averaging process across trials, this resulted in reduced N140cc amplitudes. In other words, weaker target location activations in the preattentive map for larger set-sizes gave rise to increasingly unguided search patterns, in line with previous N2pc studies that failed to observe a reliable component when the target did not pop out from the array (e.g., Dowdall et al., 2012)[1]. Thus, the present ERP findings show inefficiencies at both the preattentive and attentive stages of processing[2].

Intriguingly, in addition to the N140cc components, an early lateralization was elicited between 70 and 100 ms over central electrodes and modulated by set-size, with decreasing amplitudes for larger set-sizes. We speculate that this early component may reflect the efficiency of the preattentive analysis, similarly to the PPC component observed in visual search tasks when a singleton item is presented amongst homogeneous distractors (e.g., Jannati et al., 2013). If this is the case, this would offer further evidence for systematic differences between set-sizes during the preattentive stage of processing, before the onset of the N140ccc. While this is an interesting possibility, it is crucial that future studies further investigate the presence and the functional meaning of this lateralized component which was observed here for the first time to our knowledge.

It is worth noting that the increased perceptual noise experienced by participants with larger set-sizes may be due - at least in part - to effects of tactile masking which are known to cause distortions in the sensory representation of the tactile stimulus (e.g., Gilson, 1969). In the present study the vibrotactile frequencies chosen for target and distractor (100 Hz and 25 Hz, respectively) were specifically aimed at activating different somatosensory channels (Pacini corpuscules, for high frequency perception, >40 Hz; Rapidly adapting fibers, Meissner corpuscules for lower frequency perception, <40 Hz, respectively; e.g., Gescheider et al., 2010) with the aim of reducing fusion effects driven by shared encoding mechanoreceptors. However, recent evidence has shown that these masking effects can also be observed across different somatosensory channels and different hands (Kuroki et al., 2017). It is therefore possible that participants experienced some masking of the vibrotactile frequencies and that this masking increased with larger distractor numbers. While this study represents a first attempt to track the physiological correlates of target selection in a tactile search task as a function of set-size, future studies should systematically aim at assessing the impact of masking on the attentional strategies adopted to select the target.

The pattern of decreasing N140cc amplitude observed in the present study differs substantially from those reported in visual search studies investigating the N2pc modulation by set-size. Mazza et al. (2009) observed that the N2pc amplitudes increased as a function of set-size regardless of whether behavioral results showed efficient or inefficient searches. It was argued that the larger N2pc amplitudes reflected the greater need for target enhancement in the presence of multiple distractors. Other studies have also reported an increase in the amplitude of the N2pc when additional distractors were present in the search-array (Luck and Hillyard, 1994; Eimer, 1996; Luck et al., 1997; Salahub and Emrich, 2018; Tay et al., 2022). However, there is also evidence that set-size does not necessarily modulate the N2pc amplitude. In an inefficient visual search task in which all items in the array were color singletons, set-size was found to modulate the *duration* of the N2pc (Christie et al., 2015). The time needed to process a subtle target feature increased when the display contained additional distractors (inefficient attentional processing). The absence of N2pc onset time modulations suggested that attention was directed to the target at the same time regardless of set-size (Christie et al., 2015). By contrast, in a different inefficient task (find an O amongst Cs) no effect of set-size was observed on the N2pc and no N2pc was elicited when all trials were averaged together (Dowdall et al., 2012). However, when these were sorted according to the N2pc latency a small N2pc emerged on faster trials (Dowdall et al., 2012). This N2pc latency variability was interpreted as evidence for a serial allocation of attention to different items of the array on different trials. Together, the studies discussed above clearly show how the attentional selection of the target is strongly dependent on the specific search task used and how similar (efficient or inefficient) patterns of behavioral results can be associated with different patterns of electrophysiological correlates of target selection.

One important difference between the N140cc and the N2pc is related to the time course of these components. While the N2pc is characterized by a sharp onset, an equally abrupt offset and

1 An inefficient attentive process should lead to inefficient behavior. In the present study, we observed increasing errors but no RT modulations with increasing set-size. We believe that the lack of an effect on RTs was caused by the brief duration of the search array (400 ms) which was necessary to limit artifacts and eye movements in the present ERP study. However, previous behavioral studies using similar tactile search tasks (based on vibrotactile frequencies) found strong RT modulations as a function of set-size with longer search array durations (2,000 ms) (c.f. Groen et al., 2008; Halfen et al., 2020). The average RT in the present study was approximately 680 ms, thus participants selected and executed the response soon after the search array offset. Because frequency discrimination is complex and unfolds over time, it is possible that they were unable to maintain a mnemonic representation of the array for long after the array offset and terminated the search soon after stimuli disappeared even when they were still uncertain about the target location.

2 Alternatively, the reduced N140cc amplitudes observed with larger set-sizes could reflect a stronger attenuation of neural activity driven by the distractors (e.g., Luck et al., 1997 for the *filtering hypothesis* of the N2pc), rather than the selection/enhancement of the target (*target processing hypothesis*, e.g., Mazza et al., 2009). However, while the filtering hypothesis can account for the results of the present study, it is not consistent with studies showing that the N140cc is not affected by increased filtering demands when a salient singleton distractor is presented next to the target (e.g., Mena et al., 2020; Gherri et al., 2021).

a short duration (which can be measured through the jackknife procedure, Ulrich and Miller, 2001), the N140cc has a very shallow onset and offset and typically lasts for several tens of millisecond. Hence, it is particularly difficult to determine the exact time course of this component. It is possible that this smeared appearance of the N140cc is due to the specific stimuli used so far in ERP studies on tactile search given that the perception of vibrotactile frequencies unfolds over time, with stimulation periods alternated by non-stimulation ones. This is likely to increase the time jitter of the attentional deployment. Given the limitations related to the time course measurement of the N140cc, it is perhaps not surprising that existing studies have often reported a link between the amplitude of the N140cc and the accuracy observed in the tactile search task (Ambron et al., 2018; Gherri et al., 2021, 2022). These findings together with the results of the present study suggest that the amplitude of the N140cc may reflect the certainty (or systematicity) with which participants deploy attention to the target location within the search-array. One question that remains unexplored is whether similar properties of the N140cc can also be observed in tactile search tasks in which the target-defining feature is temporally discrete (i.e., does not change over time, line orientation, shape, etc., instead of a vibrotactile frequency). Future studies should investigate whether target selection in touch is a process intrinsically more variable across trials or whether this depends on the specific search tasks used so far.

To conclude, the present study investigated for the first time whether the N140cc component elicited in a tactile search task is modulated by set-size. The target and the homogeneous distractors differed with respect to their vibro-tactile frequencies. Results showed that both accuracy and the N140cc amplitude decreased linearly as the items in the search-array increased. We suggest that this pattern of results reflects inefficiencies already during the preattentive analysis, caused by the increasing noise that additional distractors create in the perceptual landscape. While attention is on average directed to the target side regardless of set-size, the increased number of distractors decreases attentional guidance and the certainty with which the target can be identified, resulting in reduced N140cc components.

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 Area Vasta Emilia Nord (AVEN) Ethics Committee. The patients/participants provided their written informed consent to participate in this study.

Author contributions

EG ideated the study. SR, CI, and FF contributed to the design and implementation of the research. FF programmed the experiment and collected the data. EG and FF analyzed the data and wrote a first draft of the study. CI and SR reviewed and edited the manuscript. All authors provided critical feedback, helped shape the research and the manuscript, and read and agreed to the published version of the manuscript.

Funding

CI was supported by a grant from the University of Modena and Reggio Emilia, Fondo di Ateneo per la Ricerca–Mission Oriented (FAR 2022).

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

Ambron, E., Mas-Casadesús, A., and Gherri, E. (2018). Hand distance modulates the electrophysiological correlates of target selection during a tactile search task. *Psychophysiology* 55:e13080. doi: 10.1111/psyp.13080

Bark, K., Wheeler, J. W., Premakumar, S., and Cutkosky, M. R. (2008). "Comparison of skin stretch and vibrotactile stimulation for feedback of proprioceptive information," in *Proceedings of the symposium on haptic interfaces for virtual environment and teleoperator systems*, (Piscataway, NJ: IEEE), doi: 10.1109/haptics.2008.4479916

Christie, G. J., Livingstone, A. C., and McDonald, J. F. (2015). Searching for inefficiency in visual search. *J. Cogn. Neurosci.* 27, 46–56. doi: 10.1162/jocn_a_00716

Dowdall, J. R., Luczak, A., and Tata, M. S. (2012). Temporal variability of the N2pc during efficient and inefficient visual search. *Neuropsychologia* 50, 2442–2453. doi: 10.1016/j.neuropsychologia.2012.06.015

Duncan, J. S., and Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychol. Rev.* 96, 433–458. doi: 10.1037/0033-295x.96.3.433

Egeth, H. E., Jonides, J., and Wall, S. (1972). Parallel processing of multielement displays. *Cogn. Psychol.* 3, 674–698. doi: 10.1016/0010-0285(72)90026-6

Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalogr. Clin. Neurophysiol.* 99, 225–234. doi: 10.1016/0013-4694(96)95711-9

Eimer, M. (2014). The neural basis of attentional control in visual search. *Trends Cogn. Sci.* 18, 526–535. doi: 10.1016/j.tics.2014.05.005

Folk, C. L., and Remington, R. W. (1998). Selectivity in distraction by irrelevant featural singletons: Evidence for two forms of attentional capture. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 847–858. doi: 10.1037/0096-1523.24.3.847

Forster, B., Tziraki, M., and Jones, A. (2016). The attentive homunculus: ERP evidence for somatotopic allocation of attention in tactile search. *Neuropsychologia* 84, 158–166. doi: 10.1016/j.neuropsychologia.2016.02.009

Gaspar, J. G., and McDonald, J. F. (2014). Suppression of salient objects prevents distraction in visual search. *J. Neurosci.* 34, 5658–5666.

Gescheider, G. A., Wright, J. H., and Verrillo, R. T. (2010). *Information-processing channels in the tactile sensory system: A psychophysical and physiological analysis*. New York, NY: Psychology Press.

Gherri, E., White, F., and Ambron, E. (2022). Searching on the back: Attentional selectivity in the periphery of the tactile field. *Front. Psychol.* 13:934573. doi: 10.3389/fpsyg.2022.934573

Gherri, E., Zhao, B., and Ambron, E. (2021). Behavioural and electrophysiological evidence for the effect of target-distractor separation in a tactile search task. *Biol. Psychol.* 162:108098. doi: 10.1016/j.biopsycho.2021.108098

Gilson, R. D. (1969). Vibrotactile masking: Some spatial and temporal aspects. *Percept. Psychophys.* 5, 176–180. doi: 10.3758/bf03209553

Groen, E., Oosterbeek, M., and Toet, A. (2008). "Discrimination of concurrent vibrotactile stimuli," in *Proceedings of the international conference on human haptic sensing and touch enabled computer applications*, ed. M. Ferre (Berlin: Springer).

Halfen, E. J., Magnotti, J. F., Rahman, M. S., and Yau, J. M. (2020). Principles of tactile search over the body. *J. Neurophysiol.* 123, 1955–1968.

Hilimire, M. R., Mounts, J. R., Parks, N. A., and Corballis, P. M. (2009). Competitive interaction degrades target selection: An ERP study. *Psychophysiology* 46, 1080–1089. doi: 10.1111/j.1469-8986.2009.00846.x

Hilimire, M. R., Mounts, J. R., Parks, N. A., and Corballis, P. M. (2010). Event-related potentials dissociate effects of salience and space in biased competition for visual representation. *PLoS One* 5:e12677. doi: 10.1371/journal.pone.0012677

Jannati, A., Gaspar, J. M., and McDonald, J. J. (2013). Tracking target and distractor processing in fixed-feature visual search: Evidence from human electrophysiology. *J. Exp. Psychol. Hum. Percept. Perform.* 39, 1713. doi: 10.1037/a0032251

Johansson, R., and Vallbo, A. B. (1979). Tactile sensibility in the human hand: Relative and absolute densities of four types of mechanoreceptive units in glabrous skin. *J. Physiol.* 286, 283–300. doi: 10.1113/jphysiol.1979.sp012619

Katus, T., and Eimer, M. (2019). The N2cc component as an electrophysiological marker of space-based and feature-based attentional target selection processes in touch. *Psychophysiology* 56:e13391. doi: 10.1111/psyp.13391

Katus, T., Grubert, A., and Eimer, M. (2015). Electrophysiological evidence for a sensory recruitment model of somatosensory working memory. *Cereb. Cortex* 25, 4697–4703. doi: 10.1093/cercor/bhu153

Kiesel, A., Miller, J., Jolicæur, P., and Brisson, B. (2008). Measurement of ERP latency differences: A comparison of single-participant and jackknife-based scoring methods. *Psychophysiology* 45, 250–274. doi: 10.1111/j.1469-8986.2007.00618.x

Kuroki, S., Watanabe, J., and Nishida, S. Y. (2017). Integration of vibrotactile frequency information beyond the mechanoreceptor channel and somatotopy. *Sci. Rep.* 7, 1–13.

Lederman, S. J., and Klatzky, R. L. (1997). Relative availability of surface and object properties during early haptic processing. *J. Exp. Psychol. Hum. Percept. Perform.* 23, 1680–1707. doi: 10.1037/0096-1523.23.6.1680

Luck, S. J. (2012). "Electrophysiological correlates of the focusing of attention within complex visual scenes: N2pc and related ERP components," in *The Oxford handbook of event-related potential components*, eds E. S. Kappenman and S. J. Luck (Oxford: Oxford University Press), 329–360.

Luck, S. J., and Hillyard, S. A. (1994). Spatial filtering during visual search: Evidence from human electrophysiology. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 1000–1014. doi: 10.1037/0096-1523.20.5.1000

Luck, S. J., Girelli, M., McDermott, M. T., and Ford, M. A. (1997). Bridging the gap between monkey neurophysiology and human perception: An ambiguity resolution theory of visual selective attention. *Cogn. Psychol.* 33, 64–87. doi: 10.1006/cogp.1997.0660

Mazza, V., Turatto, M., and Caramazza, A. (2009). Attention selection, distractor suppression and N2pc. *Cortex* 45, 879–890. doi: 10.1016/j.cortex.2008.10.009

McDonald, J. J., Green, J. J., Jannati, A., and Di Lollo, V. (2013). On the electrophysiological evidence for the capture of visual attention. *J. Exp. Psychol. Hum. Percept. Perform.* 39, 849–60. doi: 10.1037/A0030510

Mena, C., Lang, K., and Gherri, E. (2020). Electrophysiological correlates of attentional selection in tactile search tasks: The impact of singleton distractors on target selection. *Psychophysiology* 57:e13592. doi: 10.1111/psyp.13592

Miller, J., Patterson, T., and Ulrich, R. (1998). Jackknife-based method for measuring LRP onset latency differences. *Psychophysiology* 35, 99–115. doi: 10.1111/1469-8986.3510099

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4

Overvliet, K. E., Smeets, J. B. J., and Brenner, E. (2007). Parallel and serial search in haptics. *Atten. Percept. Psychophys.* 69, 1059–1069. doi: 10.3758/bf03193944

Palmer, J., Ames, C. T., and Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search. *J. Exp. Psychol. Hum. Percept. Perform.* 19, 108–130. doi: 10.1037/0096-1523.19.1.108

Salahub, C., and Emrich, S. M. (2018). ERP evidence for temporal independence of set size and object updating in object substitution masking. *Atten. Percept. Psychophys.* 80, 387–401. doi: 10.3758/s13414-017-1459-6

Shaw, J. M. (1984). Correlation and coherence analysis of the EEG: A selective tutorial review. *Int. J. Psychophysiol.* 1, 255–266. doi: 10.1016/0167-8760(84)90045-x

Tay, D. Y. W., McIntyre, D., and McDonald, J. F. (2022). Searching for visual singletons without a feature to guide attention. *J. Cogn. Neurosci.* 34, 2127–2143. doi: 10.1162/jocn_a_01890

Treisman, A. (1991). Search, similarity, and integration of features between and within dimensions. *J. Exp. Psychol. Hum. Percept. Perform.* 17, 652–676. doi: 10.1037/0096-1523.17.3.652

Treisman, A., and Gelade, G. A. (1980). A feature-integration theory of attention. *Cogn. Psychol.* 12, 97–136. doi: 10.1016/0010-0285(80)90005-5

Ulrich, R., and Miller, J. (2001). Using the jackknife-based scoring method for measuring LRP onset effects in factorial designs. *Psychophysiology* 38, 816–827.

Wolfe, J. M. (1994). Visual search in continuous, naturalistic stimuli. *Vis. Res.* 34, 1187–1195. doi: 10.1016/0042-6989(94)90300-x

Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. *Psychon. Bull. Rev.* 28, 1060–1092. doi: 10.3758/s13423-020-01859-9

frontiers | Frontiers in Human Neuroscience

TYPE Original Research
PUBLISHED 24 July 2023
DOI 10.3389/fnhum.2023.1195013

OPEN ACCESS

EDITED BY
Tetsuo Kida,
Aichi Developmental Disability Center, Japan

REVIEWED BY
Michael T. Willoughby,
RTI International, United States
Sabine Doebel,
George Mason University, United States

*CORRESPONDENCE
Jessica Wise Younger
✉ jessica@youngers.org
Melina R. Uncapher
✉ muncapher@aerdf.org

RECEIVED 27 March 2023
ACCEPTED 28 June 2023
PUBLISHED 24 July 2023

CITATION
Younger JW, O'Laughlin KD, Anguera JA, Bunge SA, Ferrer EE, Hoeft F, McCandliss BD, Mishra J, Rosenberg-Lee M, Gazzaley A and Uncapher MR (2023) Better together: novel methods for measuring and modeling development of executive function diversity while accounting for unity.
Front. Hum. Neurosci. 17:1195013.
doi: 10.3389/fnhum.2023.1195013

Better together: novel methods for measuring and modeling development of executive function diversity while accounting for unity

Jessica Wise Younger[1]*, Kristine D. O'Laughlin[1], Joaquin A. Anguera[1,2], Silvia A. Bunge[3], Emilio E. Ferrer[4], Fumiko Hoeft[2,5], Bruce D. McCandliss[6], Jyoti Mishra[7,8], Miriam Rosenberg-Lee[9], Adam Gazzaley[1,10] and Melina R. Uncapher[1,11]*

[1]Neuroscape, Department of Neurology, Weill Institute for Neurosciences, University of California, San Francisco, San Francisco, CA, United States, [2]Department of Psychiatry, University of California, San Francisco, San Francisco, CA, United States, [3]Department of Psychology & Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, CA, United States, [4]Department of Psychology, University of California, Davis, Davis, CA, United States, [5]Department of Psychological Sciences and Brain Imaging Research Center (BIRC), University of Connecticut, Storrs, CT, United States, [6]Graduate School of Education, Stanford University, Stanford, CA, United States, [7]Department of Psychiatry, University of California San Diego, La Jolla, CA, United States, [8]Neural Engineering & Translation Labs, University of California San Diego, La Jolla, CA, United States, [9]Department of Psychology, Rutgers University, Newark, NJ, United States, [10]Department of Psychiatry and Physiology, University of California, San Francisco, San Francisco, CA, United States, [11]Advanced Education Research and Development Fund, Oakland, CA, United States

Introduction: Executive functions (EFs) are linked to positive outcomes across the lifespan. Yet, methodological challenges have prevented precise understanding of the developmental trajectory of their organization.

Methods: We introduce novel methods to address challenges for both measuring and modeling EFs using an accelerated longitudinal design with a large, diverse sample of students in middle childhood (N = 1,286; ages 8 to 14). We used eight adaptive assessments hypothesized to measure three EFs, working memory, context monitoring, and interference resolution. We deployed adaptive assessments to equate EF challenge across ages and a data-driven, network analytic approach to reveal the evolving diversity of EFs while simultaneously accounting for their unity.

Results and discussion: Using this methodological paradigm shift brought new precision and clarity to the development of these EFs, showing these eight tasks are organized into three stable components by age 10, but refinement of composition of these components continues through at least age 14.

KEYWORDS

executive function, in-school assessment, network modeling, middle childhood, digital assessment

1. Introduction

Executive functions (EFs) comprise a variety of cognitive abilities that enable agency over one's attention (for review see e.g., Diamond, 2013; Zelazo et al., 2016). EFs are a critical set of skills as they consistently predict positive outcomes in school and across the lifespan (Moffitt et al., 2011; Schlam et al., 2013; Pascual et al., 2019; Spiegel et al., 2021). Thus, understanding how EFs emerge and change across development is critical to understanding how we might support their growth during periods of vulnerability and opportunity (from early childhood through and into adulthood). Like most complex cognitive processes, defining and measuring EFs has been complicated, and has not yielded a consistent taxonomy of EFs (see e.g., Morra et al., 2018). While a dominant conception of how EF is organized at least in adults (Miyake et al., 2000b) proposes that EFs comprise at least three components such as holding and working with information in mind ("working memory"), the flexibility to switch between multiple tasks, goals, or rules ("cognitive flexibility"), and the attentional or inhibitory control that allows one to focus on goal-relevant information while filtering out goal-irrelevant information ("attentional control"), even in their seminal 2000 paper, Miyake and colleagues suggested these were not the only EFs. To date though, the number of EFs remains undetermined; a review of the literature found as many as 18 EFs (Packwood et al., 2011). Yet, neural data suggest EF components are more alike than different, relying on similar networks, rather than operating as distinct, independent processes (see Niendam et al., 2012 for review). Indeed, the neuroimaging field has developed sophisticated methods for interrogating the complex, dynamic relationships between components of neural systems, yet, methods for measuring and modeling such dynamic interactions using behavioral input have lagged behind. Much of the developmental literature, for example, has examined components as separate constructs. We propose that to close this gap between models of neural and behavioral data, we must build on our methodological toolkit to enable the measurement and modeling of how cognitive processes, including EFs, function in concert to achieve a specific goal (Doebel, 2020). Here we introduce two novel methods, one for measurement and one for modeling, to understand how EFs manifest over development with data we collected in a large, accelerated longitudinal study with a diverse sample of students over two years. First, we show how a novel, adaptive EF assessment battery solves previous challenges to *measuring* EFs consistently. We then *model* these data using network analytic techniques to account for what is common across EFs to reveal a clear timeline of EF development during middle childhood, a particularly understudied period.

A fundamental question yet to be fully addressed by the field is how the various components of EFs are organized across development. In other words, are the three components described above the most accurate way to parse EF in both children and adults? Second, does this taxonomy of EFs change over development, and if so, when and in what way? A key developmental theory posits that EFs begin as a unitary construct in early childhood, and the differentiation of specialized components over time is initiated by experience, to become the multi-dimensional construct observed by young adulthood (Shing et al., 2010; Mungas et al., 2013). This *differentiation hypothesis* aligns with neural developmental evidence showing increased specialization of the neural systems supporting EFs (Johnson, 2011). Findings from developmental studies using latent factor analyses have been roughly consistent with the idea of increasing differentiation of EF components from preschool through adolescence (see Lee et al., 2013 for review). During middle childhood (approximately ages 7–13), reports of the number of factors of EF have varied between one (e.g., (Shing et al., 2010; Xu et al., 2013) and four (e.g., Agostino et al., 2010). However, to date the precise understanding of when individual components begin to differentiate remains unclear. Behavioral tests of this hypothesis to-date have largely almost all taken a latent variable approach, modeling EF components as related but distinct processes and failing to adequately account for the commonalities between components. Thus, despite decades of studies, there is not yet a clearly established pattern regarding the number of distinguishable components at each age. The lack of established developmental trajectories of EFs hinders progress in understanding how specific EFs might support various health and academic outcomes, and therefore how development of these skills might be supported and when in order to benefit student outcomes. The inconsistencies in the extant literature call for a paradigm shift in approach to both the measurement and modeling of EF performance to move beyond fragmented views of EF and toward treating them as a dynamic interconnected network of skills. Next, we outline the critical factors that could comprise such an approach and offer evidence in support of the promise of such an approach.

1.1. Measuring EFs

To reveal the developmental trajectory of EFs, we first need to measure EFs with assessments that are robust across developmental stages and assessment sessions. Much of the prior cross-sectional and longitudinal work has been confounded by (a) use of the *same* tasks across age ranges, which results in floor or ceiling effects in performance if the challenge level is not adjusted appropriately, or (b) use of *different* tasks with different age groups which prevents comparisons between groups (as reviewed in e.g., Best and Miller, 2010). Not only must tasks be comparable across age and ability, but EF assessments also need to be *repeatable* over multiple timepoints so developmental progress can be measured within subjects without practice or ceiling effects. Adaptive methods that use tasks that dynamically adjust to an individual's appropriate challenge level on a trial-by-trial basis, presents a compelling and simple solution to this pernicious problem (Anguera et al., 2016b; Draheim et al., 2020). Indeed, prior work with pediatric populations suggests that highly engaging assessments with adapting challenge algorithms can reveal phenotypic differences between clinical and neurotypical populations, even when group characteristics are highly variable (Anguera et al., 2016a).

We further need *multidimensional assessments* to disentangle what EFs have in common from what they uniquely contribute to performance, to ensure each component is measured validly and reliably. Any single task used to assess a component of EF will necessarily involve processes not related to EFs (e.g., visual processing, motor response), or may be related to multiple EF components, both of which will result in measurement

impurity (Miyake et al., 2000a; Diamond, 2013). To address this impurity, researchers can collect multiple measures of each hypothesized component of EF, leveraging the commonalities across tasks to extract information about EF skills, and reducing the contribution of idiosyncratic skill related to any individual task. Thus, methods that use multiple indicators to measure each hypothesized component of EF are critical for a robust and reliable understanding of how EFs develop over time.

Lastly, to understand how EFs are expressed in real-world contexts such as school or home, recent work suggests EFs should be assessed in real-world contexts (e.g., Anderson, 2002). Indeed, one study showed that while in-school EF assessments administered in group vs. individual contexts were highly correlated, only scores from group administered assessments uniquely predicted academic achievement (Obradović et al., 2018). A related study revealed that an individual's growth in EF skills over the course of the school year can be influenced by classmates' performance (Finch et al., 2019). Thus, examining EFs in real-world educational settings (in-school, group administered contexts) provides a more ecologically valid context and is thus a necessary strategy for understanding the veridical relation between EFs and academic achievement (as reviewed in McCoy, 2019).

Here we introduce a novel assessment tool—Adaptive Cognitive Evaluation Classroom (ACE-C)—that addresses these robust measurement needs. ACE-C is based on the original ACE battery described in Anguera et al. (2016b), modified for use with children and amenable to administration in large group settings. ACE-C is a battery of assessments that taps multiple EFs through several different tasks. Importantly, each task incorporates adaptive algorithms, allowing the repeated measurement of EFs across multiple timepoints, using the same tasks in different age groups without running into floor or ceiling limitations. The incorporation of adaptive algorithms across several different tasks represents a significant advancement in assessment capabilities in two significant ways. First, this work complements prior development of EF batteries that have been used across ages (e.g., NIH Toolbox; Zelazo and Bauer, 2013; Minnesota Executive Function Scale (MEFS); Carlson and Zelazo, 2014) by bringing additional dimensionality to the assessments, allowing for examination of individual EF components across individuals. Second, building off of methods that adjust task parameters at the population-level (e.g., Davidson et al., 2006), the adaptive algorithms in ACE-C operate at an individual level. As such, no assumptions are made about the individual before interacting with ACE-C, which ensures that even individuals who perform above or below what might be expected based on demographic variables (e.g., age or grade) receive the same experience as individuals on more typical developmental trajectories. Further, this individualized adjustment is done automatically, without additional input from the experimenter, which facilitates large-group assessment even across diverse groups of individuals.

1.2. Modeling EFs

Understanding the complexity of EF developmental trajectories requires not only solving measurement challenges, but also solving concomitant modeling challenges. Historically, latent variable analysis has been the most common approach to evaluating the changing organization of EFs over development (Karr et al., 2018). With latent variable analysis, we have come to understand that across the lifespan, while EFs diversify over development, they do not become completely distinct. Indeed, both behavioral and neural examinations of EFs have demonstrated the existence of a unifying umbrella construct termed "Common EF" through adulthood (Friedman et al., 2008; Reineberg et al., 2015; Friedman and Miyake, 2017; Smolker et al., 2018). Notably, cross-sectional examinations of middle childhood and adolescence using latent variable models also support the inclusion of a Common EF component (Engelhardt et al., 2015; Hatoum et al., 2020) as well. However, including such a factor to test the differentiation hypothesis and assess the dynamic development of EFs longitudinally poses severe methodological challenges.

While Common EF can be modeled from the confirmatory approach by incorporating it as a higher-order umbrella component capturing what is common among all lower-order components, such an approach is not amenable to testing the differentiation hypothesis. Models with a higher-order component would require at least three lower-order components of EF to be properly identified (Kline, 2011) and provide meaningful insight into the patterns of the behaviors being modeled. Thus a model with only one or two components differentiated from Common EF is not identified, and the earliest stages of differentiation cannot be examined. An alternative modeling approach is to incorporate Common EF not in a hierarchical fashion, but as an additional lower-order latent variable. In such a model, each observed variable measures two latent variables, Common EF and another differentiated component (a "bifactor" model). While these models can be easier to identify in some instances, it can be difficult to get such complex models to converge given the historically low power and task reliability observed in extant examinations of EF (Karr et al., 2018).

Further, confirmatory latent analyses provide limited information as to how the cognitive mechanisms supporting EF task performance may evolve over development (e.g., whether a task may index different EF components at different developmental stages). While model fit statistics can indicate whether a hypothesized organization fits the observed data well, they provide limited information on how to improve that model. For example, while one hypothesized organization might fit the data well, there could be other organizations that fit the data better that simply go untested. Additionally, methods for statistically comparing alternate hypotheses regarding which component a task measures are not straightforward. As such, alternative hypotheses around the EFs involved in different tasks are unlikely to be developed from the results of confirmatory latent analyses.

To advance our understanding of how EFs evolve over development, we need a method that (a) allows for task performance to reflect different EFs at different developmental stages, and (b) accounts for the high degree of association common to all EF tasks. Exploratory latent variable models like exploratory factor analysis (EFA) meet the first requirement but fail to account for Common EF. Conversely, confirmatory approaches such as confirmatory factor analysis (CFA) and bifactor modeling can account for commonalities among EF tasks but do not allow for task reorganization in a data-driven way. Indeed, recent evidence has suggested latent variable analysis may not be an

appropriate representation of EFs (Camerota et al., 2020). To build on the insights gained from latent variable modeling, we suggest leveraging a powerful family of techniques that provides a data-driven method for identifying unique and communal cognitive mechanisms: network analysis. Network analysis is an approach gaining traction in the psychometric field for understanding cognitive constructs comprising complex inter-related components such as intelligence, psychopathology, and personality (Borsboom and Cramer, 2013; Costantini et al., 2015; Kan et al., 2019). In network analysis, relationships between variables can be determined after accounting for what is common among all variables by examining partial correlations. Thus, through network analysis, we can understand how EF behaviors are related after what is common among all variables, including what can be attributed to Common EF, is accounted for. Further, in contrast to latent variable analysis, in which observed variables are related through the modeled unobserved latent variables, the relationships between observed variables is direct. As such, performance on one task can affect performance on any other, not just those tasks theorized to measure the same construct. Finally, after determining how each variable is related to another, we can assess which variables likely reflect the same cognitive construct by applying community detection algorithms. This data-driven approach groups together variables that are more strongly related to each other than other variables in the network, allowing us to establish a theory-agnostic organization of EFs. In this way, network analysis allows us to examine the structure of EF from a holistic perspective and arrive at the organization that best reflects the data without testing and comparing multiple competing models.

1.3. Current study

Here, we capitalize on the improved interpretability of longitudinal and cross-sectional comparisons afforded by using the same tasks across all participants (Best and Miller, 2010) with our ACE-C battery to shine light on the relatively understudied period of middle childhood (~7–12 years old), the developmental stage in which EFs may undergo the most rapid organizational development (Romine and Reynolds, 2005; Best et al., 2011). We demonstrate how network analysis can advance our understanding of the organization of three hypothesized EFs across development by first testing the differentiation hypothesis with the latent variable analysis approach and then highlighting the additional insights gained by using a network analysis approach. Specifically, we use each method to determine not only when the investigated EFs become distinct from one another but, critically, when they become distinct from the unifying Common EF component. Finally, we leverage information generated from network analyses to gain insights into the stability of the organization of these EFs across time. We show that during middle childhood, organization of these EFs begins to stabilize, yet continues to develop in a manner suggesting EFs need continued support throughout their protracted development as children transition to adolescence. Developmental insights revealed by network analyses extend those from latent variable analyses and, in line with work by Camerota et al. (2020), show how differing modeling methods can result in different conclusions regarding the number of components identified across development to date. The novel findings from network analysis lay the groundwork for new avenues of investigation to understand how to best support EFs across the lifespan.

2. Materials and methods

Participants in this study were recruited through their schools as part of Project iLEAD (in-school longitudinal executive function and academic achievement database), a two-year accelerated longitudinal study of EF development in students grades 3–8. Full details of Project iLEAD are reported in Younger et al. (2022).

The study was approved by the University of California San Francisco Institutional Review Board and conducted in accordance with the relevant guidelines and regulations. Written parental or guardian consent was obtained from all participants at the beginning of the study, and verbal assent from all participants was obtained before all in-class data collection sessions. At the end of the study, all students in participating classrooms received snacks and stickers, regardless of participation.

2.1. Participants

Nine schools (seven public, one independent, and one parochial) from northern California opted to participate in this longitudinal study, which included assessments at the Fall and Spring of two academic years for a total of four assessment periods. Two of the five public elementary schools and one of the two public middle schools were Title I schools. In total, 1,280 students participated over the course of two years. At the beginning of each school year, teachers distributed consent forms to students to take home for parental or guardian review and signature. This first round of recruitment resulted in a total of 1,088 participating students in Year 1: 284 3rd graders (M = 8.07 years old, SD = 0.35), 260 5th graders (M = 9.98 years old, SD = 0.41), and 544 7th graders (M = 11.9 years old, SD = 0.47). In the fall of Year 2, we re-opened enrollment for participating classrooms to allow new students to participate in the study, which resulted in an additional 195 students joining the study (44 4th, 147 6th, and 4 8th grade students). The Year 2 sample thus included 1,106 students: 288 4th graders (M = 9.03 years old, SD = 0.33), 336 6th graders (M = 10.9 years old, SD = 0.39), and 482 8th graders (M = 12.9 years old, SD = 0.44). For patterns of missing data across timepoints, see **Supplementary Figure 1**. Our sample was demographically diverse. Ethnically, our sample was 34% Asian, 26% Hispanic/Latinx, and 16% White. Further, 10% of the sample received Special Education services, 32% qualified as low income, and 14% were currently enrolled in English Language classes, with another 29% having previously been enrolled in English Language classes, but now considered fluent in English. See **Table 1** for additional demographics of participating students.

2.2. Procedures

We administered a series of mobile assessments of EF, math, and reading skills that took the form of digital "games", during school hours, at the beginning and end of each academic year

TABLE 1 Student demographics.

		Timepoint 1		Timepoint 2		Timepoint 3		Timepoint 4	
		n	%	*n*	%	*n*	%	*n*	%
Gender	Male	519	50.9	514	51.8	541	51.3	510	51.2
	Female	500	49.1	478	48.2	513	48.7	487	48.8
Ethnicity	American Indian or Native Alaskan	5	0.5	6	0.6	6	0.6	5	0.5
	Asian	339	33.3	332	33.5	369	35.0	350	35.1
	Black or African American	20	2.0	18	1.8	18	1.7	17	1.7
	Filipino	55	5.4	56	5.7	60	5.7	56	5.6
	Hispanic or Latinx	267	26.2	253	25.5	280	26.6	269	27.0
	Pacific Islander	7	0.7	6	0.6	5	0.5	6	0.6
	Two or more ethnicities	43	4.2	45	4.5	44	4.2	44	4.4
	White	170	16.7	165	16.6	185	17.6	169	17.0
	Unknown	113	11.1	111	11.2	87	8.26	81	8.1
Special education status	No	806	79.1	771	77.7	855	81.1	807	80.9
	Yes	101	9.9	111	11.2	113	10.7	111	11.1
	Unknown	112	11.0	110	11.1	86	8.2	79	7.9
Low income qualification	No	583	57.2	570	57.5	616	58.4	587	58.9
	Yes	324	31.8	312	31.5	352	33.4	331	33.2
	Unknown	112	11.0	110	11.1	86	8.2	79	7.9
English language fluency	English monolingual	396	38.9	391	39.4	434	41.2	413	41.4
	English multilingual, never enrolled in English classes	56	5.5	58	5.9	67	6.4	63	6.3
	English multilingual, previously enrolled in English classes	308	38.9	289	29.1	324	30.7	311	31.2
	Current English Language Learner	147	14.4	144	14.5	143	13.6	131	13.1
	Unknown	112	11.0	110	11.1	86	8.2	79	7.9

(fall and spring) over two school years. At each of the four timepoints, EF assessments were administered during one class period (approximately 50 min), with the research team returning a little over a month later to administer the math and reading assessments (M = 5.7 weeks, SD = 2.4, min. = 1.9, max. = 10). At the end of each academic year, academic performance and other relevant data were provided by the district for students whose parents consented to share district data.

2.2.1. Adaptive cognitive evaluation classroom (ACE-C)

This study used a novel mobile assessment battery, ACE-C, to assess EF skills. The original ACE battery was developed from cognitive assessments commonly used in lab-based settings and modified for real-world settings by including adaptive, psychometric staircase algorithms, highly motivating trial-wise and end-of-task feedback (Anguera et al., 2016b). ACE-C is an adaptation of this battery to include a child-friendly interface and additional instructional design to facilitate group-administration. Importantly, the adaptive algorithms enabled two critical affordances: (a) the same tasks could be used with the same students across multiple timepoints to reveal a student's changing cognitive abilities without being confounded by ceiling or floor effects or reduced motivation due to multiple assessments, and (b) the same tasks could be used across students of diverse ages to reveal individual differences in cognitive abilities across development without the confound of different tasks (Anguera et al., 2016b). This advancement in our approach to assessment enabled robust integrative data analytics within-subjects over time, and across-subjects from a wide age range without any a *priori* assumptions about any individual participant's abilities, for example, according to their age.

The assessment battery consisted of a color blindness test, a response time control task, two working memory tasks, one attentional filtering task, three context monitoring tasks, three interference resolution tasks, and one cognitive flexibility task. The attentional filtering task was excluded from the current analysis due to differential task challenge across grade levels, while the cognitive flexibility task was excluded due to technical errors that prevented consistent data reporting across timepoints. All other tasks are briefly described below along with the *a priori* defined metric of interest selected based on the psychometrics of each task. Full descriptions of each task are included in the

Supplementary material. Example stimuli and schematics for tasks are presented in **Supplementary Figure 2**.

2.2.1.1. Response time control task

The first ACE-C task was a measure of basic response time (*BRT*; **Supplementary Figure 2A**). Because improvements in EFs have also been associated with improvements in general processing speed (e.g., Fry and Hale, 1996), *BRT* was designed to serve as a covariate to be regressed from performance metrics of all other ACE-C tasks. By using BRT as a control metric, analyses were kept consistent across tasks and task-specific control metrics were not required. Mean response time (RT) collapsed across both dominant and non-dominant hands was the metric of interest for this task.

2.2.1.2. Color blindness test

The second ACE-C task was a screening assessment for red-green color blindness (Ishihara, 1972; **Supplementary Figure 2B**). We assessed whether students selected one or more responses indicating red-green color blindness according to scoring guidelines in Ishihara (1972).

2.2.1.3. Working memory

Two tasks were used to measure working memory (WM), *Forward Spatial Span* (**Supplementary Figure 2C**) and *Backward Spatial Span* (**Supplementary Figure 2D**). These two tasks were digital modifications based on the Corsi Block Task (Corsi, 1973). In this task, students were shown an array of open circles, with a target sequence cued via circles becoming filled, in sequence, with either green (*Forward Spatial Span*) or blue (*Backward Spatial Span*) color. Once students viewed the cued sequence, they were instructed to recreate the sequence in the same order (*Forward Spatial Span*) or in the reverse order (*Backward Spatial Span*). Sequence length increased according to performance. The metric of interest for both tasks was span length, or the maximum number of spatial locations attempted to be held in mind in the correct sequence.

2.2.1.4. Context monitoring

Context monitoring (CM) was measured with three tasks: Sustained and Impulsive Attention (both tasks administered within a single test called *Continuous Performance Task* [*CPT*]; **Supplementary Figure 2E**) and *Tap and Trace* (**Supplementary Figure 2F**). For all three tasks, students were instructed to respond to a target stimulus and withhold a response to non-target stimuli. *CPT* is a target detection test adapted from the Test of Variables of Attention (TOVA; Greenberg et al., 1991). This test included two tasks: a target frequent task (80% target trials) designed to assess impulse control *(Impulsive Attention)* and a target infrequent task (20% target trials) designed to test sustained attention abilities *(Sustained Attention)*. For *Sustained Attention,* we used a metric that is sensitive to lapses in attention—the standard deviation of the RT to infrequently presented targets (Leark et al., 2018). For *Impulsive Attention*, we used a metric that would measure detection of targets while accounting for withholding prepotent responses to frequent non-targets–the signal detection metric of *d'*. *Tap and Trace* is a dual-task assessment adapted from the paradigm described by Eversheim and Bock (2001). This task included three blocked conditions: one in which students used their dominant hand to tap when they detected a target stimulus, a second in which they traced a shape with their non-dominant hand, and a third in which they performed both tasks simultaneously. To differentiate this task from the CPT and better address task impurity concerns by assessing context monitoring when EFs are challenged by divided attention, we included performance only on the dual-task block. For this task, the metric of interest was how reliably students could detect a target vs. a distractor during the dual-task portion of the task; thus, we again deployed *d'*.

2.2.1.5. Interference resolution

Interference resolution (IR) was measured with three tasks: *Stroop* (**Supplementary Figure 2G**), *Flanker* (**Supplementary Figure 2H**), and *Boxed* (**Supplementary Figure 2I**). *Stroop* is based on the computerized version of the color-word *Stroop* task as described by Mead et al. (2002) in which students selected the text color (e.g., green) of a centrally presented color word (e.g., BLUE). On 30% of trials, the text and word were incongruent, and on 70% of trials they were congruent. *Flanker* is a letter flanker task based on the paradigm described by Eriksen and Eriksen (1974) in which students are instructed to indicate the middle letter of a string of five letters. On 50% of trials, the central and flanking letters were congruent, and on 50% of trials they were incongruent. Finally, *Boxed* is a top-down/bottom-up attention task based on the visual search paradigm first described by Treisman and Gelade (1980) in which students must identify a target stimulus in an array of distractor stimuli. This task included four blocked conditions that varied on search condition and number of distractors. In each condition, the target was either identifiable by one feature (color) or by the conjunction of two features (color of target and location of opening of the target box) and either a low (3) or high (11) number of distractor stimuli. For tasks in which students were expected to respond to each trial, we used Rate Correct Score (RCS) to index performance on both RT and accuracy. Task-level RCS was calculated by dividing the number of correct responses by the product of mean RT for all trials and the total number of trials responded to Woltz and Was (2006), Vandierendonck (2017) across all conditions. To achieve a high RCS, participants must perform quickly and accurately across all trials, regardless of condition. This approach thus takes into account how participants perform on both congruent and incongruent conditions without introducing reliability issues frequently cited when using more traditional subtraction methods (Enkavi et al., 2019). RCS was used for *Stroop* and *Flanker*, however, a technical error in *Boxed* prevented RCS from being calculated in the same manner as the other tasks. Instead, we used mean RT to all correct trials for *Boxed*. The grouping of tasks into these three components differs slightly from some extant literature in an effort to bring greater precision to the EFs measured. For extended discussion on the battery design and component grouping, see the "4. Discussion" section.

2.3. Analysis methods

2.3.1. Data cleaning procedures

A very small number of students who were red-green colorblind as indicated by the colorblind screener ($n = 16$) were excluded from analysis, given that several tasks required students to discriminate between colors. Trials with no response when a response was

expected and anticipatory trials (RT < 200 ms) were excluded from analyses (1.8% of all trials).

Data from each student were evaluated and cleaned on a task-level basis at each timepoint. In this way, participants were not wholly excluded from analysis, but only task data for which we could not be confident that the participant understood or complied with the task instructions were excluded. For each task, to be included in data analysis, students must have answered a minimum of five trials per condition and achieved above-chance accuracy on the easiest condition (i.e., the condition that required lowest cognitive load). Data from each task were then evaluated for outlier students based on performance within each cohort and timepoint. Outlier performance was defined as performance falling outside three median absolute deviations (MADs) of the median performance of the relevant cohort at a given timepoint (Leys et al., 2013). Finally, additional outlier analyses to identify influential observations in the larger regression analysis of task performance were conducted by computing Cook's distance. Observations with Cook's d > 1 were removed. These cleaning procedures resulted in exclusion of 1.9% of task-level data collected. See **Supplementary Table 1** for *N* datasets excluded per task per cleaning step. For patterns of missing data across timepoints, see **Supplementary Figure 2**.

2.3.2. Effects of age and time on task performance

For each task's metric of interest, we sought to understand the developmental trajectory of performance across different age ranges. We used linear mixed effects models to examine how an individual student's performance over time may differ depending on age after controlling for multiple demographic variables. To index the variable of time more precisely, it was coded as the number of months since last assessment. In this way, the first instance of a participant's engagement with ACE-C was always indexed as 0, regardless of whether that occurred during the first timepoint of the study Fall 2016 or later (due to later enrollment, absence on data collection day, etc.). Age was indexed as participant age in months at the time of assessment. Control variables in these models included mean RT on the BRT task (continuous) as an indicator of general processing speed, cohort (3 categories: 3rd–4th grade, 5th–6th grade, 7th–8th grade), and gender (2 categories; male, female). Random effects included school (9 categories), the random intercept of participant, and the random slope of time. Models were run using the "lme4" package in R (Bates et al., 2014) and significance of each variable was evaluated using Satterthwaite's degree of freedom method as implemented in the "lmerTest" package in R (Kuznetsova et al., 2017).

2.3.3. Confirmatory factor analysis

We conducted latent variable modeling using confirmatory factor analysis (CFA). We chose to use CFA over exploratory factor analysis (EFA) because, while data-driven organizations of variables are possible using EFA, exploratory approaches do not provide a straightforward way to account for the high degree of overlap between performance on EF tasks, and assignment of a behavior to a latent variable is dubious, often resulting in uninterpretable organizations (Brocki and Bohlin, 2004). We conducted separate CFAs for the three cohorts at the four timepoints to avoid assuming the structure of EF remained the same across timepoints and to assess the stability of these structures over a two-year measurement period. We evaluated five models of EF: the maximally differentiated structure with three distinct factors, all possible permutations of a two-factor model in which two of the three factors are collapsed into one, and the simplest structure in which all tasks represent a single, undifferentiated EF factor (see **Supplementary Figure 3**). Although the longitudinal stability of models can be tested with a CFA approach, such statistical tests for longitudinal network analysis have not yet been developed. To keep the results of the two modeling techniques comparable, we do not account for the dependencies in observations across timepoints.

After assessing covariance coverage to ensure sufficient available data for all tasks across all cohorts and timepoints, all CFAs were conducted in Mplus version 8.1 (Muthén and Muthén, 2017) with the robust maximum likelihood estimation method. To statistically compare nested models, we used Satorra-Bentler scaled chi-square tests with degrees of freedom equal to the difference in number of free parameters between the comparison and nested models (Satorra and Bentler, 2010). These tests help us to determine whether more complex representations of EF are a better fit to EF task performance across middle childhood. Because these statistics are meant to compare nested models, the 1-factor model was compared to each of the 2-factor models, and each of the 2-factor models were compared to the 3-factor model, but the 2-factor models cannot be statistically compared to each other in this manner. In interpreting these results, we took a conservative approach in which a more complex model would be selected over a less complex model only if a more complex model would always be preferred, regardless of which 2-factor permutation was considered. The results of chi-square difference testing were corroborated by converging evidence from the Comparative Fit Index (CFI), root mean square error of approximation (RMSEA), Akaike Information Criteria (AIC), and sample-size adjusted Bayesian Information Criterion (BICc). CFI values > 0.90 were considered excellent model fit, with values closer to 1 indicating better model fit. RMSEA values less than or equal to 0.06 were considered adequate model fit (Hu and Bentler, 1999), with lower values indicative of better model fit.

Models explicitly incorporating a Common EF factor were not tested here, as models in which Common EF is a higher-order factor are not amenable to testing the differentiation hypothesis. While Common EF could be incorporated as a higher-order umbrella component reflecting what is common among all lower-order components, structures with any fewer than three differentiated components would not be considered properly identified (i.e., it would not be possible to uniquely estimate each component's association with Common EF). While it is possible to test the differentiation hypothesis with an alternative approach incorporating Common EF as an additional lower-order latent variable, taking such an approach was not possible with our dataset. In such "bifactor" models, each observed variable measures two latent variables: Common EF and another differentiated component. Such a model would not be identified for this dataset without assuming performance on the WM tasks contributes equally to both the WM and Common EF factors (see Limitations), which has not been supported in the literature (Friedman et al., 2008, 2011).

2.3.4. Network analysis

Replicating the general approach used for the latent variable models, we created separate models of EF performance for each cohort and timepoint. All network analyses were conducted in R 4.1.2 (R Core Team, 2020). Network models were estimated using the bootnet package (Epskamp, 2015). All models were fully saturated partial correlation networks (non-regularized Gaussian Markov random fields), and missing data were handled via full information maximum likelihood. After estimating each network model, the Spinglass algorithm (Reichardt and Bornholdt, 2006) from the igraph package (Csardi and Nepusz, 2006) was applied separately to each network to determine communities of tasks. We employed the Spinglass algorithm rather than other community detection algorithms, such as the Louvain algorithm, because it can handle negative partial correlations in a network. To ensure the stability of groupings, community detection was performed 1,000 times and the most frequent grouping is reported here. Resulting network and community detection results were displayed graphically using the qgraph package (Epskamp et al., 2012). For graphing purposes, nodes were fixed to the same positions across networks and partial correlations between -0.1 and 0.1 are not displayed. To understand network stability over time, edge weights from each network were correlated with each other. Because these edge weights represent partial correlations, edge weights were first Fischer transformed before computing correlations between networks.

3. Results

We first show how the use of novel, adaptive assessments can robustly measure EFs longitudinally across a wide age range without floor and ceiling effects. We then demonstrate how a holistic modeling approach that accounts for Common EF can enhance our current understanding of the emergence and development of EFs by testing the differentiation hypothesis using two analytic approaches, latent variable analysis and network analysis. Using a latent variable approach, we replicate the ambiguous, difficult to interpret results found in prior investigations. We then critically extend our understanding using a network analytic approach, revealing developmental insights missed under the latent variable approach that could not appropriately take into account Common EF.

3.1. Novel EF measurement

To examine the utility of our novel adaptive assessment, we performed two analyses, one to assess task performance, and another to assess challenge level. We had different predictions for each analysis. We predicted the adaptive response window would equate task challenge level across cohorts and timepoints as supported by similar percent of responses for which participants received "correct" feedback across cohorts and timepoints. However, we expected that task performance as measured by the metric of interest for each task noted above, which did not take into account whether the response was within the adaptive response window and may have included other aspects of performance such as response time (e.g., RCS, standard deviation of response time, d', etc.), would show traditional developmental improvements in performance over time.

To confirm the effectiveness of the adaptive response window across tasks, we examined percent of responses with "correct" feedback only. In tasks with an adaptive response window (Impulsive Attention, Sustained Attention, Tap and Trace, Stroop, Flanker, and Boxed), participants only received "correct" feedback if they provided the correct answer within a limited time frame. All other responses resulted in feedback indicating the response was correct but "late" or "incorrect". This adaptive algorithm was designed to produce ∼75% of responses resulting in "correct" feedback for all participants and while this target accuracy was not achieved across all tasks, it was confirmed in practice to produce an average of 72.04% across tasks. Additionally, the adaptive algorithm did not completely eliminate developmental effects; while linear models examining the effect of cohort and time on percentage of trials with "correct" feedback did show significant differences between cohort and timepoint However, the significance of these effects is likely driven by the large sample size used in the current study; model effect sizes were small, accounting for less than 20% of the variance across all tasks (average R^2 = 0.10 see **Table 2** and **Supplementary Figure 5**). Together, these results suggest the adaptive tasks successfully presented a similar challenge across ages and measurement occasions.

We next examined the potential developmental effect on task performance as measured by the task-specific metric of interest described above which did not take into account whether the response was within the adaptive response window and may have included other aspects of performance such as response time (e.g., RCS, standard deviation of response time, d', etc.). We found each adaptive EF assessment captured predicted developmental improvements in performance. Linear mixed effects models examining task performance for each metric of interest allowing random effects for participant, school, and time, showed that, across tasks, performance significantly improved with age and time after controlling for BRT, cohort, and gender except the two span tasks. Both Forward and Backward Spatial Span showed significant effects of time, but only trended towards main effects of age, possibly due to the ordinal nature of the metric of task performance for these tasks which leaves little room for variation. Plots of raw scores not accounting for these control variables are shown in **Figure 1** and effect sizes of each control variable are shown in **Supplementary Figure 4**. For between task correlations as well as the mean and standard deviation of task performance after accounting for BRT for each cohort and timepoint (the metric used in modeling analyses), see **Supplementary Tables 2–4**. Beyond these predicted EF performance improvements with age, performance on all but two tasks (Tap and Trace and Backwards Spatial Span) showed a significant interaction between age and time, suggesting that younger participants tended to improve more over time compared to older participants. Across tasks, the two control variables that most frequently had a significant effect on performance were BRT and gender. The consistently strong effect of BRT on all tasks was expected as this variable was included to capture potential differences in an individual's pattern of responses, which might also capture such variance due to familiarity with responding on a touch-screen device, etc. Further, for all but Sustained Attention and Forward Spatial Span, there was

TABLE 2 Mean (standard deviation) percent of responses that received correct for tasks with an adaptive response window and variance explained by model of cohort and time.

	3rd–4th grade cohort				5th–6th grade cohort				7th–8th grade cohort				Model F (NumDF, DenDF)	Model R^2	95% confidence interval
Timepoint	1	2	3	4	1	2	3	4	1	2	3	4			
Sustained attention	87.4 (6.6)	87.7 (5.6)	88.0 (6.8)	88.6 (5.2)	89.9 (3.2)	89.5 (3.3)	90.1 (4.3)	89.4 (5.2)	90.4 (3.5)	90.4 (3.1)	90.4 (5.7)	89.4 (7.0)	44.75 (3, 3918)	0.033	0.023, 0.045
Impulsive attention	73.9 (4.3)	74.2 (5.2)	75.1 (5.0)	75.0 (4.9)	76.0 (4.8)	75.9 (4.2)	76.0 (4.2)	76.2 (4.0)	76.9 (4.0)	77.0 (3.9)	76.8 (4.3)	76.6 (4.4)	78.06 (3,3954)	0.056	0.043, 0.071
Tap & trace	68.6 (8.3)	71.7 (8.4)	72.7 (8.7)	74.0 (9.0)	74.2 (7.4)	75.2 (7.7)	75.5 (8.4)	77.2 (8.2)	78.3 (6.0)	78.3 (8.1)	78.0 (8.0)	78.8 (6.6)	153.38 (3, 3049)	0.131	0.111, 0.154
Stroop	72.0 (6.7)	61.3 (9.5)	61.9 (9.9)	60.9 (9.7)	72.7 (5.3)	59.4 (9.3)	59.3 (9.7)	58.6 (9.5)	71.3 (5.3)	57.1 (10.0)	56.6 (9.4)	58.1 (9.0)	304.90 (3,3800)	0.194	0.173, 0.216
Flanker	66.8 (7.2)	64.1 (7.6)	63.9 (6.8)	63.3 (6.8)	68.7 (6.8)	63.3 (6.5)	63.7 (7.0)	62.3 (6.0)	69.3 (5.2)	62.2 (7.4)	62.2 (7.0)	62.3 (7.3)	111.86 (3, 3267)	0.093	0.076, 0.113
Boxed	62.2 (4.4)	63.0 (4.6)	64.0 (4.3)	64.3 (3.8)	64.3 (4.0)	65.4 (3.5)	64.9 (3.9)	65.7 (3.8)	65.2 (4.2)	65.6 (3.9)	66.2 (3.7)	66.1 (3.7)	95.51 (3, 3755)	0.071	0.057, 0.088

Data are collapsed across the four timepoints. NumDF, numerator degrees of freedom; DenDF, denominator degrees of freedom.

a significant effect of gender, with students self-identifying as female showing better task performance compared to those identifying as male.

3.2. Novel EF modeling

After solving for persistent challenges to measuring EFs through the use of our novel tool, ACE-C, we demonstrate how network analysis can build on the findings from latent variable analysis and generate new hypotheses regarding the organization of EFs by accounting for what is common between EF components.

3.2.1. Latent variable analysis

To directly test the differentiation hypothesis using latent variable modeling, we compared a series of models to establish the number of distinguishable EF components at each stage of development using CFA. In accordance with the differentiation hypothesis, we expected more complex models with more unique factors would provide better model fit for older students. Based on prior adult literature and the tasks used in the current study, the number of components could range from one to three, with the maximally-differentiated organization of EFs representing WM, IR, and CM grouping components. As noted in the methods, we did not explicitly incorporate Common EF into these models and instead examined correlations between factors to assess when these components could be differentiated beyond the unifying Common EF factor. Correlations greater than 0.70 between factors indicate that components represent redundant information (sharing more than 49% of variance) and are therefore likely not fully differentiated from one another.

Overall, the latent variable approach revealed an indeterminate developmental progression of differentiation of EF components. Model fit statistics (**Supplementary Table 5**) tended to indicate a 2-factor model was the best fitting model for the 3rd–4th grade cohort at all timepoints, though a different 2-factor model was the best fitting at each timepoint. At timepoint 1, the model with WM as distinct fit best, the model with IR as distinct fit best at timepoints 2 and 3, and the model with CM as distinct fit best at timepoint 4. However, it should be noted that at timepoints 1, 2, and 4, a 3-factor solution had similar fit statistics to these 2-factor solutions. Fit statistics were similarly mixed at timepoint 1 for the 5th–6th grade cohort, with both the 2-factor model in which WM is distinct and the 3-factor model showing best fit statistics. After timepoint 1 though, fit statistics pointed towards the 3-factor solution being the best fit through timepoint 4 for the 7th–8th grade cohort. However, statistical comparisons of the models indicated that while more complex models may have better fit indices, they may not be necessary to model the data well.

Generally, results of statistical comparisons (**Table 3**) suggest that a single component best describes the organization of EF from 3rd through 4th grade, after which at least three distinct EF components can be identified. However, this pattern is not unequivocal, and many open questions remain. Within the 3rd–4th grade cohort, at least two out of three 2-factor models did not provide significantly better model fit than a 1-factor model with the exception of timepoint 2. At this timepoint, even the 3-factor model provided better fit than all but the 2-factor model in which IR is

10.3389/fnhum.2023.1195013

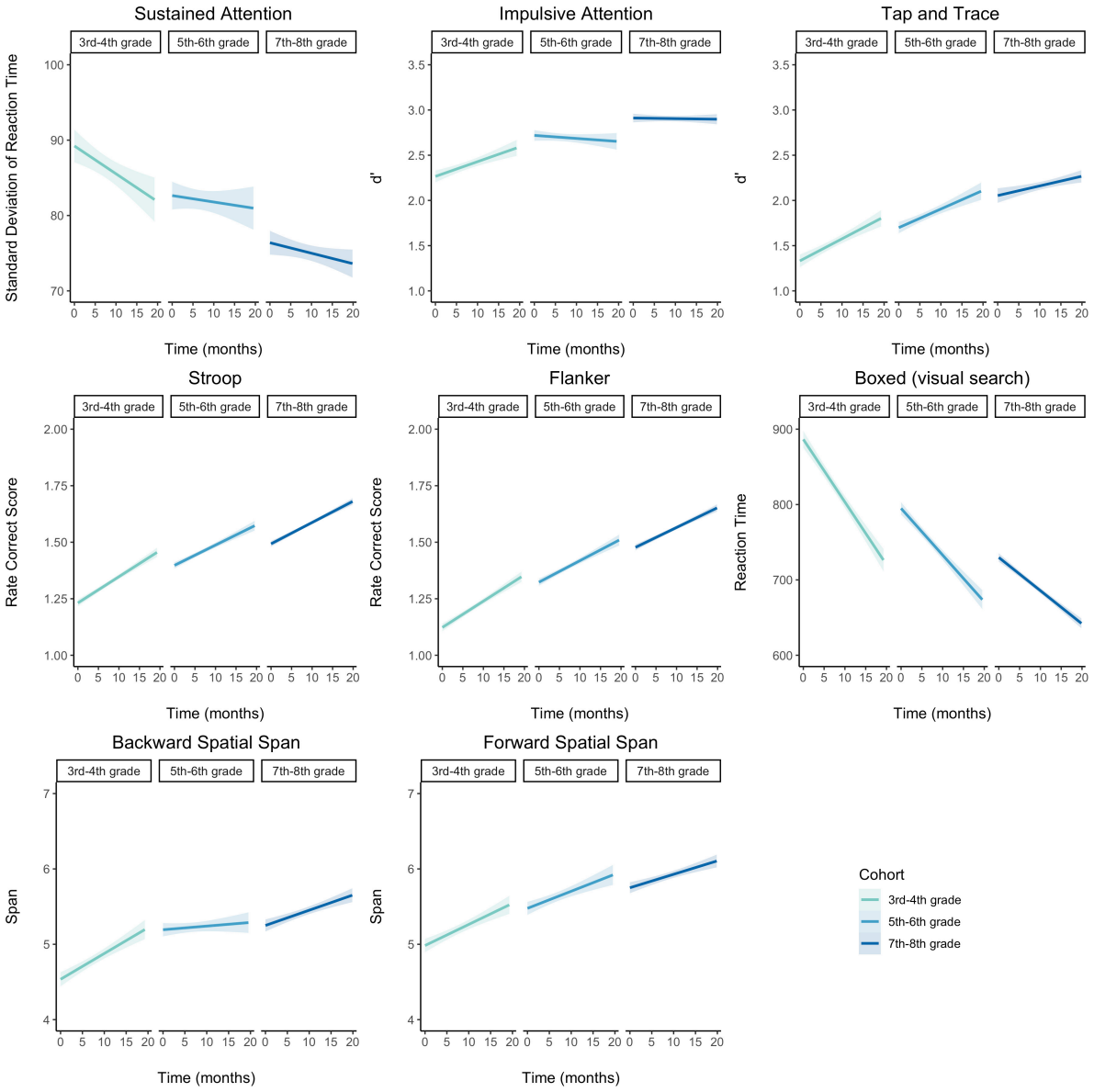


FIGURE 1
Growth in performance on executive function metrics of interest for each task and cohort. With few exceptions, all participants improved over time, and younger students tended to show the most gains over time, as indicated by significant main effects. Shaded region represents 95% confidence interval of linear regression of time on performance.

distinct. Yet, at timepoint 3, a more complex model never provided significantly better fit compared to a single component, leaving the developmental trajectory unclear. Further, at the first timepoint for both the 5th–6th and 7th–8th grade cohort, the 2-factor model combining CM and IR fit significantly better than a single-factor model, but other potential 2-factor configurations did not fit the data better than models with a single component. Additionally, the 3-factor model did not fit better than the WM-distinct 2-factor model, indicating EFs may not be well-differentiated at timepoint 1 for any age group. Moreover, alternative hypotheses around the EFs involved in different tasks are unlikely to be developed from these results. Different structures from those tested here may fit the data better (e.g., a task may index a different EF component at different developmental stages), but methods for statistically comparing such alternate hypotheses regarding which EF component a task draws on are not straightforward and would not be feasible to test without additional theoretical guidance.

Finally, the degree of differentiation of these factors from Common EF was unclear; factor correlations for structures in which a 3-factor solution was selected suggest WM differentiates by 5th grade ($M_{WMandCM}$ = 0.40; $M_{WMandIR}$ = 0.54), however, a persistent high degree of overlap between CM and IR ($M_{IRandCM}$ = 0.69) across cohorts leaves open the question of whether one or both of these components would be distinguishable from Common EF (see **Supplementary Table 6–10** for full list of factor loadings and correlations). Without statistical methods to determine when components become distinct from both other EFs and Common EF, the use of latent variable models to

TABLE 3 Satorra–Bentler scaled χ^2 tests comparing 1-, 2-, and 3-factor models of executive function.

			1- versus 2-Factor (IR with CM)		1- versus 2-Factor (WM with CM)		1- versus 2-Factor (WM with IR)		2- (IR with CM) versus 3-Factor		2- (WM with CM) versus 3-Factor		2- (WM with IR) versus 3-Factor	
Cohort	Time-point	*n*	$\Delta\chi^2(\Delta df)$	*p*	$\Delta\chi^2(\Delta df)$	*p*	$\Delta\chi^2(\Delta df)$	*p*	$\Delta\chi^2(\Delta df)$	*p*	$\Delta\chi^2(\Delta df)$	*p*	$\Delta\chi^2(\Delta df)$	*p*
3rd-4th grade cohort	1	210	10.602 (1)	**0.001**	0.157 (1)	0.692	0.225 (1)	0.635	0.277 (2)	0.871	10.642 (2)	**0.005**	11.199 (2)	**0.004**
	2	209	4.539 (1)	**0.033**	28.671 (1)	<**0.001**	20.036 (1)	<**0.001**	21.504 (2)	<**0.001**	3.486 (2)	0.175	12.052 (2)	**0.002**
	3	217	0.002 (1)	0.962	2.975 (1)	0.085	1.449 (1)	0.229	3.662 (2)	0.160	0.255 (2)	0.880	2.155 (2)	0.340
	4	234	2.804 (1)	0.094	3.513 (1)	0.061	11.037 (1)	**0.001**	10.544 (2)	**0.005**	10.116 (2)	**0.006**	2.397 (2)	0.302
5th–6th grade cohort	1	211	9.077 (1)	**0.003**	1.5 (1)	0.221	3.059 (1)	0.080	3.056 (2)	0.217	10.232 (2)	**0.006**	9.389 (2)	**0.009**
	2	201	9.685 (1)	**0.002**	18.086 (1)	<**0.001**	9.551 (1)	**0.002**	12.839 (2)	**0.002**	7.828 (2)	**0.020**	14.254 (2)	**0.001**
	3	281	10.905 (1)	**0.001**	2.194 (1)	0.139	17.954 (1)	<**0.001**	15.696 (2)	<**0.001**	30.375 (2)	<**0.001**	10.741 (2)	**0.005**
	4	273	9.23 (1)	**0.002**	14.339 (1)	<**0.001**	38.936 (1)	<**0.001**	18.365 (2)	<**0.001**	13.323 (2)	**0.001**	7.003 (2)	**0.030**
7th–8th grade cohort	1	447	10.761 (1)	**0.001**	2.335 (1)	0.126	5.982 (1)	**0.014**	5.445 (2)	0.066	14.184 (2)	**0.001**	10.575 (2)	**0.005**
	2	453	13.883 (1)	<**0.001**	5.112 (1)	**0.024**	6.9 (1)	**0.009**	6.739 (2)	**0.034**	15.111 (2)	**0.001**	13.203 (2)	**0.001**
	3	432	26.161 (1)	<**0.001**	7.093 (1)	**0.008**	21.051 (1)	<**0.001**	18.276 (2)	<**0.001**	38.066 (2)	<**0.001**	20.456 (2)	<**0.001**
	4	410	42.235 (1)	<**0.001**	20.249 (1)	<**0.001**	37.355 (1)	<**0.001**	33.262 (2)	<**0.001**	57.776 (2)	<**0.001**	41.058 (2)	<**0.001**

IR, interference resolution; CM, context monitoring; WM, working memory; $\Delta\chi^2$, difference in Satorra–Bentler scaled χ^2 between nested and comparison models; Δdf, difference in degrees of freedom between nested and comparison models. Bolded values *p*-values represent cases where the more complex model shows significantly better model fit compared to the simpler model.

10.3389/fnhum.2023.1195013

answer questions about the differentiation hypothesis becomes even more untenable.

3.2.2. Network analysis

Next, we demonstrate how using network analysis to treat EF task performances as an interconnected set of cognitive processes leads to insights into their development, which were not revealed using latent variable modeling. Network analysis provided a data-driven method for grouping task performance according to strength of in-group performance compared to out-group performance, resulting in EF component construction that was not restricted by theoretical assumptions of which tasks draw on each EF component. Further, because we used partial correlations to form networks, the degree of differentiation of components identified with this method is unambiguous; components are only identified if they are distinct from the unifying Common EF component. Thus, network analyses allow for the examination of component grouping after Common EF is accounted for.

Concerning the number of components, community detection results (**Figure 2**) revealed that the EFs examined in this study were organized into two communities through grades 3 and 4, then stabilized into a three-community structure by 5th grade. Yet even through 8th grade, the relationships between tasks continued to evolve over time. Both the CFA and network analytic methods indicated the organization of EF task performances was most variable early in development through grades 3 and 4. However, unlike latent variable modeling, network analysis showed that while the *number* of communities for the 3rd–4th grade cohort was consistent across timepoints, the *composition* of these communities was variable. In this youngest cohort, community detection analysis consistently suggested two of the three theorized components combined into a single component, though similar to the 2-factor solutions tested in the CFA, which component was distinct differed across all four timepoints. Network analysis showed WM was distinct at timepoint 1, IR at timepoint 2, both IR and CM at timepoint 3, and CM at timepoint 4. EF organization for the older cohorts, though, was relatively stable. For both the 5th–6th grade and 7th–8th grade cohorts, the tasks almost always formed three communities with groupings consistent with those predicted by theory. However, for the 5th-6th grade cohort, at timepoint 1, Sustained Attention and Flanker switched communities, grouping with IR and CM communities respectively. Further, at timepoint 2, Tap and Trace was grouped with IR tasks for the 7th–8th grade cohort. Thus, while the EFs examined in this study can be organized into at least three distinct components by about 5th grade, network analysis suggests organization of the IR and CM components in particular continue to undergo refinement across the developmental period examined here. See **Supplementary material** for additional analyses supporting the results of the community detection analysis.

As indicated by the varying line thickness connecting tasks across models in **Figure 2**, connections between tasks both within and between communities waxed and waned over development, suggesting the organization of these EFs continued to be refined over time. See **Supplementary Figure 6** for estimates for all edge weights with parametric bootstrapped 95% confidence intervals. A unique benefit of network analysis is our ability to leverage the resulting network metrics to quantify and compare the degree of network stability across cohorts. Specifically, we can determine how stable a network is by examining how strongly individual network connections correlate across timepoints for a given cohort. For example, while the strength of individual connections between task performances (e.g., Flanker and Stroop) might increase or decrease over time, these changes are occurring in similar ways over time for a given cohort, the network would be considered more stable in that the organization of task performance is unlikely to change. We used a one-way ANOVA to directly interrogate whether correlations between network connections (**Table 4**) were more variable in younger cohorts compared to older cohorts. Results revealed these correlations indeed significantly differed across cohorts ($F(2,15) = 11.29$, $p = 0.001$, $\eta^2 = 0.60$). Tukey post-hoc tests showed correlations between the 3rd–4th grade cohort networks connections were significantly lower than correlations between both the 5th–6th grade cohort networks ($M_{difference} = 0.34$, 95% CI [0.09–0.58], $p = 0.008$) and the 7th–8th grade cohort networks ($M_{difference} = 0.43$, 95% CI [0.18, 0.68], $p = 0.001$). Correlations between network connections over time did not differ significantly between these two older cohorts though ($M_{difference} = 0.09$, 95% CI [-0.15, 0.35], $p = 0.60$). Thus, the period between 3rd and 4th grade is further supported as one in which the organization of EFs is undergoing larger degrees of change compared to the period between 5th and 8th grade, which may show more incremental change. Together, the results of the community detection analysis and the between-cohort differences in network connection correlations illustrate how a holistic examination of the EF system that accounts for Common EF can reveal novel insights into how these processes develop, beginning to resolve the inconsistencies across the literature that have emerged from the use of a reductionist framework that treats components as distinct, but correlated constructs.

4. Discussion

This study exemplified a feasible analytical technique for testing the differentiation hypothesis and for revealing new insights into the developmental trajectories of EFs. It further demonstrated how methodological choices can influence conclusions and interpretations around the organization of EFs, particularly in developmental populations. By comparing and contrasting the results across analytic techniques, we can bring a new lens to the inconsistencies in the number of EF components in children reported in the literature to date and, with further investigation, resolve them. Ultimately, this work can lay the groundwork towards building a clearer consensus on which EFs emerge on what timeline, and what factors might influence their development.

By applying network analysis techniques, we established a clear developmental timeline of EF organization in our sample and revealed several critical insights into how three EFs examined in the current study evolve over time. First, while both modeling methods used in our analyses point to organization of the examined EFs stabilizing around 5th grade, network analyses were unambiguous in the number of EF components at each timepoint. Network analyses of this sample revealed that a single, undifferentiated component of EF is an unlikely organization for any age in grades 3 through 8. Second, both methods suggest greater variability in the 3rd–4th grade cohort and continued refinement from 5th through

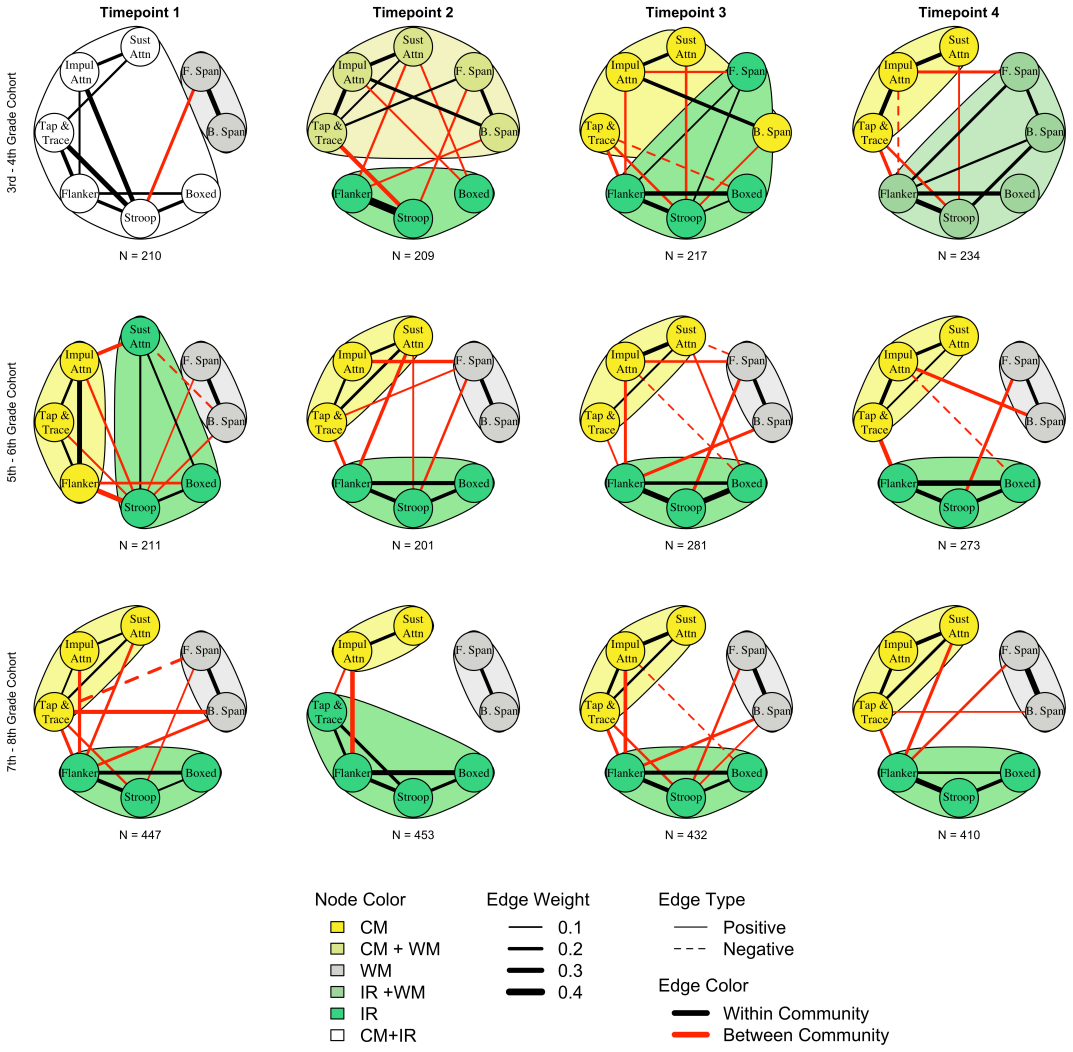


FIGURE 2
Results of network analysis and community detection for each cohort and timepoint. Strength of the connection between task performance is indicated by line thickness; thicker lines indicate a stronger relationship between two tasks. Edges between -0.1 and 0.1 are not shown for visualization purposes. Connections between tasks are further categorized as either within community (black) or between community (red); weaker and fewer between community connections compared to within community connections is consistent with more distinct communities. Community detection algorithms indicate a two-community organization for the 3rd–4th grade cohort that differentiates into a three community structure by about 5th grade. Fluctuations in grouping and magnitude of edge weights across older cohorts suggests continued subtle development for older students. WM, working memory; CM, context monitoring; IR, interference resolution; B. span, backward spatial span; F. span, forward spatial span; Sust Attn, sustained attention; Impul Attn, impulsive attention.

at least 8th grade, but only network analysis revealed which EFs are developing and in what way. Our methods revealed that the variability in the 3rd– 4th grade cohort sample was likely due to development and not to traditional constraints such as sample size and measurement differences. Finally, unlike latent variable analysis, the metrics generated from network analyses were used to gain further insight into the development of EFs and develop new hypotheses around their trajectories.

This study presents innovative methods for understanding precisely how EFs differentiate across middle childhood. Adaptive algorithms in our EF assessments allow us to meet the learner where they are, regardless of ability and without making assumptions about skill level according to demographic variables such as age and allow for multiple assessments within-subject over time. Using our novel technology, we administered assessments to large groups of children at once, affording us a larger sample size for each age group studied. These large samples of students, who completed the same tasks that presented a similar degree of challenge according to individual performance, represent a unique dataset from which to understand three EFs. Paired with simulation results, we can be more confident that differences seen between cohorts are developmentally-related—not merely due to differences in sample size or task difficulty. In this way, we overcame one pernicious limitation in the extant literature, which has commonly had to use different tasks for different age ranges (e.g., McAuley and White, 2011; Camerota et al., 2020; though see Van der Ven et al., 2012; Boelema et al., 2014) or seen ceiling effects in performance by older students (e.g., Lee et al., 2013).

Using a network analytic approach and leveraging the power of this dataset, we were able to explore new avenues for understanding

TABLE 4 Correlations of network connections between network models.

	Timepoint	3rd–4th grade cohort				5th–6th grade cohort				7th–8th grade cohort			
		1	2	3	4	1	2	3	4	1	2	3	4
3rd–4th grade cohort	1	1.00											
	2	0.40 [0.03, 0.67]	1.00										
	3	0.36 [-0.02, 0.64]	0.32 [-0.06, 0.62]	1.00									
	4	0.41 [0.05, 0.68]	0.32 [-0.06, 0.62]	0.45 [0.09, 0.7]	1.00								
5th–6th grade cohort	1	0.48 [0.13, 0.72]	0.56 [0.23, 0.77]	0.54 [0.2, 0.76]	0.43 [0.07, 0.69]	1.00							
	2	0.37 [0, 0.65]	0.34 [-0.04, 0.63]	0.41 [0.04, 0.68]	0.51 [0.17, 0.74]	0.39 [0.01, 0.66]	1.00						
	3	0.43 [0.07, 0.69]	0.48 [0.13, 0.72]	0.47 [0.12, 0.72]	0.43 [0.07, 0.69]	0.75 [0.52, 0.88]	0.67 [0.4, 0.84]	1.00					
	4	0.36 [-0.01, 0.65]	0.51 [0.18, 0.74]	0.64 [0.34, 0.82]	0.55 [0.22, 0.76]	0.62 [0.32, 0.81]	0.58 [0.27, 0.79]	0.63 [0.33, 0.81]	1.00				
7th–8th grade cohort	1	0.36 [-0.02, 0.65]	0.55 [0.23, 0.77]	0.42 [0.05, 0.68]	0.43 [0.07, 0.69]	0.57 [0.25, 0.78]	0.41 [0.04, 0.68]	0.60 [0.3, 0.8]	0.52 [0.18, 0.75]	1.00			
	2	0.34 [-0.04, 0.63]	0.54 [0.21, 0.76]	0.50 [0.16, 0.74]	0.42 [0.05, 0.68]	0.71 [0.46, 0.86]	0.60 [0.29, 0.79]	0.72 [0.47, 0.86]	0.69 [0.42, 0.84]	0.66 [0.38, 0.83]	1.00		
	3	0.54 [0.21, 0.76]	0.62 [0.33, 0.81]	0.39 [0.02, 0.67]	0.47 [0.12, 0.72]	0.72 [0.47, 0.86]	0.65 [0.37, 0.82]	0.78 [0.58, 0.9]	0.69 [0.42, 0.84]	0.61 [0.31, 0.8]	0.85 [0.69, 0.93]	1.00	
	4	0.55 [0.22, 0.77]	0.63 [0.34, 0.81]	0.36 [-0.02, 0.65]	0.51 [0.17, 0.74]	0.55 [0.22, 0.76]	0.63 [0.33, 0.81]	0.54 [0.21, 0.76]	0.62 [0.33, 0.81]	0.61 [0.3, 0.8]	0.64 [0.35, 0.82]	0.64 [0.35, 0.82]	1.00

10.3389/fnhum.2023.1195013

the development of EF as a dynamic interconnected network of skills that can align behavioral and neural models. Our series of analyses provide converging evidence that the period from third to fourth grade is one of great change in the structure and organization of EFs compared to later periods in development. Not only did both latent variable and network analysis show a greater degree of variability in the model that best represents organization of EFs, but the between network correlations between edge weights support characterizing organizations as "unstable". This pattern of findings may suggest individual differences in component differentiation that should be explored in future research. Experience may drive differentiation rates to differ across children. Further, individual differences in differentiation rates may also explain differences in the number of EF components found in this age range. For example, studies have shown students might differentially employ EFs based on, for example, pubertal stage, socioeconomic background, et cetera (Haft and Hoeft, 2017; Doebel, 2020). Such differences in the way individuals employ EFs may also impact the trajectory of differentiation of these EFs. Understanding the potential paths in development and how they can be influenced by life experience will be critical in fostering continued growth of EF skills (Best and Miller, 2010).

In the current study, network analysis allowed us to go beyond assessing the stability of the number of components across development and extend our assessment to the stability of component *composition*. As discussed, the 3rd–4th grade cohort in this study was highly variable across time, showing a different combination of components at each of the four timepoints examined. However, this variation in organization was not restricted to the youngest cohort; the 5th–6th grade cohort studied here showed a non-hypothesized organization at timepoint 1, namely, Flanker grouping with context monitoring rather than interference resolution, and Sustained Attention grouping with interference resolution. This finding, consistent with prior work showing protracted development of EF skill (Davidson et al., 2006), emphasizes that EFs may manifest or be deployed differently across development, and tasks shown to measure one construct in adults may measure a different one in children (Morra et al., 2018). Such potential differences in how EFs might be employed to accomplish a task across development were missed when using a latent variable model approach, and may help explain the inconsistencies in the extant literature regarding the number of components in this age range (Lee et al., 2013). Latent variable analysis does not allow for statistically comparing models with different configurations of the same indicators. As such, alternate configurations are often not investigated. During a period of such developmental instability, the differences in the tasks used to measure each component and the metric of skill on each of those tasks across studies could result in many acceptable models of the data. Without a data-driven method for determining which EF component a task reflects, researchers are left with an untenable number of configurations to test. Indeed in the current study, such configurational differences were missed with factor analysis, since the theory-driven configuration of EFs fit reasonably well, and there was no indication a different configuration might better represent EF constructs. Considering alternative approaches such as the network analysis shown here can add to our understanding of the measurement approach that best represents EFs across the lifespan (Camerota et al., 2020).

Importantly, the use of network analysis to test the differentiation hypothesis allowed for the examination of how different EFs become distinct from not only each other, but from Common EF. To date, only one other investigation to our knowledge has used analytic methods that support such an investigation (Hartung et al., 2020). While this investigation examined different EFs than those studied here (specifically, Working Memory, Switching, Updating, and Inhibition), the results are largely complementary. Specifically, Hartung and colleagues analyses indicated that in younger children age 8–10, EFs were highly correlated with one another, suggesting little differentiation between Common EFs and individual components at this age. Further, Inhibition, most similar to the CM and IR examined here, became increasingly differentiated by about age 10, consistent with the finding from the current study that organization of CM, IR, and WM stabilized around 5th grade, or age 10. Finally, a primary finding from Hartung et al.'s (2020) investigation was the lack of a uniform pattern of development across either components or individual tasks, suggesting a more nuanced pattern of developmental trajectories, consistent with the findings from the current study. Both studies underscore the importance of carefully considering which components are measured in what way, and whether the relationships between tasks and EFs seen in adults holds true for childhood populations.

4.1. Limitations and future directions

This study makes significant strides in our approach to measure and model EFs, improving on several critical limitations in the field. Yet, further advancements are needed to build upon and address limitations of this work, particularly regarding the scope of EFs assessed and the availability of statistical methods to compare network models longitudinally.

4.1.1. EF measurement

Developing a novel, adaptive battery of EF tasks for all ages and abilities was not without its challenges, and a future iteration of this battery that addresses many of the challenges encountered here is already underway. This iteration, called ACE Explorer (ACE-X) is currently undergoing large-scale norming and validation with a nationally representative sample across ages 7–107. A key challenge with using ACE-C concerned the design decisions made when modifying tasks for large-group assessment and to incorporate adaptive algorithms. Specifically, in this study, the WM component was only indexed by two measures, which limited the type of latent variable model that could be constructed and tested here. While a third task hypothesized to measure WM, Filter, was originally included in the ACE-C battery, it used a different adaptive mechanism, which resulted in age-related differences in challenge level, and ultimately its exclusion from the current analysis. Consequently, we could not test certain factor configurations without rendering the models uninformative. In ACE-X, we have aligned the adaptive mechanism to use the response window in the same manner as the majority of other tasks in the battery, which has resulted in more consistent challenge-levels across age groups.

Further, as with any investigation that does include an exhaustive assessment of all potential EFs, the conclusions

concerning developmental trajectory of EFs can only be applied to what was examined. The components examined via ACE-C were not intended to be an exhaustive list of potential EF components, and notably, not all components intended to be measured with ACE-C were able to be included. Due to time constraints associated with in-school testing sessions, we were limited in the number of tasks that could be administered, and we chose to focus on tasks commonly used to assess EFs and cognitive control across both the adult and developmental literature to better bridge our understanding of these constructs across the lifespan. Further, while we did administer a task intended to assess the cognitive flexibility component of EF, a technical malfunction in the analytics for this task prevented its inclusion in the current study. As such, one prominent EF component was not assessed here, though this issue has been corrected in the ACE-X battery.

Additionally, careful consideration must be given to the terms that are used to discuss EF components, and how those terms are reflected by the task designs used in each investigation. For example, the Stroop task has been considered to measure inhibition when the verbal response mechanism is used, but interference resolution when a motor response mechanism is used, as is the case in the current study. Further, while the components put forth by Miyake and colleagues (Miyake et al., 2000b; "updating", "inhibitory control", and "cognitive flexibility") are the most frequently examined components (Karr et al., 2018) they are often inconsistently defined across the literature. In particular, the "inhibitory control" component is often measured with a combination of tasks that involve both the "interference resolution" and "context monitoring" aspects of cognitive control (see Diamond, 2013 for review). However, neural data from both children and adults indicate these are indeed two separate components (Bunge et al., 2002). By including additional tasks (e.g., Boxed and Tap and Trace) and separating inhibitory control-related tasks into those in which a response must always be made (interference resolution) and those in which a participant must decide whether to make a response or not (context monitoring), the ACE-C battery is able to bring further specificity to the characterization of EFs in middle childhood.

Similarly, though, the "working memory" component of EF would benefit from increased precision around its definition, and therefore measurement. The field has not yet reached a consensus on whether "short term memory" is distinct from "working memory" and whether these constructs might differ across development as this component evolves. While the inclusion of both a forward and backward span in the ACE-C battery was done in keeping with their widespread use in clinical practice to assess what is referred to as "working memory" (see Berch et al., 1998), these two tasks do not exhaustively capture all potential aspects of the construct. Indeed, in this investigation we used the term "working memory" rather than "updating" as is used in the most commonly cited model of EF components (Miyake et al., 2000b) because the Forward Spatial Span task does not strictly fit with the component conceptualized as "updating". By including additional tasks that tap different aspects of "working memory", such as the Filter task that examines the ability to remember task-relevant information while ignoring task irrelevant information (Luck and Vogel, 1997), we can further understand the composition of this construct and bring increased specificity to how it is discussed and measured.

Finally, future directions for the ACE battery include increasing its capabilities as a measurement tool of multiple components of EF. First, ACE-X has been made more inclusive by using a color palette compliant with the Americans with Disabilities Act to ensure individuals who are colorblind can use the battery. Second, to build on to the engaging design that afforded us high retention and compliance rates in this study, ACE-X incorporates the battery of tasks into a cohesive story to further motivate participants to complete the full battery. Finally, the large-scale validation efforts and norming with a nationally representative population will further allow us to replicate the results shown here in additional populations, including within sub-populations represented but not separately examined in this study (e.g., students with learning disabilities). In this way, we will be able to replicate and extend the results of the current study, to better understand additional factors that may impact the developmental trajectory described here.

4.1.2. EF modeling

This study demonstrated a new approach to modeling EFs that better accounts for the unity while examining the diversity of EFs. Yet, this methodological approach must continue to be built upon to fully model the development of EFs. Indeed, there were methodological challenges related to comparing two analytical approaches in testing the differentiation hypothesis. For example, we intentionally did not explicitly model the dependency of multiple observations per student that occurs with longitudinal data in either analytic approach. While it is possible to model using factor analysis, development of network models that can handle longitudinal data are still in their infancy (though see Deserno et al., 2021). To keep the general modeling strategy consistent and inferences comparable, we treated all observations as independent in both approaches. However, this strategy is unlikely to have affected the results for two reasons. First, without accounting for within-person changes, within-cohort comparisons were more conservative than necessary. Second, we did not perform tests that were likely to be affected by treating observations as independent. Nonetheless, as network analytic methodology continues to advance, so too must the methods used to reveal the evolution of EF structure advance.

Further, neither modeling approach was able to simultaneously account for Common EF *and* provide statistical comparisons between models of differing complexity. With latent variable analysis, it is a straightforward process to compare whether a model with more factors fits statistically better than a model with fewer factors. These capabilities, though, are currently limited with network models (though see Epskamp et al., 2021). Community detection algorithms provide a likely grouping for task performances, but there is no index to statistically determine whether a two-community network explains EFs just as well as a three-community network, for example. However, existing methods for accounting for Common EF in the latent variable approach preclude such statistical comparisons between models, leaving the theoretical problem of how to account for Common EF in the context of differentiation of components with this approach unresolved. To date, the benefits of the network analysis approach, which accounts for commonality among all EF task performances rather than treating it as a separate component entirely, presents a promising solution for accounting for Common EF. The rapidly emerging statistical approaches for testing network

model complexity position this technique as the path forward in establishing the developmental trajectory of EFs.

The future potential for network analysis to help us understand complex cognitive constructs is bright. Researchers in related fields have already begun to capitalize on information gained from taking a network analytic perspective to understand other cognitive processes. For example, Kan et al. (2020) demonstrated how fit statistics can be obtained for network models, allowing a direct comparison between network and latent variable models. As such, future research could directly compare a variety of configurations of EF modeled using latent variable analysis to those using network analysis to determine which organization best fits observed EF performance. While outside the scope of the current paper, researchers in the field of intelligence have used this approach to show that modeling aspects of intelligence as being mutually and reciprocally related through a network framework is favored over modeling an overarching umbrella component ("g") in a latent variable framework (Kan et al., 2019). Given this field's similar dilemma around how to quantify developmental differentiation in the presence of task commonality (Molenaar et al., 2010), we anticipate such investigations in EFs will be similarly fruitful for determining which modeling approach better reflects the unity and diversity of EFs and for elucidating the mechanisms through which skill changes arise.

Further, as methods for appropriately modeling longitudinal data emerge, network analysis provides an avenue for understanding the potential reciprocal relationships among EFs over time (Deserno et al., 2021). For example, in a separate study we are examining how growth in performance on individual tasks are connected. By using a network framework for investigating EF skill growth, we can evaluate whether the same communities formed when modeling contemporaneous ties between task performances also emerge when looking at their patterns of growth across time. Such evidence would reinforce the identity of the communities as distinct components of EF and allow us to answer whether components of EF emerge independently or in tandem with other components.

Such insights into the development of EFs are critical for advancing our understanding of how they influence, and can be influenced by, internal and external factors. For example, EFs are often the focus of educational interventions with the goal of improving academic-related outcomes (see e.g., Diamond and Lee, 2011; Titz and Karbach, 2014; Jacob and Parkinson, 2015). Network analysis is well-poised to generate hypotheses regarding which EF tasks or components might be more likely to transfer outside a training regime, which can then guide future training studies. Indeed, the findings from the current study provide a clear set of testable hypotheses: given that the cross-sectional network models found here suggest that WM is less strongly connected to other EF components, future training studies should test the hypothesis that training a highly connected component such as IR would be more likely to result in transfer to other EFs compared to training on the less-well connected WM component.

4.2. Conclusion

The findings from this study showcase how advances in assessing EFs and an increasingly popular modeling technique, network analysis, can be applied to the field of EFs to better align behavioral and neural investigations. The dual paradigm shifts to network analysis using adaptive measures provide a promising pathway for refining and specifying our understanding of how EFs develop. These insights can in turn be applied to advance our understanding of EFs' wide-reaching impact on factors related to physical and cognitive health across the lifespan (Zelazo et al., 2016). Together, our improved methodological approaches to *measuring* EFs can lead to the development of improved methods for *supporting* EFs and providing students the proper foundation they need for learning and future educational success.

Data availability statement

The original contributions presented in this study are publicly available. This data can be found here: https://osf.io/scpkm/.

Ethics statement

The studies involving human participants were reviewed and approved by The University of California San Francisco Institutional Review Board. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

JA, SB, FH, BM, JM, MR-L, AG, and MU conceived of and designed the study. JY and MU collected the data. JY, KO'L, EF, and MU analyzed the data. JY, KO'L, and MU wrote the manuscript. All authors discussed the results and contributed to editing the manuscript.

Funding

This research was supported by funding from the National Science Foundation, Science of Learning Collaborative Networks Grant (NSFSLCN-1540854) awarded to MU [lead Principal Investigator (PI)] and AG, and co-PIs JA, SB, FH, BM, JM, and MR-L.

Acknowledgments

We authors would like to thank the research staff, Jordin Rodondi, Caleb Banks, Zoe D'Esposito, John David Lorentz, and the large team of UCSF volunteers as well as the students, teachers, parents, and school and district administrators who made this research program possible. The authors are grateful for consultation from Stephanie Haft, Ariel Starr, Joshua Jordan, and Britte Cheng. We authors are also grateful to the team of developers who have made the ACE assessment possible including WoWLabz, Zynga.org, and Rose Feldman, and the programmers that created the aceR processing code Jose Gallegos and Monica Thieu.

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.2023.1195013/full#supplementary-material

References

Agostino, A., Johnson, J., and Pascual-Leone, J. (2010). Executive functions underlying multiplicative reasoning: Problem type matters. *J. Exp. Child Psychol.* 105, 286–305. doi: 10.1016/J.JECP.2009.09.006

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychol.* 8, 71–82. doi: 10.1076/chin.8.2.71.8724

Anguera, J. A., Jordan, J. T., Castaneda, D., Gazzaley, A., and Areán, P. A. (2016b). Conducting a fully mobile and randomised clinical trial for depression: Access, engagement and expense. *BMJ Innov.* 2, 14–21. doi: 10.1136/bmjinnov-2015-000098

Anguera, J. A., Brandes-Aitken, A. N., Rolle, C. E., Skinner, S. N., Desai, S. S., Bower, J. D., et al. (2016a). Characterizing cognitive control abilities in children with 16p11.2 deletion using adaptive 'video game' technology: A pilot study. *Transl. Psychiatry* 6:e893. doi: 10.1038/tp.2016.178

Bates, D., Mächler, M., Zurich, E., Bolker, B. M., and Walker, S. C. (2014). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48.

Berch, D., Krikorian, R., and Huha, E. (1998). The Corsi Block-Tapping Task: Methodological and theoretical considerations. *Brain Cogn.* 38, 317–338. doi: 10.1006/brcg.1998.1039

Best, J. R., and Miller, P. H. (2010). A developmental perspective on executive function. *Child Dev.* 81, 1641–1660. doi: 10.1111/j.1467-8624.2010.01499.x

Best, J. R., Miller, P. H., and Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learn. Individ. Differ.* 21, 327–336. doi: 10.1016/j.lindif.2011.01.007

Boelema, S. R., Harakeh, Z., Ormel, J., Hartman, C. A., Vollebergh, W. A. M., and Van Zandvoort, M. J. E. (2014). Executive functioning shows differential maturation from early to late adolescence: Longitudinal findings from a TRAILS study. *Neuropsychology* 28, 177–187. doi: 10.1037/neu0000049

Borsboom, D., and Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annu. Rev. Clin. Psychol.* 9, 91–121. doi: 10.1146/ANNUREV-CLINPSY-050212-185608

Brocki, K. C., and Bohlin, G. (2004). Executive functions in children aged 6 to 13: A dimensional and developmental study. *Dev. Neuropsychol.* 26, 571–593. doi: 10.1207/s15326942dn2602_3

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., and Gabrieli, J. D. E. (2002). Immature frontal lobe contributions to cognitive control in children: Evidence from fMRI. *Neuron* 33, 301–311. doi: 10.1016/S0896-6273(01)00583-9

Camerota, M., Willoughby, M. T., and Blair, C. B. (2020). Measurement models for studying child executive functioning: Questioning the status quo. *Dev. Psychol.* 56, 2236–2245. doi: 10.1037/dev0001127

Carlson, S. M., and Zelazo, P. D. (2014). *Minnesota executive function scale: Test manual.* Saint Paul, MN: Reflection Sciences.

Corsi, P. M. (1973). *Human memory and the medial temporal region of the brain.* Montreal, QC: McGill University.

Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Mõttus, R., Waldorp, L. J., et al. (2015). State of the aRt personality research: A tutorial on network analysis of personality data in R. *J. Res. Pers.* 54, 13–29. doi: 10.1016/j.jrp.2014.07.003

Csardi, G., and Nepusz, T. (2006). The igraph software package for complex network research. *Int. J. Complex Syst.* 1695:1–9.

Davidson, M., Amso, D., Anderson, L., and Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia* 44, 2037–2078.

Deserno, M., Sachisthal, M., and Epskamp, S. (2021). A magnifying glass for the study of coupled developmental changes: Combining psychological networks and latent growth models. *PsyArXiv* [Preprint]. doi: 10.31234/osf.io/ngfxq

Diamond, A. (2013). Executive functions. *Annu. Rev. Psychol.* 64, 135–168. doi: 10.1146/annurev-psych-113011-143750

Diamond, A., and Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science* 333, 959–64. doi: 10.1126/science.1204529

Doebel, S. (2020). Rethinking executive function and its development. *Perspect. Psychol. Sci.* 15, 942–956. doi: 10.1177/1745691620904771

Draheim, C., Tsukahara, J. S., Martin, J. D., Mashburn, C. A., and Engle, R. W. (2020). A toolbox approach to improving the measurement of attention control. *J. Exp. Psychol. Gen.* 150, 242–275. doi: 10.1037/xge0000783

Engelhardt, L. E., Briley, D. A., Mann, F. D., Harden, K. P., and Tucker-Drob, E. M. (2015). Genes unite executive functions in childhood. *Psychol. Sci.* 26, 1151–1163. doi: 10.1177/0956797615577209

Enkavi, Z. A., Eisenberg, I. W., Bissett, P. G., Mazza, G. L., MacKinnon, D. P., Marsch, L. A., et al. (2019). Large-scale analysis of test–retest reliabilities of self-regulation measures. *Proc. Natl. Acad. Sci. U.S.A.* 116, 5472–5477.

Epskamp, S. (2015). *bootnet: Bootstrap methods for various network estimation routines. R-Package.*

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., and Borsboom, D. (2012). Qgraph: Network visualizations of relationships in psychometric data. *J. Stat. Softw.* 48, 1–18. doi: 10.18637/jss.v048.i04

Epskamp, S., Isvoranu, A.-M., and Cheung, M. W.-L. (2021). Meta-analytic gaussian network aggregation. *Psychometrika* 87, 12–46. doi: 10.1007/S11336-021-09764-3

Eriksen, B., 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.

Eversheim, U., and Bock, O. (2001). Evidence for processing stages in skill acquisition: A dual-task study. *Learn. Mem.* 8, 183–189. doi: 10.1101/lm.39301

Finch, J. E., Garcia, E. B., Sulik, M. J., and Obradović, J. (2019). Peers matter: Links between classmates' and individual students' executive functions in elementary school. *AERA Open* 5, 1–14. doi: 10.1177/2332858419829438

Friedman, N. P., and Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex* 86, 186–204. doi: 10.1016/j.cortex.2016.04.023

Friedman, N. P., Miyake, A., Robinson, J. L., and Hewitt, J. K. (2011). Developmental trajectories in toddlers' self-restraint predict individual differences in executive functions 14 years later: A behavioral genetic analysis. *Dev. Psychol.* 47, 1410–1430. doi: 10.1037/a0023750

Friedman, N. P., Miyake, A., Young, S. E., Defries, J. C., Corley, R. P., and Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *J. Exp. Psychol. Gen.* 137, 201–225. doi: 10.1037/0096-3445.137.2.201

Fry, A. F., and Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychol. Sci.* 7, 231–241. doi: 10.1111/j.1467-9280.1996.tb00366.x

Greenberg, L., Leark, R., Dupuy, T., and Corman, C. (1991). *The test of variables of attention (TOVA).* Los Alamitos, CA: *Universal Attention Disorders.*

Haft, S. L., and Hoeft, F. (2017). Poverty's impact on children's executive functions: Global considerations. *New Direct. Child Adolesc. Dev.* 2017, 69–79. doi: 10.1002/cad.20220

Hartung, J., Engelhardt, L. E., Thibodeaux, M. L., Harden, K. P., and Tucker-Drob, E. M. (2020). Developmental transformations in the structure of executive functions. *J. Exp. Child Psychol.* 189:104681. doi: 10.1016/j.jecp.2019.104681

Hatoum, A. S., Morrison, C. L., Mitchell, E. C., Lam, M., Benca-Bachman, C. E., Reineberg, A. E., et al. (2020). Genome-wide association study of over 427,000 individuals establishes executive functioning as a neurocognitive basis of psychiatric disorders influenced by GABAergic processes. *bioRxiv* [Preprint]. doi: 10.1101/674515

Hu, L. T., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct. Equat. Model.* 6, 1–55. doi: 10.1080/10705519909540118

Ishihara, S. (1972). *Tests for colour-blindness*. Bunkyo-Ku: Kanehara Shuppan Co.

Jacob, R., and Parkinson, J. (2015). The potential for school-based interventions that target executive function to improve academic achievement: A review. *Rev. Educ. Res.* 85, 512–552. doi: 10.3102/0034654314561338

Johnson, M. H. (2011). Interactive specialization: A domain-general framework for human functional brain development? *Dev. Cogn. Neurosci.* 1, 7–21. doi: 10.1016/J.DCN.2010.07.003

Kan, K. J., de Jonge, H., van der Maas, H. L. J., Levine, S. Z., and Epskamp, S. (2020). How to compare psychometric factor and network models. *J. Intell.* 8:35. doi: 10.3390/jintelligence8040035

Kan, K. J., van der Maas, H. L. J., and Levine, S. Z. (2019). Extending psychometric network analysis: Empirical evidence against g in favor of mutualism? *Intelligence* 73, 52–62. doi: 10.1016/J.INTELL.2018.12.004

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., and Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychol. Bull.* 144, 1147–1185. doi: 10.1037/bul0000160

Kline, R. B. (2011). *Principles and practice of structural equation modeling*, 3rd Edn. New York, NY: Guilford Press.

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). lmerTest Package: Tests in linear mixed effects models. *J. Stat. Softw.* 82, 1–26. doi: 10.18637/jss.v082.i13

Leark, R. A., Greenberg, L. M., Kindschi, C. L., Dupuy, T. R., and Hughes, S. J. (2018). *The TOVA Professional Manual*, 9th Edn. Los Alamitos, CA: The Tova Company.

Lee, K., Bull, R., and Ho, R. M. H. (2013). Developmental changes in executive functioning. *Child Dev.* 84, 1933–1953. doi: 10.1111/cdev.12096

Leys, C., Ley, C., Klein, O., Bernard, P., and Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *J. Exp. Soc. Psychol.* 49, 764–766. doi: 10.1016/j.jesp.2013.03.013

Luck, S. J., and Vogel, E. K. (1997). The capacity of visual working memory for scenes. *Nature* 390:1297. doi: 10.1167/18.10.1297

McAuley, T., and White, D. A. (2011). A latent variables examination of processing speed, response inhibition, and working memory during typical development. *J. Exp. Child Psychol.* 108, 453–468. doi: 10.1016/j.jecp.2010.08.009

McCoy, D. C. (2019). Measuring young children's executive function and self-regulation in classrooms and other real-world settings. *Clin. Child Fam. Psychol. Rev.* 22, 63–74. doi: 10.1007/s10567-019-00285-1

Mead, L. A., Mayer, A. R., Bobholz, J. A., Woodley, S. J., Cunningham, J. M., Hammeke, T. A., et al. (2002). Neural basis of the Stroop interference task: Response competition or selective attention? *J. Int. Neuropsychol. Soc.* 8, 735–742. doi: 10.1017/S1355617702860015

Miyake, A., Emerson, M. J., and Friedman, N. P. (2000a). Assessment of executive functions in clinical settings: Problems and recommendations. *Semin. Speech Lang.* 21, 169–183. doi: 10.1055/s-2000-7563

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000b). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: A latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/COGP.1999.0734

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H. L., et al. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2693–2698. doi: 10.1073/pnas.1010076108

Molenaar, D., Dolan, C. V., Wicherts, J. M., and van der Maas, H. L. J. (2010). Modeling differentiation of cognitive abilities within the higher-order factor model using moderated factor analysis. *Intelligence* 38, 611–624. doi: 10.1016/j.intell.2010.09.002

Morra, S., Panesi, S., Traverso, L., and Usai, M. C. (2018). Which tasks measure what? Reflections on executive function development and a commentary on Podjarny, Kamawar, and Andrews (2017). *J. Exp. Child Psychol.* 167, 246–258. doi: 10.1016/j.jecp.2017.11.004

Mungas, D., Widaman, K., Zelazo, P. D., Tulsky, D., Heaton, R. K., Slotkin, J., et al. (2013). NIH toolbox cognition battery (CB): Factor structure for 3 to 15 year olds. *Monogr. Soc. Res. Child Dev.* 78, 103–118. doi: 10.1111/mono.12037

Muthén, L. K., and Muthén, B. O. (2017). *User's guide manual*, 7th Edn. Los Angeles, CA: Muthen and Muthen, 1–13. doi: 10.1111/j.1600-0447.2011.01711.x

Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., and Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cogn. Affect. Behav. Neurosci.* 12, 241–268. doi: 10.3758/s13415-011-0083-5

Obradović, J., Sulik, M. J., Finch, J. E., and Tirado-Strayer, N. (2018). Assessing students' executive functions in the classroom: Validating a scalable group-based procedure. *J. Appl. Dev. Psychol.* 55, 4–13. doi: 10.1016/j.appdev.2017.03.003

Packwood, S., Hodgetts, H. M., and Tremblay, S. (2011). A multiperspective approach to the conceptualization of executive functions. *J. Clin. Exp. Neuropsychol.* 33, 456–470. doi: 10.1080/13803395.2010.533157

Pascual, A. C., Moyano, N., and Robres, A. Q. (2019). The relationship between executive functions and academic performance in primary education: Review and meta-analysis. *Front. Psychol.* 10:1582. doi: 10.3389/FPSYG.2019.01582/BIBTEX

R Core Team (2020). *A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Reichardt, J., and Bornholdt, S. (2006). Statistical mechanics of community detection. *Phys. Rev. E* 74:016110. doi: 10.1103/PhysRevE.74.016110

Reineberg, A. E., Andrews-Hanna, J. R., Depue, B. E., Friedman, N. P., and Banich, M. T. (2015). Resting-state networks predict individual differences in common and specific aspects of executive function. *NeuroImage* 104, 69–78. doi: 10.1016/J.NEUROIMAGE.2014.09.045

Romine, C. B., and Reynolds, C. R. (2005). A model of the development of frontal lobe functioning: Findings from a meta-analysis. *Appl. Neuropsychol.* 12, 190–201. doi: 10.1207/s15324826an1204_2

Satorra, A., and Bentler, P. M. (2010). Ensuring positiveness of the scaled difference Chi-square test statistic. *Psychometrika* 75, 243–248. doi: 10.1007/s11336-009-9135-y

Schlam, T. R., Wilson, N. L., Shoda, Y., Mischel, W., and Ayduk, O. (2013). Preschoolers' delay of gratification predicts their body mass 30 years later. *J. Pediatr.* 162, 90–93. doi: 10.1016/j.jpeds.2012.06.049

Shing, Y. L., Lindenberger, U., Diamond, A., Li, S. C., and Davidson, M. C. (2010). Memory maintenance and inhibitory control differentiate from early childhood to adolescence. *Dev. Neuropsychol.* 35, 679–697. doi: 10.1080/87565641.2010.508546

Smolker, H. R., Friedman, N. P., Hewitt, J. K., and Banich, M. T. (2018). Neuroanatomical correlates of the unity and diversity model of executive function in young adults. *Front. Hum. Neurosci.* 12:283. doi: 10.3389/FNHUM.2018.00283

Spiegel, J. A., Goodrich, J. M., Morris, B. M., Osborne, C. M., and Lonigan, C. J. (2021). Relations between executive functions and academic outcomes in elementary school children: A meta-analysis. *Psychol. Bull.* 147, 329–351. doi: 10.1037/BUL0000322

Titz, C., and Karbach, J. (2014). Working memory and executive functions: Effects of training on academic achievement. *Psychol. Res.* 78, 852–868. doi: 10.1007/s00426-013-0537-1

Treisman, A. M., and Gelade, G. (1980). A feature-integration theory of attention. *Cogn. Psychol.* 12, 97–136.

Van der Ven, S. H. G., Kroesbergen, E. H., Boom, J., and Leseman, P. P. (2012). The development of executive functions and early mathematics: A dynamic relationship. *Br. J. Educ. Psychol.* 82, 100–119. doi: 10.1111/j.2044-8279.2011.02035.x

Vandierendonck, A. (2017). A comparison of methods to combine speed and accuracy measures of performance: A rejoinder on the binning procedure. *Behav. Res. Methods* 49, 653–673. doi: 10.3758/s13428-016-0721-5

Woltz, D. J., and Was, C. A. (2006). Availability of related long-term memory during and after attention focus in working memory. *Mem. Cogn.* 34, 668–684. doi: 10.3758/BF03193587

Xu, F., Han, Y., Sabbagh, M. A., Wang, T., Ren, X., and Li, C. (2013). Developmental differences in the structure of executive function in middle childhood and adolescence. *PLoS One* 8:e77770. doi: 10.1371/journal.pone.0077770

Younger, J., O'Laughlin, K. D., Anguera, J. A., Bunge, S., Ferrer, E. E., Hoeft, F., et al. (2022). Development of executive function in middle childhood: A large-scale, in-school, longitudinal investigation. *PsyArXiv* [Preprint]. doi: 10.31234/OSF.IO/XF489

Zelazo, D. P., and Bauer, P. (2013). *National Institutes of Health Toolbox cognition battery (NIH Toolbox CB): Validation for children between 3 and 15 years*. Hoboken, NJ: Wiley.

Zelazo, P. D., Blair, C. B., and Willoughby, M. T. (2016). *Executive function: Implications for education*. Available online at: http://ies.ed.gov/ncer/pubs/20172000/ (accessed November 16, 2017).

frontiers | Frontiers in Human Neuroscience

TYPE Original Research
PUBLISHED 10 August 2023
DOI 10.3389/fnhum.2023.1219189

OPEN ACCESS

EDITED BY
Hidehiko Okamoto,
International University of Health and Welfare (IUHW), Japan

REVIEWED BY
Linling Li,
Shenzhen University, China
Deokjong Lee,
Yonsei University Health System, Republic of Korea

*CORRESPONDENCE
Feiqiu Wen
✉ fwen62@126.com

RECEIVED 08 May 2023
ACCEPTED 12 July 2023
PUBLISHED 10 August 2023

CITATION
Huang W, Fateh AA, Zhao Y, Zeng H, Yang B, Fang D, Zhang L, Meng X, Hassan M and Wen F (2023) Effects of the SNAP-25 MnlI variant on hippocampal functional connectivity in children with attention deficit/hyperactivity disorder. *Front. Hum. Neurosci.* 17:1219189. doi: 10.3389/fnhum.2023.1219189

Effects of the SNAP-25 MnlI variant on hippocampal functional connectivity in children with attention deficit/hyperactivity disorder

Wenxian Huang[1,2], Ahmed Ameen Fateh[3], Yilin Zhao[1,3], Hongwu Zeng[3], Binrang Yang[2], Diangang Fang[3], Linlin Zhang[2], Xianlei Meng[3], Muhammad Hassan[3] and Feiqiu Wen[4]*

[1]Department of Pediatric China Medical University, Shenyang, China, [2]Healthy Care Center, Shenzhen Children's Hospital, Shenzhen, China, [3]Department of Radiology, Shenzhen Children's Hospital, Shenzhen, China, [4]Department of Pediatrics, Shenzhen Children's Hospital, Shenzhen, China

Objectives: Attention-deficit/hyperactivity disorder (ADHD) is one of the most widespread and highly heritable neurodevelopmental disorders affecting children worldwide. Although synaptosomal-associated protein 25 (SNAP-25) is a possible gene hypothesized to be associated with working memory deficits in ADHD, little is known about its specific impact on the hippocampus. The goal of the current study was to determine how variations in ADHD's SNAP-25 MnlI polymorphism (rs3746544) affect hippocampal functional connectivity (FC).

Methods: A total of 88 boys between the ages of 7 and 10 years were recruited for the study, including 60 patients with ADHD and 28 healthy controls (HCs). Data from resting-state functional magnetic resonance imaging (rs-fMRI) and clinical information were acquired and assessed. Two single nucleotide polymorphisms (SNP) in the SNAP-25 gene were genotyped, according to which the study's findings separated ADHD patients into two groups: TT homozygotes (TT = 35) and G-allele carriers (TG = 25).

Results: Based on the rs-fMRI data, the FC of the right hippocampus and left frontal gyrus was evaluated using group-based comparisons. The corresponding sensitivities and specificities were assessed. Following comparisons between the patient groups, different hippocampal FCs were identified. When compared to TT patients, children with TG had a lower FC between the right precuneus and the right hippocampus, and a higher FC between the right hippocampus and the left middle frontal gyrus.

Conclusion: The fundamental neurological pathways connecting the SNAP-25 MnlI polymorphism with ADHD via the FC of the hippocampus were newly revealed in this study. As a result, the hippocampal FC may further serve as an imaging biomarker for ADHD.

KEYWORDS

ADHD, SNAP-25, functional connectivity, hippocampus, fMRI

1. Introduction

One of the most common mental conditions, attention-deficit/hyperactivity disorder (ADHD), affects almost one in 20 children and adolescents globally and is characterized by fundamental symptoms of hyperactivity, impulsivity, and inattention (Bauermeister et al., 2007; Polanczyk et al., 2007; Mohammadi et al., 2021). ADHD is associated with various structural and functional abnormalities in the hippocampus, a brain region known for its importance in memory functions (Rapport et al., 2008; Kasper et al., 2012; Irwin et al., 2021).

The hippocampus plays a crucial role in spatial navigation and consolidation of information from short-term to long-term memory (Bird and Burgess, 2008; Zhong et al., 2020). Studies have reported mixed findings on hippocampal morphology in ADHD patients, with some suggesting larger volumes as a compensatory reaction to impaired temporal processing (Plessen et al., 2006), while others have related lower hippocampal volumes to more severe ADHD symptomatology (Papadopoulos et al., 2021).

As research progresses, increasing attention has been placed on the relevance of genetic factors and their relationship to ADHD, particularly through the application of neuroimaging techniques, such as resting-state functional magnetic resonance imaging (rs-fMRI), structural MRI, and diffusion tensor imaging (DTI). These studies have investigated various gene polymorphisms that could influence brain function, connectivity, and structure related to ADHD. For instance, certain dopaminergic genes, like DAT1 (SLC6A3) and DRD4, have shown significant associations with ADHD (Shaw et al., 2007; Bralten et al., 2013). rs-fMRI investigations found altered fronto-striatal FC in the presence of these gene polymorphisms, which is partly consistent with the dopamine hypothesis of ADHD (Durston et al., 2004; Posner M. I. et al., 2014). It is also worth mentioning the alterations found in striatal volumetry, where the volumes of the caudate and putamen have been tied to the DAT1 genotype (Greven et al., 2015).

Another gene of interest in ADHD neuroimaging is the brain-derived neurotrophic factor (BDNF) gene. Reductions in gray matter volume in prefrontal and limbic structures were associated with the BDNF Val66Met polymorphism, which plays a role in neuronal survival, growth, and differentiation (Gerritsen et al., 2012). Moreover, this polymorphism showed altered default mode network connectivity in individuals with ADHD, providing evidence for the involvement of BDNF in ADHD-specific alterations in FC (Lawrie, 2020; Woelfer et al., 2020).

Additionally, the catechol-O-methyltransferase (COMT) gene has been investigated in the context of ADHD. The COMT Val158Met polymorphism is associated with altered fronto-striatal connectivity (Nackley et al., 2006), and it has been linked to cognitive performance and behavioral ratings in ADHD children (Hoogman et al., 2013). Furthermore, COMT has been found to modulate the influence of the DAT1 gene on striatal volumes, suggesting a potential interaction between the two genes (Onnink et al., 2015).

Similarly, serotonin-related genes, such as the serotonin transporter (5-HTT) gene and the serotonin 2A receptor (HTR2A) gene, have also been examined, as imbalances in the serotonergic system are thought to contribute to ADHD (Oades, 2007). A study by van Rooij et al. (2015) found that both the 5-HTTLPR polymorphism of the 5-HTT gene and the T102C polymorphism of the HTR2A gene were associated with alterations in the fronto-insula-parietal network.

Furthermore, the SNAP-25 gene, encoding the SNAP-25 protein, has been shown to play a significant role in the process of synaptic vesicle fusion. This process is crucial for communication between neurons, which in turn is critical for memory and learning (Jahn and Scheller, 2006). The protein's role is particularly highlighted in the hippocampus, a brain region known for its importance in the formation of new memories. SNAP-25's involvement in synaptic plasticity- a key mechanism in learning and memory- has been reported in several studies. For instance, a study by Osen-Sand et al. (1996) discovered that SNAP-25 is a critical component of the synaptic vesicle fusion machinery that enables the fast, calcium-triggered release of neurotransmitters. Furthermore, SNAP-25 levels were found to be elevated in the hippocampus of rats during learning tasks, suggesting their involvement in memory consolidation (Hardingham et al., 2010). The protein's impact on axonal growth could also indirectly affect memory and learning processes. A study by Martinez-Arca et al. (2003) showed that SNAP-25 regulates axonal elongation and specification, suggesting its role in the establishment of neural networks, which form the physical substrate of memory. Moreover, SNAP-25's association with ADHD has been linked to its role in cognitive functions, including memory. ADHD is characterized by attention deficits and hyperactivity but often involves impaired memory function as well. Therefore, the association between SNAP-25 and ADHD could provide an indirect line of evidence for the role of SNAP-25 in memory and learning (Gizer et al., 2009). A study by Gosso et al. (2006) further strengthens the case for the involvement of SNAP-25 in human intelligence. By studying two different Dutch cohorts, the researchers found that variations in the SNAP-25 gene were associated with differences in human intelligence. Since intelligence is closely tied to learning and memory, this provides additional evidence for SNAP-25's role in these processes.

While there have been some studies investigating hippocampus-related functional connectivity (FC) alterations in ADHD (Posner J. et al., 2014; Kowalczyk et al., 2022), the relationship between hippocampal FC and ADHD symptomatology requires further exploration. Our study aims to expand existing knowledge with a focus on hippocampal FC alterations and their connection to deficiencies in working memory. Furthermore, we consider the potential effects of the SNAP-25 MnlI variant on hippocampal FC in children with ADHD, building on previous findings that demonstrated significant relationships between rs3746544 polymorphisms and brain connectivity, as well as working memory (Wang et al., 2018; Yang et al., 2022).

Consequently, the SNAP-25 gene has been shown to play a role in ADHD. In this study, we further investigate its impact by categorizing ADHD patients into subtypes based on their SNAP-25 genotypes. We hypothesize that these genotypic subtypes might exhibit different patterns of functional connectivity, thus providing a more nuanced understanding of ADHD.

We expect our findings to provide additional insights into the underlying neurobiological mechanisms linking SNAP-25 gene variations with ADHD symptomatology and working memory deficits. This deeper understanding may not only enhance the accuracy of ADHD diagnosis but also contribute to the development of personalized intervention strategies. For instance,

if the effects of the SNAP-25 MnlI variant on hippocampal FC are found to be significant, targeted therapies could be designed to address these changes and improve working memory in children with ADHD.

2. Materials and methods

2.1. Participants

This study recruited 60 children with ADHD aged between 7 and 10 years from Shenzhen Children's Hospital through a carefully planned collaboration with local schools and pediatric clinics. Our primary goal was to investigate the relationship between hippocampal functional connectivity (FC) alterations and ADHD symptomatology, as well as assess the impact of these alterations on working memory performance in children with ADHD. Prior to enrollment, clinicians at the Shenzhen Children's Hospital identified potential participants and informed their parents about the research study. All participants and their parents were interviewed by experienced clinicians to confirm or exclude a diagnosis of ADHD or any other psychiatric disorder using a clinical interview and the Schedule for Affective Disorders and Schizophrenia for School-Age Children–present and lifetime version (K-SADS-PL) (Kaufman et al., 1997), based on the DSM-V criteria (Association and Others, 2013).

Children with ADHD were required to meet the following inclusion criteria: (1) 7–10 years old, (2) educated in private or public schools, and (3) diagnosed with ADHD. Healthy control subjects had the same age and education requirements as ADHD subjects. The exclusion criteria for both groups included a history of head injury with loss of consciousness, severe physical disease or neurological abnormalities, drug or substance misuse, full-scale IQ measured by Wechsler Intelligence Scale for Chinese Children-IV (WISC-IV-Chinese) below 70, prescription medications for ADHD or other medical conditions used over the long term, and comorbid conduct disorder or Oppositional Defiant Disorder (ODD).

To better understand the relationship between ADHD symptoms and participant characteristics, we assessed ADHD symptomatology using the Conners' Parent Rating Scale (Conners et al., 2011), which encompassed various factors such as delinquent behaviors, learning problems, psychosomatic disorders, hyperactivity, anxiety, and impulsivity.

The MRI scans were only performed on participants who were right-handed dominant, had no visible abnormalities on their MRI images, and did not have a history of claustrophobia. ADHD participants presented with six or more inattentive symptoms as well as six or more hyperactive/impulsive symptoms, and in subsequent statistical analysis, the summing severity scores of each symptom were used as indicators of symptom severity.

As discussed in section "2.3. Genotyping for the detection of SNAP-25 MnII variants" below, we subdivided these ADHD subjects into two subgroups based on their SNAP-25 rs3746544 genotypes, to study the potential correlation between certain genotypes and FC alterations in ADHD symptomatology. The breakdown was as follows:

1. TT homozygotes: 35 participants
2. G-allele carriers (TG): 25 participants

The healthy control group did not undergo genotyping, as our study was focused exclusively around ADHD symptomatology and the potential connection to specific genotypes.

All children and their parents were informed about the purpose and procedures of the study. Children gave their assent to participate, and their parents provided written informed consent on their behalf. The Shenzhen Children's Hospital Medical Ethics Committee approved this study. All methods were performed in accordance with relevant guidelines and regulations.

2.2. Assessments of ADHD symptoms, cognitive function, and clinical outcomes

Attention-deficit/hyperactivity disorder patients, along with TT homozygotes, G-allele carriers, and HCs, underwent an exhaustive set of cognitive and behavioral evaluations. Parents of the participating children completed the Conners' Parent Rating Scale (Conners et al., 2011). This comprehensive questionnaire assesses a variety of behavioral and cognitive concerns, capturing key factors such as learning problems, hyperactivity, anxiety, impulsivity, and delinquent behavior.

Cognitive function was evaluated using the Wechsler Intelligence Scale for Children, Fourth Edition- Chinese version (WISC-IV-Chinese) (Yang P. et al., 2013), a globally accepted intelligence test for children aged 6 to 16 years. The WISC-IV-Chinese, administered by trained professionals, provides a Full-Scale Intelligence Quotient (FSIQ) along with several other indices including the Verbal Comprehension Index (VCI), the Perceptual Reasoning Index (PRI), the Working Memory Index (WMI), and the Processing Speed Index (PSI).

These evaluations are crucial to understanding the cognitive and behavioral profiles of our participants. In this study, we employed the original summary scores for each index from the WISC-IV-Chinese and each factor from the Conners' Parent Rating Scale. Our comprehensive assessments of ADHD symptoms and cognitive functions provide us with a rich dataset that can be correlated with the observed FC alterations in our study. Such correlations may elucidate the clinical implications of the FC alterations we identified and may provide further insights into the symptomatology and cognitive profile associated with ADHD.

2.3. Genotyping for the detection of SNAP-25 MnII variants

Following the manufacturer's guidelines, peripheral venous blood samples were collected from the participants, and SNAP-25 rs3746544 genotyping was carried out using the Flexi Gene DNA Kit (QIAGEN, Germany). Rs3746544 had a forward primer of 5′ TTCTCCTCCAAATGCTGTCG 3′ and reverse primer of 5′ CCACCGAGGAGAAAATG 3′. EX-Taq polymerase and GC buffer (Takara, Dalian, China) were used in a thermocycler to perform the polymerase chain reaction (PCR) amplification (Biometra, Germany). A denaturing cycle at 94°C for 2 min was followed by 30 cycles of 94°C for 30 s, 52°C for 30 s, 72°C for 45 s, and finally an extension step at 72°C for 8 min, in the PCR technique. The TT homozygote group had 35 members (TT group = 35), whereas the TG group had 25 G-allele carriers (TG group = 25).

We decided to concentrate on the TT homozygotes and the TG G-allele carriers. The distribution of these genotypes in our sample is approximately balanced (TT group = 35, TG group = 25). This distribution was not designed to reflect population prevalence or vulnerability to ADHD. Instead, it aimed to provide a balanced dataset for comparing functional connectivity patterns.

2.4. Resting-state fMRI data acquisition

The Radiology Department of Shenzhen Children's Hospital in Shenzhen, China, used a 3.0-T system scanner (Siemens Magnetom Skyra) to collect rs-fMRI data from each participant. The following parameters were used in the echo-planar imaging (EPI) process to acquire the rs-fMRI data: repetition time (TR) = 2000∼ms; echo time = 30∼ms; flip angle = 90°; matrix size = 64 \times 64; 32 axial slices; field of view = 24 \times 24 cm^2; slice thickness = 3 mm; no gap. Structure 3D-MPRAGE: T1 Repetition Time [TR, ms] = 2300 ms, Echo Time [TE, ms] = 2.26; Number of Averages = 1.0, Slice Thickness = 1.0 mm, Field of View (FOV) = 256 mm.

2.5. Data pre-processing

The DPARSF (v5.1) toolkit (Yan et al., 2016) was used to preprocess the data using SPM12.[1] The image preprocessing in the analysis as follow: the first 10 volumes were dropped because of the instability of the initial magnetic resonance imaging data and the participants' adaptation to the experimental apparatus. First, the slice time corrected, remaining 230 volumes were realigned to account for the head movement. The head motion criterion of translation <3 mm or rotation <3° in any direction was used to retain all patients with ADHD. Subsequently, the data were normalized and resampled into 3 mm× 3 mm × 3 mm voxels. The time course of each voxel was regressed to remove unwanted factors such as the global signal, white matter signal, cerebrospinal fluid signal, and Friston-24 parameters of head motion. To lessen the impact of low-frequency drift and high-frequency physiological noise, the data were linearly detrended, filtered at 0.01–0.08 Hz, and smoothed with a 6 mm full-width-at-half-maximum Gaussian kernel. The calculation of frame-wise displacement (FD) across the time point for each participant was then performed to evaluate head motion. The bad time point and its 1-back and 2-forward volumes were finally estimated by cubic spline interpolation using scrubbing methods (Power et al., 2012) at an FD threshold of 0.5 mm.

2.6. Head of motion

Jenkinson et al. (2002) relative root-mean-square method was used to eliminate the mean FD produced during the scanning procedure. The mean FD (Jenkinson) was calculated to assess voxel-wise motion differences between the three groups. The mean FD did not differ significantly between the [HC (0.05 ± 0.02), TT homozygotes (0.05 ± 0.04) and G-allele carriers (0.05 ± 0.1)] groups ($P < 0.6$).

1 https://www.fil.ion.ucl.ac.uk/spm/software/spm12/

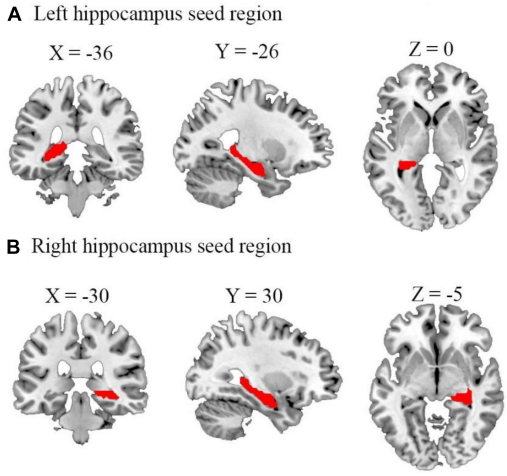


FIGURE 1
Representation of the seed regions for left and right hippocampus. The figure shows axial, coronal, and sagittal views of the brain with seed regions for **(A)** left and **(B)** right hippocampus highlighted. These regions were selected as the regions of interest for our functional connectivity analyses. Coordinates are presented in the Montreal neurological institute (MNI) standard space (left hippocampus: $x = -36$, $y = -26$, $z = 0$; right hippocampus: $x = -30$, $y = 30$, $z = -5$).

2.7. Hippocampus FC statistical analysis

To compare the FC maps between the three groups (TT, TG, and HC), one-way ANOVA was performed. Age and FD were regressed as covariates. The basic threshold used for the data provided here was voxel-wise $p < 0.001$, cluster-level $p < 0.05$, and GRF-adjusted. Then, by averaging the Z-scores of each peak coordinate, we identified the functional connectivity signals that revealed distinct variations across the three groups. Brain regions that had undergone numerous compression corrections were subjected to *post hoc* analysis of two-tailed ANOVA tests to ascertain the direction of FC change between the three. Statistical significance was defined as $p < 0.05/4$ (Bonferroni's corrected). Using the Montreal Neurological Institute (MNI) atlas, a standardized and internationally recognized spatial framework, we identified the exact location and boundaries of the hippocampus. We aligned our subjects' brain images to this atlas, effectively superimposing the predefined hippocampal region onto each subject's brain image. We then visually confirmed the alignment and made necessary adjustments to ensure an accurate fit. This rigorous process ensured the consistency of our hippocampal ROI across all subjects. **Figure 1** illustrates the process of alignment and confirmation. **Table 2** provides a summary of the different brain areas.

2.8. Partial correlation analysis

The FC values that significantly deviated from the baseline were extracted, and partial correlation analysis was used to examine the relationship between the altered FC (FC values for FSIQ and the mean value of the right precuneus) and ADHD symptom severity, working memory scores, and Conners' scale scores. This was done

while controlling for nuisance covariates such as age, grade, and mean FD. An uncorrected $p < 0.05$ threshold was used as the statistical significance criterion.

It is important to note that IQ was not included as a covariate in this analysis. While IQ is often used as a covariate in many neurodevelopmental studies, there are compelling reasons to avoid such an approach. As argued by Dennis et al. (2009), IQ scores are highly dynamic, reflecting an individual's overall functional outcomes shaped by a myriad of factors, including genetics, biological status, cognitive capabilities, educational attainment, and personal experiences. Therefore, employing IQ as a covariate could potentially result in an "overcorrection" of our data, yielding skewed or counterintuitive results. Given these methodological considerations, we opted not to control for IQ in our analysis.

2.9. ROC analysis

The feasibility of employing z-variance as a diagnostic biomarker for differentiating patients with ADHD from HCs was evaluated using a support vector machine (SVM) built into the LIBSVM library [LIbrary for Support Vector Machines (Chang and Lin, 2011)]. Based on the data from the primary sample, ROIs were specifically determined from group analysis.

Our ADHD patient groups utilized in the ROC analysis comprised two genotypes: 35 TT homozygotes and 25 G-allele carriers (TG). For the SVM classification, these ADHD patient groups were split into two: one for training the classifier ($n = 30$) and another for testing ($n = 30$). These groups were randomly selected from our total pool of 60 ADHD patients and were statistically comparable in terms of demographic information and clinical characteristics. As detailed in **Table 1**, which provides a breakdown of demographic and clinical characteristics for the training and testing groups, we found no significant differences between these two ADHD groups ($p > 0.05$ for all variables), confirming the comparability of these groups for SVM classification.

Although the genotypes were not differentiating factors in the classifier training, they represent significant data that contribute to a comprehensive understanding of ADHD symptomatology and FC alterations.

For feature selection, the average ROI value for each participant in the primary sample was retrieved and utilized as a feature. Prior to SVM classification, feature data were normalized within each ROI using z-score normalization (mean subtraction followed by standard deviation scaling). This normalization process ensured that features had equal weighting and did not introduce bias to the classification performance.

To train the classifier, the normalized averaged ROI values were used. In order to distinguish patients with ADHD from HCs in the independent sample, the averaged z-variance values of the same ROIs were obtained for each participant, normalized, and then input into the classifier as features.

The statistical significance of the classification performance [area under the curve (AUC) of the receiver operating characteristic curves (ROC)] was determined using a non-parametric permutation test. The actual group labels (ADHD and HC) were randomly shuffled in each permutation test trial, and the same classification process was used to determine the classification accuracy score based on the shuffled dataset. This process was repeated 5,000 times to assess the level of statistical significance and obtain the *p*-value.

3. Results

3.1. Demographic characteristics and clinical variables and clinical variables

Three TT homozygotes patients were excluded from further analyses due to excessive head motion. The final cohort contained 35 TT homozygotes, 25 G-allele carriers, and 28 HC (**Table 2**). No significant differences are found in term of age, gender,

TABLE 1 Demographic and clinical characteristics of ADHD training and testing sets.

Variables	Training set ($n = 30$)	Testing set ($n = 30$)	Statistics	*p*-value
Age, mean ± SD	8.4 ± 0.81	8.6 ± 0.79	$t = 0.8$	0.43
Sex (male)	M	M	$\chi^2 = 0.0$	1.00
Grade, mean ± SD	2.9 ± 0.51	2.7 ± 0.55	$t = 1.2$	0.23
IQ scores, mean ± SD	85 ± 7.05	86 ± 8.61	$t = 0.6$	0.55
WMI, mean ± SD	89.5 ± 9.31	88.2 ± 9.75	$t = 0.5$	0.61
VCI, mean ± SD	82.8 ± 8.61	83.5 ± 8.41	$t = 0.3$	0.76
PRI, mean ± SD	94.9 ± 11.50	96.2 ± 11.75	$t = 0.4$	0.69
PSI, mean ± SD	93.5 ± 11.40	92.9 ± 11.80	$t = 0.2$	0.84
Delinquent behaviors, mean ± SD	1.13 ± 0.51	1.15 ± 0.49	$t = 0.2$	0.83
Learning problem, mean ± SD	1.95 ± 0.57	1.98 ± 0.59	$t = 0.2$	0.84
Psychosomatic disorder, mean ± SD	0.27 ± 0.31	0.26 ± 0.30	$t = 0.1$	0.91
Hyperactivity, mean ± SD	1.62 ± 0.68	1.61 ± 0.65	$t = 0.1$	0.92
Anxiety, mean ± SD	0.67 ± 0.56	0.68 ± 0.55	$t = 0.1$	0.91
Impulsivity, mean ± SD	1.59 ± 0.49	1.60 ± 0.48	$t = 0.1$	0.92

ADHD, Attention deficit hyperactivity disorder; HC, Healthy control; SD, Standard deviation; WMI, Working Memory Index; VCI, Verbal Comprehension Index; PRI, Perceptual Reasoning Index; PSI, Processing Speed Index. The statistics column refers to *t*-values obtained by independent samples *t*-test (for continuous variables) or χ^2-values obtained by chi-square test (for categorical variables).

TABLE 2 Characteristics of demographics of the three genotypic groups.

Variables	HC (n = 28)	TT homozygotes (n = 35)	G-allele carriers (n = 25)	Statistics	p-value
Age, mean ± SD	8.9 ± 0.97	8.31 ± 0.79	8.84 ± 0.71	F = 0.8	0.01[1]
Sex (male)	M	M	M		
Grade, mean ± SD	3.03 ± 0.83	2.8 ± 0.45	2.6 ± 0.57	F = 0.7	0.02[1]
IQ scores, mean ± SD	90 ± 8.40	84 ± 6.93	87 ± 9.79	F = 1.9	0.04[1]
WISC-IV-Chinese					
Working memory index (WMI), mean ± SD	97.64 ± 9.95	90.11 ± 8.70	85.02 ± 9.87	F = 24.80	0.05[1]
Verbal comprehension index (VCI)	103.61 ± 8.82	82.11 ± 7.91	83.25 ± 11.93	F = 48.130	0.048[1]
Perceptual reasoning index (PRI)	109.36 ± 11.61	96.78 ± 11.60	93.85 ± 10.76	F = 13.737	0.07[1]
Processing speed index (PSI)	103.64 ± 10.67	93.72 ± 11.24	92.35 ± 14.36	F = 7.263	0.02[1]
Corners' parent rating scale					
Delinquent behaviors	0.63 ± 0.21	1.08 ± 0.55	1.18 ± 0.47	F = 8.322	0.05[1]
Learning problem	0.74 ± 0.25	2.03 ± 0.54	1.86 ± 0.63	F = 43.050	0.032[1]
Psychosomatic disorder	0.15 ± 0.16	0.27 ± 0.31	0.26 ± 0.33	F = 1.200	0.307[1]
Hyperactivity	0.61 ± 0.31	1.59 ± 0.70	1.64 ± 0.64	F = 20.153	0.050[1]
Anxiety	0.45 ± 0.21	0.66 ± 0.55	0.68 ± 0.58	F = 1.444	0.243[1]
Impulsivity	0.44 ± 0.22	1.61 ± 0.51	1.47 ± 0.46	F = 18.660	0.02[1]

HC, healthy control; ADHD, attention deficit hyperactivity; SD, standard deviation.
[1]One-way analysis of variance. F-values were obtained by 1-way ANOVA.

education level, handedness, mean FD, and the variance of FD among the three groups (Table 2). A total of 35 TT homozygotes, 25 G-allele carriers, and 28 HC groups differed significantly in clinical variables (Table 2). Patient demographic and descriptive statistics are listed in Table 2. Based on the psychometric evaluations presented in Table 2, there are some noticeable differences between the TT homozygotes and the G-allele carriers. TT homozygotes demonstrated lower scores in all WISC-IV-Chinese indices compared to HC, and G-allele carriers showed even lower scores in the Working Memory Index. In terms of behavioral assessments from Conners' Parent Rating Scale, both TT homozygotes and G-allele carriers exhibited higher levels of delinquent behaviors, learning problems, hyperactivity, and impulsivity compared to HC, while no significant differences were found in the manifestation of psychosomatic disorder and anxiety.

3.2. Main effect diagnosis FC between the three groups TT, TG, and HCs in the right/left

The 1-way ANOVA results indicated significant differences in FC variability among the three groups (TT, TG, and HCs) for the hippocampus seed regions (Figure 2 and Table 3; voxel $P < 0.001$, cluster $P < 0.05/2$, controlling for age, grade level, and the mean FD). Significant differences were observed between the three groups in both the left and right hippocampus seeds. Notably, significant differences in FC were found between the right hippocampus seed and right precuneus (Figure 2A), as well as between the left hippocampus seed and left middle frontal gyrus (Figure 2B).

To further examine these differences, we conducted *post hoc* two-sample t-tests for each significantly different region (Figures 2A, B) among the two groups. Table 3 presents the significant effects on FC for both the right and left hippocampus seeds. A comparison of the TT and HC groups revealed a decreased FC between the right hippocampal seed and precuneus, whereas the TG group exhibited a decreased FC in the right hippocampal seed when compared to the HC group (Figure 2A). Additionally, the TT group demonstrated a decrease in FC between the left middle frontal gyrus and the left hippocampus seed when compared with the HC group (Figure 2B). In contrast, the TG group displayed increased FC between the left hippocampus seed and left middle frontal gyrus compared with the HC group as shown in Figure 2B.

3.3. Correlation between FC and clinical variables

As shown in Figure 3, the partial correlation analysis, controlling for age, showed that altered FC between the right hippocampus and right precuneus in the TG group was negatively correlated with IQ scores ($r = -0.40$, $p = 0.04$) (Figure 3A). It also showed that altered FC between left hippocampus and left middle frontal gyrus in TT group was negatively correlated ($r = -0.6$, $p = 0.02$) (Figure 3B). No significant correlations were found between other changes in the FC and other regions when accounting for the control covariates.

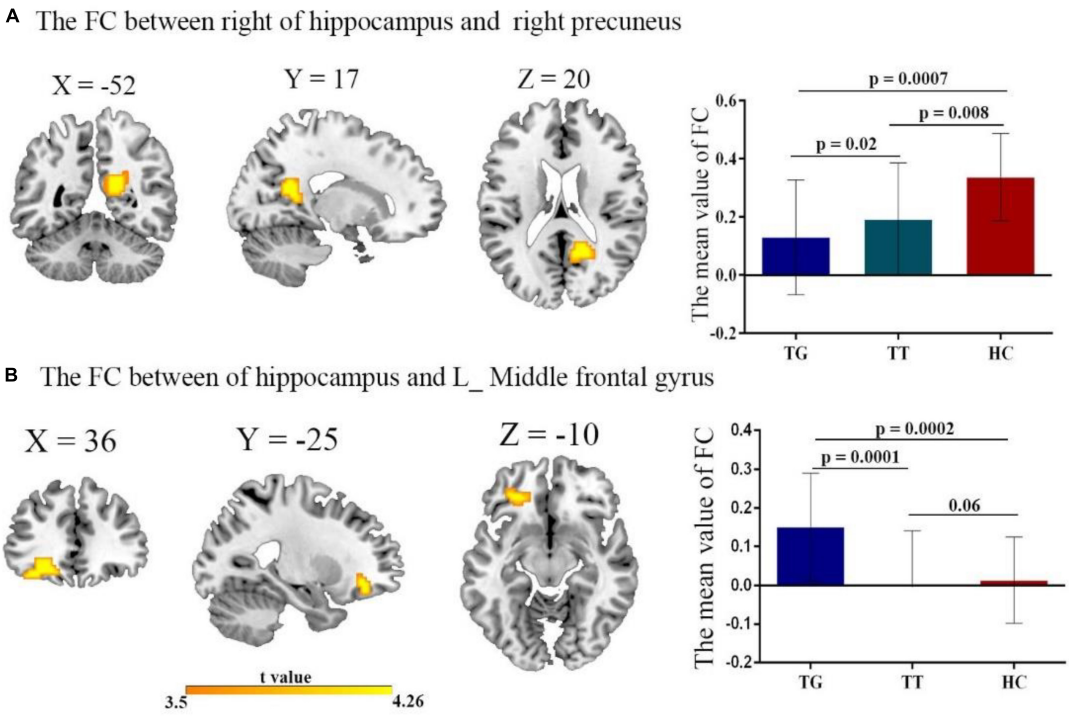


FIGURE 2
The FC of the left/right hippocampus with brain regions after comparing all three groups TT, TG, and HC with *t*-value of the FC connectivity variance; **(A)** right hippocampus with right precuneus. Decreased FC in both TT group, and TG group, compared with HC group. **(B)** Left hippocampus with left middle frontal gyrus. Reduced FC in both TT and TG group.

TABLE 3 Brain clusters showing a significant effect in the FC with right and left of the hippocampus.

Seed region	Group differences	Cluster size	Z-score	M N I	TT homozygotes (*n* = 35)	G-allele carriers (*n* = 25)	HC (*n* = 28)
				X Y Z	M ± SD	M ± SD	M ± SD
Right hippocampus	Right precuneus	60	−5.09	16−56.18	0.19 ± 0.20	0.33 ± 0.19	0.33 ± 0.14
Left hippocampus	Left middle frontal gyrus	59	3.3	−27 ± 38.13	−0.04 ± 0.14	0.15 ± 0.10	0.01 ± 0.10

SD, standard deviation; M, mean value.

3.4. ROC analysis

Receiver operating characteristic analysis confirmed that HCs and patients with ADHD can be distinguished based on FC between the right hippocampus, right precuneus, and left middle frontal gyrus. As shown in **Figure 4**, the AUCs of these connection variability were significantly higher than those expected by chance (Bonferroni's correction). We obtained AUCs proportions between 0.7 and 0.82, which introduces an acceptable discrimination.

4. Discussion

In this study, we built upon our previous research (Wang et al., 2018; Yang et al., 2022) to investigate the hippocampal involvement in working memory deficits in children with ADHD, considering the role of SNAP-25 MnII variants. Using a novel diagnostic subtyping based on SNAP-25 MnII variant genotyping, we identified distinct ADHD patient groups: ADHD-TT and ADHD-TG. Our results highlighted altered hippocampal FC among the three groups (TT, TG, and HCs), with both TG and TT patients displaying decreased FC between the right hippocampus and precuneus compared to HCs. Furthermore, we found that TG patients exhibited greater FC between the left hippocampus and left middle frontal gyrus than both TT patients and HCs. In our discussion, we delve into these findings and their implications on cognitive functioning and IQ-scores in childhood ADHD.

4.1. Bilateral heterogeneous hippocampus FC profiles and their effects on cognitive functioning in children with ADHD

Memory and spatial navigation are both attributed to the hippocampus. The left and right hippocampi have often been

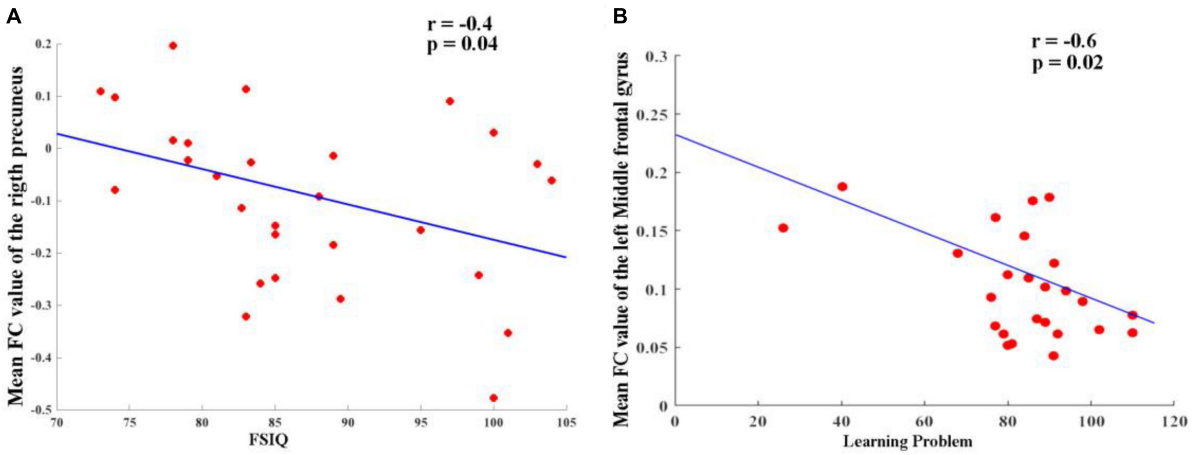


FIGURE 3
(A) Altered FC between right hippocampus and right precuneus in TG group and it was negatively correlated ($r = -0.40$, $p = 0.04$). **(B)** Altered FC between left hippocampus and left middle frontal gyrus in TT group and it was negatively correlated ($r = -0.6$, $p = 0.02$).

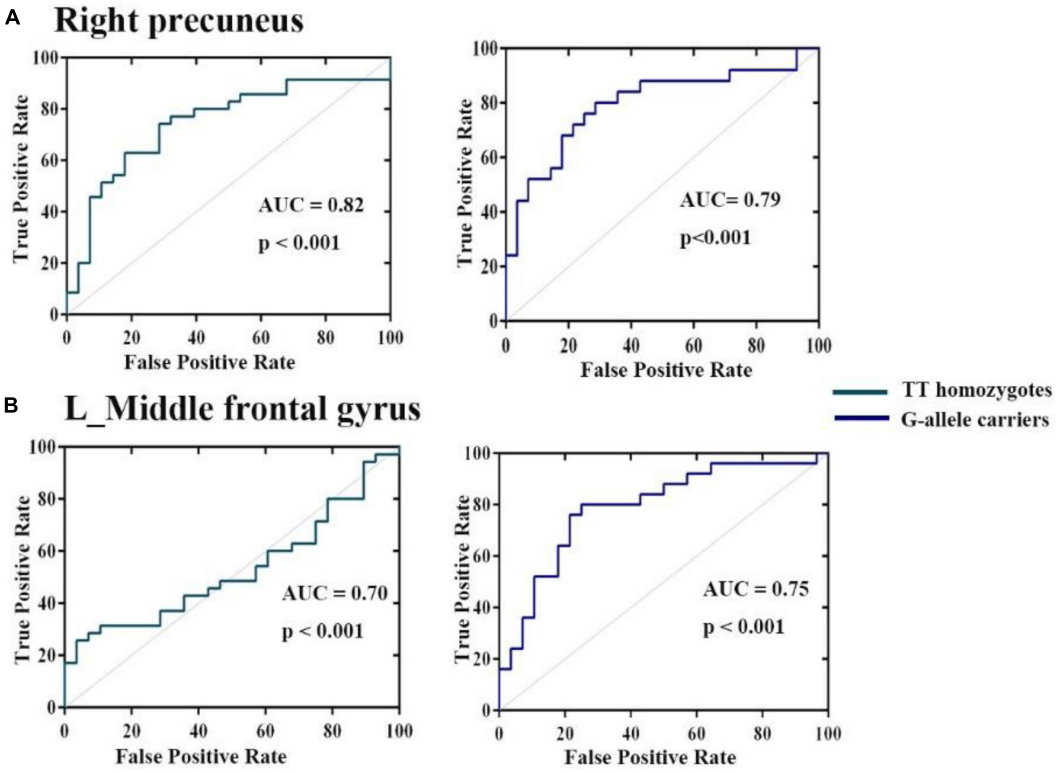


FIGURE 4
Operating characteristic (ROC) curves for discrimination between three groups for mean variance of FC **(A)** right hippocampus **(B)** left middle frontal gyrus. The AUCs of these ROIs were higher than those expected by chance (Bonferroni's corrected).

regarded as functionally equal in rats, and this bilateral brain region has been widely researched. Recently, the molecular and morphological properties of the neural connections in the brain hemisphere have been found to exhibit unanticipated asymmetries (Shipton et al., 2014). Similarly, Gu et al. (2022) found that only aberrant left hippocampal connectivity was related to cognitive function in patients with relapsing-remitting multiple sclerosis (RRMS). Our study supports two major interrelated findings: (1) we observed a bilateral distinct hippocampal FC profile with heterogeneous connections to different brain regions, and (2) these FCs may impact cognitive functioning in children with ADHD, taking into account the obtained IQ scores.

We found altered FC between the right hippocampus and the (right) precuneus in both the TT and TG patient groups compared to the healthy control group (HCs). Disrupted FC between the left hippocampus and the (left) middle frontal gyrus was observed

in patient groups. The precuneus, involved in a wide range of complex tasks such as memory, information integration, mental imagery, and affective reactions to pain, has been extensively reported in correlation with hippocampal FC, especially regarding mild cognitive impairment and Alzheimer's disease (Kim et al., 2013; Xue et al., 2019).

Ren et al. (2018) highlighted the reciprocal connectivity between the hippocampus and precuneus, emphasizing their relevance to metacognition in settings similar to daily life. Our study added to this evidence by suggesting a negative correlation between changes in FC between the right hippocampus and the right precuneus in the TG group and reduced IQ scores ($r = -0.40$, $p = 0.04$). In addition to the previously mentioned findings, our study has further identified a significant negative correlation between altered FC in the left hippocampus and the left middle frontal gyrus in the TT group ($r = -0.6$, $p = 0.02$). This adds another dimension to our understanding of how ADHD may impact brain connectivity and cognitive abilities. It is becoming increasingly clear that genetic variations in ADHD patients play a pivotal role in their brain function, contributing to the symptomatic heterogeneity observed in this disorder. The results underscore the importance of considering the interplay of genetic factors and functional connectivity when examining the complex manifestations of ADHD and its associated cognitive deficits.

The left middle frontal gyrus is crucial for literacy development. In Chinese reading, the left middle frontal gyrus serves as a specialized hub region that connects ventral and dorsal pathways (Guo et al., 2022). In dyslexic Chinese readers, structural and functional impairments in the left middle frontal gyrus contrast with the left temporoparietal regions in alphabetic languages (Yang Y. et al., 2013). This might explain why we observed disrupted FC between the hippocampus and the left middle frontal gyrus in our Chinese ADHD patients. Future research should further explore the implications of these FC alterations on other aspects of cognitive and behavioral functioning, as well as their potential as targets for intervention. Our findings suggest that different genetic groups within ADHD may require tailored treatment approaches taking into account their specific patterns of brain connectivity.

Regarding the cognitive implications of these FC alterations, the co-occurrence of lower IQ scores and academic achievement deficits and learning problems in ADHD children is well-established. We observed decreased FC values between the left hippocampus and the left middle frontal gyrus in the TT group compared to the TG group, although not statistically significant. Similar findings were reported by Siok et al. (2009), who discussed how dyslexic reading in Chinese individuals is represented by phonological impairments assessed through the weak activity of the left middle frontal gyrus in a rhyme-judgment task.

These altered hippocampus-precuneus-middle frontal gyrus connections may clarify the compromised attention-related encoding and retrieval processes leading to cognitive deficits in ADHD (Ortega et al., 2020). A thesis by Roya (2016) hypothesized that ADHD patients possess decreased resting-state activity in pathways through the hippocampus because of reduced volume and executive functioning frequency. However, that study revealed no abnormal connectivity in the hippocampus. These contradicting results highlight the need for further hippocampus research in ADHD populations.

4.2. SNAP-25 as a sensitive marker of disrupted patterns of connectivity in ADHD

Our first main finding was that FC alterations within the hippocampal FC in Chinese children with ADHD are genetically driven through observations of SNAP-25 variants. If replicated, this result indicates that SNAP-25 is a sensitive marker of disrupted connectivity patterns in ADHD. Although the relationship between SNAP-25 and altered structure or connectivity patterns has been repeatedly reported in patients with bipolar disorder (Houenou et al., 2017), autism (Braida et al., 2015), schizophrenia, and major depressive disorder (Najera et al., 2019), it is remarkable that only a few studies have investigated the implication of SNAP-25 in connectivity alterations in ADHD (Wang et al., 2018; Yang et al., 2022). One possible reason for the relationship between SNAP-25 and these psychiatric disorders is that SNAP-25 has a genetic basis that is linked to some symptoms that co-occur and are common among these disorders. For the SNAP-25 itself, although changes in neurotransmitter release have been suggested as potential causative processes, the mechanisms by which abnormalities in SNAP-25 may contribute to certain mental illnesses, including ADHD, remain unclear. Intriguingly, in line with our findings, Braida et al. (2015), suggested that the SNP rs363050 has a regulatory region based on analysis of transcriptional activity, which resulted in a reduction in protein expression. This reduction in protein expression affected the teenage mice, whose levels of SNAP-25 were reduced. Accordingly, they exhibited hyperactivity, cognitive and social dysfunction, and irregular EEG signals with numerous spikes.

On the other hand, Houenou et al. (2017) demonstrated that compared to non-risk carriers (of a promoter variant in SNAP25, rs6039769 at-risk allele), male risk carriers had a larger amygdala and increased FC between the amygdala and ventromedial prefrontal cortex. Therefore, this study supports the finding that this allelic variation of SNAP25 has a functional effect on modulating the development and plasticity of the prefrontal-limbic network, which may increase vulnerability to both early onset bipolar disorder and schizophrenia. Based on these studies, to better understand the abnormalities of the hippocampal FC in ADHD patients, we also need to analyze more about the genetic variations of SNAP-25 and ADHD.

The presynaptic plasma membrane protein SNAP-25 is abundantly and selectively expressed in nerve cells (Söllner et al., 1993). Considering the functions of SNAP-25, it is likely that any variation in this protein, which is mainly and distinctively encountered in axons and nerve terminals (Wang et al., 2014), can influence vulnerability to ADHD by affecting neurotransmitter release and the formation of neural circuits throughout the central nervous system (CNS). Based on its physiological significance in the docking and fusion of synaptic vesicles in presynaptic neurons as well as in axonal growth and synaptic plasticity, SNAP-25 is a potential candidate gene for ADHD. The C allele of rs1051312 is increasingly being transmitted in Canada cases (Barr et al., 2000). Using a transmission disequilibrium test (TDT), Brophy et al. (2002) demonstrated favored transmission of the T allele of rs1051312 among Irish ADHD patients. Chinese (Gao et al., 2009) and Colombian (Gálvez et al., 2014) populations showed a

significant association between rs3746544 (1065T > G) and ADHD in case-control studies, but the Irish (Brophy et al., 2002), Indian (Sarkar et al., 2012), Canadian (Barr et al., 2000), US Caucasian (Feng et al., 2005), and UK Caucasian (Mill et al., 2004) groups showed no such association. In two separate samples of families with ADHD, 12 SNPs were examined by Feng et al. (2005). They discovered significant over-transmission of the rs66039806-C, rs362549-A, rs362987-A, and rs362998-C alleles in a Canadian sample but not in a southern California sample. These alleles were located in introns 2, 4, and 6. When they used quantitative analysis to assess a Canadian population for the behavioral ADHD subtypes of inattention and hyperactivity, they discovered relationships between both categories and SNAP-25. However, several studies (Ilott et al., 2010) have found no evidence to support a link between these polymorphisms and ADHD.

However, consistent with our findings, SNAP-25 was found to encode a protein that is crucial for synaptic vesicle fusion and neurotransmitter release. Furthermore, recent studies have indicated that SNAP-25 is involved in learning and memory, two processes essential for human cognition and intelligence (Noor and Zahid, 2017). Single nucleotide polymorphisms (SNPs) in genes associated with cognitive function have been reported in patients with ADHD. Barr et al. (2000) were the first to detect the MnlI polymorphism (rs3746544), a relatively often researched SNP found in the 3′-untranslated region (3′-UTR), which is linked to ADHD. Early in Chen et al. (2008) indicated that SNPs in the 3′-UTR constitute a crucial microRNA-binding site and may alter binding sites while demolishing the operating site or generating another illegitimate site (Chen et al., 2008), which could impact the expression of the SNAP-25 gene and may ultimately increase the susceptibility for the progression of ADHD (Németh et al., 2013; Ye et al., 2016).

Although our study does not provide direct clinical applications of SNAP-25 genotyping in ADHD, it suggests the possibility of subtype-specific FC patterns. This might mean that TT homozygotes and G-allele carriers could have differences in ADHD presentation, which could potentially inform future diagnostic and treatment strategies. Further research is necessary to verify these findings and explore their clinical implications in more detail.

4.3. Limitations

We acknowledge that our study has several limitations despite providing evidence from many levels. First, the sample size of the fMRI study was rather small; therefore, caution should be exercised when interpreting data. Whole-brain structural and connectivity analyses could potentially be performed using independent samples with a larger sample size. Second, replication of our findings is problematic because of the small number of accessible brain samples. The investigation of SNAP-25 expression levels in various regions depending on the genotype would be of significant interest because of the critical role of the hippocampus in ADHD and the regions that have been demonstrated to be functionally connected. However, changes in the SNAP-25 level and its SNARE complex binding partners have frequently been observed in mouse models and humans with psychiatric illnesses, highlighting the significance of this complex's control. Moreover, replication from multiple polymorphisms in subjects among genetically different ethnic groups can also reveal significant findings. Even though the analyses took age into account, further evaluations of individuals who had reached adulthood could provide more light on developmental features. Another limitation is that although we examined the relationship between the hippocampus and cognitive deficits, our study did not take into account a more comprehensive psychological assessment of the cognitive parameters associated with ADHD. Therefore, future research can include more features, particularly those related to children's performance at school or at home. Other patient-specific factors were also reported to limit our study, such as male preponderance, existence of comorbidities, and heterogeneity in clinical presentation, which are typically prevalent in ADHD studies. Therefore, to obtain more accurate results, further research should use more complex patient groups, in which both genetic and symptomatology-based subtypes are considered.

5. Conclusion

A major obstacle in understanding neuropsychiatric illnesses is the functional characterization of disease-associated variants, which will open an avenue for the creation of personalized therapies. According to a growing body of research, the SNARE complex, and more specifically, the SNAP-25 protein, may play a role in mental diseases. Here, as a first step, we were able to corroborate the association of one of the SNAP-25 variants with ADHD by providing modest evidence for the *Mnl*I marker of SNAP-25 with ADHD. Based on genotyping for the detection of SNAP-25 MnII variants, we were able to obtain a new subtyping of patient groups (TT and TG) for further application in functional imaging analysis. In other words, instead of adopting the typical symptomatology-based ADHD subtypes, we use genetics-based subtypes, which we believe provides more accurate findings, especially for heritable and genetic disorders such as ADHD. For our FC analysis, we investigated the role of the hippocampal FC in childhood ADHD and its possible association with cognitive impairment. In view of the influence of SNAP-25 on disease processes, an additional thorough integration of genetics-based investigations with neuroimaging is essential to pave the way for larger, more varied, and in-depth genome-wide association studies.

Data availability statement

The datasets generated and/or analyzed during the current study are not publicly available because of Chinese Ethics Committee regulations but are available from the corresponding author (FW, fwen62@126.com) on reasonable request.

Ethics statement

The studies involving human participants were reviewed and approved by the Medical Ethics Committee of the Shenzhen

Children's Hospital. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

FW, HZ, BY, and AF designed the study. WH, YZ, DF, and LZ collected the data and organized the clinical information. FW, AF, BY, and HZ reviewed the methods. AF, WH, and DF analyzed the data. WH and AF wrote the manuscript. All authors discussed, approved, and proofread the results and the final manuscript.

Funding

This work was supported by a grant from Shenzhen Medical and Health Project (No. SZSM202011005).

Acknowledgments

We would like to thank all individuals who participated in this research.

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

References

Association A. P. and Others (2013). DSM 5 diagnostic and statistical manual of mental disorders. In *DSM 5 diagnostic and statistical manual of mental disorders*. Washington, DC: American Psychiatric Association Publishing. doi: 10.1176/appi.books.9780890425787

Barr, C. L., Feng, Y., Wigg, K., Bloom, S., Roberts, W., Malone, M., et al. (2000). Identification of DNA variants in the SNAP-25 gene and linkage study of these polymorphisms and attention-deficit hyperactivity disorder. *Mol. Psychiatry* 5, 405–409. doi: 10.1038/sj.mp.4000733

Bauermeister, J. J., Shrout, P. E., Chávez, L., Rubio-Stipec, M., Ramírez, R., Padilla, L., et al. (2007). ADHD and gender: are risks and sequela of ADHD the same for boys and girls? *J. Child Psychol. Psychiatry* 48, 831–839.

Bird, C. M., and Burgess, N. (2008). The hippocampus and memory: insights from spatial processing. *Nat. Rev. Neurosci.* 9, 182–194.

Braida, D., Guerini, F. R., Ponzoni, L., Corradini, I., De Astis, S., Pattini, L., et al. (2015). Association between SNAP-25 gene polymorphisms and cognition in autism: functional consequences and potential therapeutic strategies. *Trans. Psychiatry* 5:e500. doi: 10.1038/tp.2014.136

Bralten, J., Franke, B., Waldman, I., Rommelse, N., Hartman, C., Asherson, P., et al. (2013). Candidate genetic pathways for attention-deficit/hyperactivity disorder (ADHD) show association to hyperactive/impulsive symptoms in children With ADHD. *J. Am. Acad. Child Adolesc. Psychiatry* 52, 1204.e1–1212.e1. doi: 10.1016/j.jaac.2013.08.020

Brophy, K., Hawi, Z., Kirley, A., Fitzgerald, M., and Gill, M. (2002). Synaptosomal-associated protein 25 (SNAP-25) and attention deficit hyperactivity disorder (ADHD): evidence of linkage and association in the Irish population. *Mol. Psychiatry* 7, 913–917. doi: 10.1038/sj.mp.4001092

Chang, C. C., and Lin, C. J. (2011). LIBSVM: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.* 2, 1–27. doi: 10.1145/1961189.1961199

Chen, K., Song, F., Calin, G. A., Wei, Q., Hao, X., Zhang, W., et al. (2008). Polymorphisms in microRNA targets: A gold mine for molecular epidemiology. *Carcinogenesis* 29, 1306–1311.

Conners, C. K., Pitkanen, J., and Rzepa, S. R. (2011). "Conners 3rd Edition (Conners 3; Conners 2008)," in *Encyclopedia of clinical neuropsychology*, eds J. S. Kreutzer, J. DeLuca, and B. Caplan (New York, NY: Springer New York), 675–678. doi: 10.1007/978-0-387-79948-3_1534

Dennis, M., Francis, D. J., Cirino, P. T., Schachar, R., Barnes, M. A., and Fletcher, J. M. (2009). Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *J. Int. Neuropsychol. Soc.* 15, 331–343. doi: 10.1017/S1355617709090481

Durston, S., Pol, H. E. H., Schnack, H. G., Buitelaar, J. K., Steenhuis, M. P., Minderaa, R. B., et al. (2004). Magnetic resonance imaging of boys with attention-deficit/hyperactivity disorder and their unaffected siblings. *J. Am. Acad. Child Adolesc. Psychiatry* 43, 332–340. doi: 10.1097/00004583-200403000-00016

Feng, Y., Crosbie, J., Wigg, K., Pathare, T., Ickowicz, A., Schachar, R., et al. (2005). The SNAP25 gene as a susceptibility gene contributing to attention-deficit hyperactivity disorder. *Mol. Psychiatry* 10, 998–1005. doi: 10.1038/sj.mp.4001722

Gálvez, J. M., Forero, D. A., Fonseca, D. J., Mateus, H. E., Talero-Gutierrez, C., and Velez-van-Meerbeke, A. (2014). Evidence of association between SNAP25 gene and attention deficit hyperactivity disorder in a Latin American sample. *ADHD Attent. Deficit Hyperact. Disord.* 6, 19–23. doi: 10.1007/s12402-013-0123-9

Gao, X. P., Su, L. Y., Zhao, A. L., Luo, X. R., and Xia, K. (2009). Association of 14 polymorphisms in the five candidate genes and attention deficit hyperactivity disorder. *Chinese J. Contemp. Pediatr.* 11, 617–622.

Gerritsen, L., Tendolkar, I., Franke, B., Vasquez, A. A., Kooijman, S., Buitelaar, J., et al. (2012). BDNF Val66Met genotype modulates the effect of childhood adversity on subgenual anterior cingulate cortex volume in healthy subjects. *Mol. Psychiatry* 17, 597–603. doi: 10.1038/mp.2011.51

Gizer, I. R., Ficks, C., and Waldman, I. D. (2009). Candidate gene studies of ADHD: a meta-analytic review. *Hum. Genet.* 126, 51–90. doi: 10.1007/s00439-009-0694-x

Gosso, M. F., de Geus, E. J. C., van Belzen, M. J., Polderman, T. J. C., Heutink, P., Boomsma, D. I., et al. (2006). The SNAP-25 gene is associated with cognitive ability: evidence from a family-based study in two independent Dutch cohorts. *Mo. Psychiatry* 11, 878–886. doi: 10.1038/sj.mp.4001868

Greven, C. U., Bralten, J., Mennes, M., O'Dwyer, L., van Hulzen, K. J. E., Rommelse, N., et al. (2015). Developmentally stable whole-brain volume reductions and developmentally sensitive caudate and putamen volume alterations in those with attention-deficit/hyperactivity disorder and their unaffected siblings. *JAMA Psychiatry* 72, 490–499. doi: 10.1001/jamapsychiatry.2014.3162

Gu, X.-Q., Liu, Y., Gu, J.-B., Li, L.-F., Fu, L.-L., and Han, X.-M. (2022). Correlations between hippocampal functional connectivity, structural changes, and clinical data in patients with relapsing-remitting multiple sclerosis: a case-control study using multimodal magnetic resonance imaging. *Neural Regener. Res.* 17, 1115–1124. doi: 10.4103/1673-5374.324855

Guo, W., Geng, S., Cao, M., and Feng, J. (2022). The brain connectome for chinese reading. *Neurosci. Bull.* 38, 1097–1113.

Hardingham, N. R., Read, J. C. A., Trevelyan, A. J., Nelson, J. C., Jack, J. J. B., and Bannister, N. J. (2010). Quantal analysis reveals a functional correlation between presynaptic and postsynaptic efficacy in excitatory connections from rat Neocortex. *J. Neurosci.* 30, 1441–1451. doi: 10.1523/JNEUROSCI.3244-09.2010

Hoogman, M., Onnink, M., Cools, R., Aarts, E., Kan, C., Arias Vasquez, A., et al. (2013). The dopamine transporter haplotype and reward-related striatal responses in adult ADHD. *Eur. Neuropsychopharmacol.* 23, 469–478. doi: 10.1016/j.euroneuro.2012.05.011

Houenou, J., Boisgontier, J., Henrion, A., D'Albis, M.-A., Dumaine, A., Linke, J., et al. (2017). A multilevel functional study of a SNAP25 at-risk variant for bipolar

disorder and schizophrenia. *J. Neurosci.* 37, 10389–10397. doi: 10.1523/JNEUROSCI.1040-17.2017

Ilott, N. E., Saudino, K. J., and Asherson, P. (2010). Genetic influences on attention deficit hyperactivity disorder symptoms from age 2 to 3: a quantitative and molecular genetic investigation. *BMC Psychiatry* 10:102. doi: 10.1186/1471-244X-10-102

Irwin, L. N., Soto, E. F., Chan, E. S. M., Miller, C. E., Carrington-Forde, S., Groves, N. B., et al. (2021). Activities of daily living and working memory in pediatric attention-deficit/hyperactivity disorder (ADHD). *Child Neuropsychol.* 27, 468–490. doi: 10.1080/09297049.2020.1866521

Jahn, R., and Scheller, R. H. (2006). SNAREs — engines for membrane fusion. *Nat. Rev. Mol. Cell Biol.* 7, 631–643. doi: 10.1038/nrm2002

Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage* 17, 825–841. doi: 10.1016/S1053-8119(02)91132-8

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., et al. (1997). Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): Initial reliability and validity data. *J. Am. Acad. Child Adolesc. Psychiatry* 36, 980–988. doi: 10.1097/00004583-199707000-00021

Kasper, L. J., Alderson, R. M., and Hudec, K. L. (2012). Moderators of working memory deficits in children with attention-deficit/hyperactivity disorder (ADHD): a meta-analytic review. *Clin. Psychol. Rev.* 32, 605–617. doi: 10.1016/J.CPR.2012.07.001

Kim, J., Kim, Y.-H., and Lee, J.-H. (2013). Hippocampus–precuneus functional connectivity as an early sign of Alzheimer's disease: a preliminary study using structural and functional magnetic resonance imaging data. *Brain Res.* 1495, 18–29. doi: 10.1016/j.brainres.2012.12.011

Kowalczyk, O. S., Mehta, M. A., O'Daly, O. G., and Criaud, M. (2022). Task-based functional connectivity in attention-deficit/hyperactivity disorder: a systematic review. *Biol. Psychiatry Glob. Open Sci.* 2, 350–367. doi: 10.1016/j.bpsgos.2021.10.006

Lawrie, S. M. (2020). Translational neuroimaging of ADHD and related neurodevelopmental disorders. *World J. Biol. Psychiatry* 21, 659–661. doi: 10.1080/15622975.2020.1823694

Martinez-Arca, S., Rudge, R., Vacca, M., Raposo, G., Camonis, J., Proux-Gillardeaux, V., et al. (2003). A dual mechanism controlling the localization and function of exocytic v-SNAREs. *Proc. Natl. Acad. Sci. U.S.A.* 100, 9011–9016. doi: 10.1073/pnas.1431910100

Mill, J., Richards, S., Knight, J., Curran, S., Taylor, E., and Asherson, P. (2004). Haplotype analysis of SNAP-25 suggests a role in the aetiology of ADHD. *Mol. Psychiatry* 9, 801–810. doi: 10.1038/sj.mp.4001482

Mohammadi, M. R., Zarafshan, H., Khaleghi, A., Ahmadi, N., Hooshyari, Z., Mostafavi, S. A., et al. (2021). Prevalence of ADHD and its comorbidities in a population-based sample. *J. Attenti. Disord.* 25, 1058–1067. doi: 10.1177/1087054719886372

Nackley, A. G., Shabalina, S. A., Tchivileva, I. E., Satterfield, K., Korchynskyi, O., Makarov, S. S., et al. (2006). Human Catechol- O -Methyltransferase haplotypes modulate protein expression by altering mRNA secondary structure. *Science* 314, 1930–1933. doi: 10.1126/science.1131262

Najera, K., Fagan, B. M., and Thompson, P. M. (2019). SNAP-25 in major psychiatric disorders: a review. *Neuroscience* 420, 79–85. doi: 10.1016/j.neuroscience.2019.02.008

Németh, N., Kovács-Nagy, R., Székely, A., Sasvári-Székely, M., and Rónai, Z. (2013). Association of impulsivity and polymorphic microRNA-641 target sites in the SNAP-25 gene. *PLoS One* 8:e84207. doi: 10.1371/journal.pone.0084207

Noor, A., and Zahid, S. (2017). A review of the role of synaptosomal-associated protein 25 (SNAP-25) in neurological disorders. *Int. J. Neurosci.* 127, 805–811.

Oades, R. D. (2007). Role of the serotonin system in ADHD: treatment implications. *Expert Rev. Neurother.* 7, 1357–1374. doi: 10.1586/14737175.7.10.1357

Onnink, A. M. H., Zwiers, M. P., Hoogman, M., Mostert, J. C., Dammers, J., Kan, C. C., et al. (2015). Deviant white matter structure in adults with attention-deficit/hyperactivity disorder points to aberrant myelination and affects neuropsychological performance. *Prog. Neuro Psychopharmacol. Biol. Psychiatry* 63, 14–22. doi: 10.1016/j.pnpbp.2015.04.008

Ortega, R., López, V., Carrasco, X., Escobar, M. J., García, A. M., Parra, M. A., et al. (2020). Neurocognitive mechanisms underlying working memory encoding and retrieval in attention-deficit/hyperactivity disorder. *Sci. Rep.* 10:7771. doi: 10.1038/s41598-020-64678-x

Osen-Sand, A., Staple, J. K., Naldi, E., Schiavo, G., Rossetto, O., Petitpierre, S., et al. (1996). Common and distinct fusion proteins in axonal growth and transmitter release. *J. Comp. Neurol.* 367, 222–234. doi: 10.1002/(SICI)1096-9861(19960401)367:2<222::AID-CNE5>3.0.CO;2-7

Papadopoulos, A., Seguin, D., Correa, S., and Duerden, E. G. (2021). Peer victimization and the association with hippocampal development and working memory in children with ADHD and typically-developing children. *Sci. Rep.* 11:16411. doi: 10.1038/s41598-021-95582-7

Plessen, K. J., Bansal, R., Zhu, H., Whiteman, R., Amat, J., Quackenbush, G. A., et al. (2006). Hippocampus and Amygdala Morphology in Attention-Deficit/Hyperactivity Disorder. *Arch. Gen. Psychiatry* 63, 795–807. doi: 10.1001/archpsyc.63.7.795

Polanczyk, G., de Lima, M. S., Horta, B. L., Biederman, J., and Rohde, L. A. (2007). The worldwide prevalence of ADHD: a systematic review and metaregression analysis. *Am. J. Psychiatry* 164, 942–948. doi: 10.1176/ajp.2007.164.6.942

Posner, J., Siciliano, F., Wang, Z., Liu, J., Sonuga-Barke, E., and Greenhill, L. (2014). A multimodal MRI study of the hippocampus in medication-naive children with ADHD: what connects ADHD and depression? *Psychiatry Res.* 224, 112–118. doi: 10.1016/j.pscychresns.2014.08.006

Posner, M. I., Rothbart, M. K., Sheese, B. E., and Voelker, P. (2014). Developing attention: behavioral and brain mechanisms. *Adv. Neurosci.* 2014, 1–9. doi: 10.1155/2014/405094

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., and Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59, 2142–2154. doi: 10.1016/j.neuroimage.2011.10.018

Rapport, M. D., Alderson, R. M., Kofler, M. J., Sarver, D. E., Bolden, J., and Sims, V. (2008). Working memory deficits in boys with attention-deficit/hyperactivity disorder (ADHD): the contribution of central executive and subsystem processes. *J. Abnorm. Child Psychol.* 36, 825–837.

Ren, Y., Nguyen, V. T., Sonkusare, S., Lv, J., Pang, T., Guo, L., et al. (2018). Effective connectivity of the anterior hippocampus predicts recollection confidence during natural memory retrieval. *Nat. Commun.* 9:4875. doi: 10.1038/s41467-018-07325-4

Roya, G. (2016). *An investigation of abnormal brain connectivity associated with regions implicated in ADHD*. University, MS: University of Mississippi.

Sarkar, K., Bhaduri, N., Ghosh, P., Sinha, S., Ray, A., Chatterjee, A., et al. (2012). Role of SNAP25 explored in eastern Indian attention deficit hyperactivity disorder probands. *Neurochem. Res.* 37, 349–357. doi: 10.1007/s11064-011-0618-8

Shaw, P., Gornick, M., Lerch, J., Addington, A., Seal, J., Greenstein, D., et al. (2007). Polymorphisms of the dopamine D4 receptor, clinical outcome, and cortical structure in attention-deficit/hyperactivity disorder. *Arch. Gen. Psychiatry* 64, 921–931. doi: 10.1001/archpsyc.64.8.921

Shipton, O. A., El-Gaby, M., Apergis-Schoute, J., Deisseroth, K., Bannerman, D. M., Paulsen, O., et al. (2014). Left–right dissociation of hippocampal memory processes in mice. *Proc. Nal. Acad. Sci. U.S.A.* 111, 15238–15243. doi: 10.1073/pnas.1405648111

Siok, W. T., Spinks, J. A., Jin, Z., and Tan, L. H. (2009). Developmental dyslexia is characterized by the co-existence of visuospatial and phonological disorders in Chinese children. *Current Biol.* 19, R890—-R892.

Söllner, T., Whiteheart, S. W., Brunner, M., Erdjument-Bromage, H., Geromanos, S., Tempst, P., et al. (1993). SNAP receptors implicated in vesicle targeting and fusion. *Nature* 362, 318–324.

van Rooij, D., Hoekstra, P. J., Mennes, M., von Rhein, D., Thissen, A. J. A. M., Heslenfeld, D., et al. (2015). Distinguishing adolescents With ADHD from their unaffected siblings and healthy comparison subjects by neural activation patterns during response inhibition. *Am. J. Psychiatry* 172, 674–683. doi: 10.1176/appi.ajp.2014.13121635

Wang, C., Yang, B., Fang, D., Zeng, H., Chen, X., Peng, G., et al. (2018). The impact of SNAP25 on brain functional connectivity density and working memory in ADHD. *Biol. Psychol.* 138, 35–40.

Wang, W., Wang, F., Liu, J., Zhao, W., Zhao, Q., He, M., et al. (2014). SNAP25 ameliorates sensory deficit in rats with spinal cord transection. *Mol. Neurobiol.* 50, 290–304. doi: 10.1007/s12035-014-8642-8

Woelfer, M., Li, M., Colic, L., Liebe, T., Di, X., Biswal, B., et al. (2020). Ketamine-induced changes in plasma brain-derived neurotrophic factor (BDNF) levels are associated with the resting-state functional connectivity of the prefrontal cortex. *World J. Biol. Psychiatry* 21, 696–710. doi: 10.1080/15622975.2019.1679391

Xue, J., Guo, H., Gao, Y., Wang, X., Cui, H., Chen, Z., et al. (2019). Altered directed functional connectivity of the hippocampus in mild cognitive impairment and alzheimer's disease: a resting-state fMRI study. *Front. Aging Neurosci.* 11:326. doi: 10.3389/fnagi.2019.00326

Yan, C. G., Wang, X.-D., Zuo, X. N., and Zang, Y. F. (2016). DPABI: data processing & analysis for (Resting-State) brain imaging. *Neuroinformatics* 14, 339–351. doi: 10.1007/S12021-016-9299-4/FIGURES/6

Yang, P., Cheng, C.-P., Chang, C.-L., Liu, T.-L., Hsu, H.-Y., and Yen, C.-F. (2013). Wechsler Intelligence Scale for Children 4th edition-Chinese version index scores in Taiwanese children with attention-deficit/hyperactivity disorder. *Psychiatry Clin. Neurosci.* 67, 83–91. doi: 10.1111/pcn.12014

Yang, Y., Bi, H.-Y., Long, Z.-Y., and Tao, S. (2013). Evidence for cerebellar dysfunction in Chinese children with developmental dyslexia: an fMRI study. *Int. J. Neurosci.* 123, 300–310.

Yang, Y., Peng, G., Zeng, H., Fang, D., Zhang, L., Xu, S., et al. (2022). Effects of the SNAP25 on integration ability of brain functions in children with ADHD. *J. Attent. Disord.* 26, 88–100.

Ye, C., Hu, Z., Wu, E., Yang, X., Buford, U. J., Guo, Z., et al. (2016). Two SNAP-25 genetic variants in the binding site of multiple microRNAs and susceptibility of ADHD: a meta-analysis. *J. Psychiatric Res.* 81, 56–62.

Zhong, S., Ding, W., Sun, L., Lu, Y., Dong, H., Fan, X., et al. (2020). Decoding the development of the human hippocampus. *Nature* 577, 531–536.

frontiers | Frontiers in Human Neuroscience

TYPE Original Research
PUBLISHED 23 August 2023
DOI 10.3389/fnhum.2023.1218437

OPEN ACCESS

EDITED BY
Tetsuo Kida,
Institute for Developmental Research, Japan

REVIEWED BY
Aneta Brzezicka,
University of Social Sciences and Humanities, Poland
Hugo Lehmann,
Trent University, Canada

*CORRESPONDENCE
Koichi Yokosawa
✉ yokosawa@med.hokudai.ac.jp

RECEIVED 07 May 2023
ACCEPTED 08 August 2023
PUBLISHED 23 August 2023

CITATION
Onishi H and Yokosawa K (2023) Differential working memory function between phonological and visuospatial strategies: a magnetoencephalography study using a same visual task.
Front. Hum. Neurosci. 17:1218437.
doi: 10.3389/fnhum.2023.1218437

Differential working memory function between phonological and visuospatial strategies: a magnetoencephalography study using a same visual task

Hayate Onishi[1] and Koichi Yokosawa[2]*

[1]Graduate School of Health Sciences, Hokkaido University, Sapporo, Japan, [2]Faculty of Health Sciences, Hokkaido University, Sapporo, Japan

Previous studies have reported that, in working memory, the processing of visuospatial information and phonological information have different neural bases. However, in these studies, memory items were presented via different modalities. Therefore, the modality in which the memory items were presented and the strategy for memorizing them were not rigorously distinguished. In the present study, we explored the neural basis of two working memory strategies. Nineteen right-handed young adults memorized seven sequential directions presented visually in a task in which the memory strategy was either visuospatial or phonological (visuospatial/phonological condition). Source amplitudes of theta-band (5–7 Hz) rhythm were estimated from magnetoencephalography during the maintenance period and further analyzed using cluster-based permutation tests. Behavioral results revealed that the accuracy rates showed no significant differences between conditions, while the reaction time in the phonological condition was significantly longer than that in the visuospatial condition. Theta activity in the phonological condition was significantly greater than that in the visuospatial condition, and the cluster in spatio-temporal matrix with $p < 5\%$ difference extended to right prefrontal regions in the early maintenance period and right occipito-parietal regions in the late maintenance period. The theta activity results did not indicate strategy-specific neural bases but did reveal the dynamics of executive function required for phonological processing. The functions seemed to move from attention control and inhibition control in the prefrontal region to inhibition of irrelevant information in the occipito-parietal region.

KEYWORDS

working memory, phonological strategy, visuospatial strategy, magnetoencephalography (MEG), executive functions

Abbreviations: MEG, magnetoencephalography; ROI, region of interest; WM, working memory.

1. Introduction

Working memory (WM) is responsible for the temporary maintenance and manipulation of information to carry out certain behavioral goals (Goldman-Rakic, 1995; Fuster and Bressler, 2012; Miller et al., 2018, as reviews), which is a higher cognitive function that is essential in daily life.

In the classical psychological model of WM (Baddeley and Hitch, 1974; Baddeley, 2012), WM comprises a phonological loop that stores phonological information, a visuospatial sketchpad that stores visuospatial information, and an executive function (central executive) that manipulates the information stored by these storage mechanisms. The neural basis of WM has been studied using a variety of tasks. For instance, studies using auditory stimuli or visual words have revealed brain activity in temporo-parietal and inferior frontal regions (Deng et al., 2011; Fegen et al., 2015; Albouy et al., 2019), while studies using objects or spatial positions revealed brain activity in occipital and parietal regions (Todd and Marois, 2004; Curtis, 2006; Sobczak-Edmans et al., 2016). These studies have typically adopted either phonological or visuospatial information, and have not always excluded the possibility that different subjects used different memory strategies to manipulate the information. In addition, because these studies were based on functional magnetic resonance imaging, the temporal variation of WM function has not been clarified. As described above, it remains unclear how WM function and its temporal variation differ when the same information is manipulated using different strategies.

Non-invasive electrophysiological methods [i.e., electroencephalography and magnetoencephalography (MEG)] are appropriate for recording temporal variation in brain activities. Various bands of brain rhythms have been studied to investigate memory processing (Klimesch, 1999; Crivelli-Decker et al., 2018, for review). Couplings between bands of brain rhythms have also been investigated, on the basis of Lisman and Idiart's model (Lisman and Idiart, 1995). For example, Bahramisharif et al. (2018) investigated not only the theta/alpha-band (7–13 Hz) power during maintenance, but also the couplings between gamma-band (>30 Hz) power and the theta/alpha-band phase. The researchers concluded that the increase in theta/alpha-band power reflects the inhibition of upcoming sensory irrelevant information and the protection of the information that is already held in WM (Bahramisharif et al., 2018). Malenínská et al. (2021) reported that the theta to gamma cycle length ratio predicted memory performance using the digit span test. In contrast, within single bands of brain rhythms, an amplitude increase of the hippocampal theta-band (5–7 Hz) rhythm and an amplitude decrease of the cortical alpha-band (8–13 Hz) rhythm have been reported in learning and memory (Parish et al., 2018). Costers et al. (2020) demonstrated that theta activity reflects sensory processing in a study using the n-back task. Mizuhara et al. (2004) and Magosso et al. (2021) suggested that theta activity is associated with executive function. The alpha-band rhythm has advantageous characteristics for study, with a large amplitude and wide modulation, and there is considerable evidence indicating its involvement in short-term memory and WM processing (Foster et al., 2016; Wianda and Ross, 2019). Therefore, we compared the roles of alpha- and theta-band rhythm in our previous study (Takase et al., 2019). When young participants performed a sequential memory task similar to that used in the current study, the alpha-band rhythm contributed exclusively to the active inhibition of task-irrelevant inputs, whereas cortical theta-band rhythm was associated with memory performance. Thus, in the current study, we focused on the cortical theta-band rhythm.

In the current study, we recorded MEG while subjects performed a WM task that required two different memory strategies while presenting the same visual stimuli. We estimated the source amplitude of the theta-band rhythm during the period in which subjects maintained the information and rehearsed it in two different ways. By applying the cluster-based permutation test and effect sizes method (Meyer et al., 2021), we aimed to identify brain regions in which theta activities differed between strategies and how these differences varied over time. We expected to find a specific neural basis for each strategy.

Visually presented words are recognized by their shape, sound, and meaning, whereas objects placed in space are recognized by their shape, position, and angle. The recognition processes for visually presented words and objects involve the temporo-parietal and inferior frontal regions, and occipital and parietal regions, respectively. The previous studies mentioned above have generally reported effects that correspond to these brain areas (e.g., Sobczak-Edmans et al., 2016; Albouy et al., 2019). However, it remains unclear which brain areas are involved when the same visual stimuli are presented. Additionally, it is not known whether there is a difference in temporal variation between the different memory strategies during maintenance. When different processes are performed on the same visual stimuli, executive function is likely to be more important. Recent neuroscientific findings have suggested that maintenance and manipulation are accomplished by executive functions involving attention control, prioritizing information, inhibiting irrelevant information, and updating information (Diamond, 2013; D'Esposito and Postle, 2015; Myers et al., 2017; Chiu et al., 2018). Internal attention to representations is particularly important for maintenance (Banich et al., 2000; Magosso et al., 2021). Executive functions have been reported to involve prefrontal and parietal regions (Leung and Zhang, 2004; Mogadam et al., 2018; Schumacher et al., 2019). To detect the temporal dynamics of memory processing of different memory strategies, including executive function, we considered that it was necessary to take a spatial-temporal exploratory approach.

As a WM task, we designed a visual sequential memory task in which subjects memorized one of four (up, down left, right) directions. Subjects watched a white circle indicating the direction and a Kanji character indicating a direction at the same time and memorized one of them (**Figure 1**). Thus, subjects were forced to choose either a phonological or visuospatial strategy to maintain the information and perform rehearsal. Subjects reported the direction by pressing one of four directional buttons in our task, in contrast to traditional WM tasks in which subjects typically give two-choice (i.e., yes/no) responses. Thus, we were able to assess accuracy more rigorously using our WM task, because subjects were given four choices.

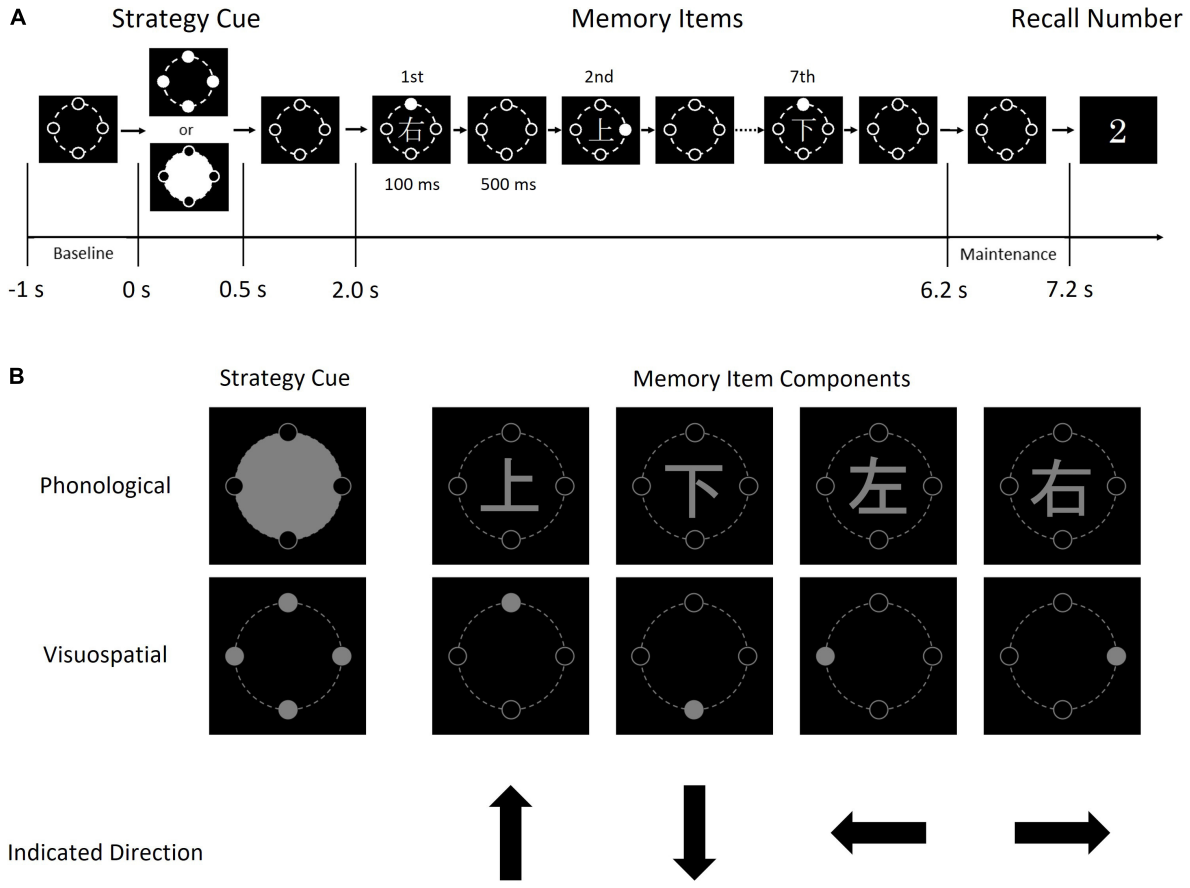


FIGURE 1
Schematic diagram of one trial of the sequential memory task **(A)** and the details of each image **(B)**. Each memory item consisted of a Kanji character and a gray circle. The condition cue indicates the condition (i.e., phonological or visuospatial); when the upper image (a large gray circle with four black circles) was presented, each subject was required to memorize the direction indicated by the Kanji character, by reading and memorizing it. In contrast, when the lower image (four gray circles) was presented, each subject was required to memorize the direction indicated by the position of the solid gray circle.

2. Materials and methods

2.1. Subjects

Nineteen right-handed young adults took part in the experiment [22.1 ± 1.1 years old (mean ± standard deviation), 12 males, seven females]. Subjects reported no history of neurological or psychiatric disorders and all reported normal or corrected-to-normal vision. The study was approved by the Ethics Committees of Faculty of Health Sciences at Hokkaido University. Written informed consent was obtained from each subject prior to the experiments.

2.2. Sequential memory task

Subjects were instructed to perform the visual sequential memory task shown in **Figure 1**, involving static images of a condition cue, memory items, and recall number. The task was created using Presentation (Neurobehavioral Systems, ver. 20.3, Berkeley, CA, USA), and memory items were presented in gray on a black background. The fonts of Kanji characters and recall numbers were Yu Gothic and Lucida Console, respectively. These images were presented using a back-projection screen. The screen was located in the magnetically shielded room for MEG recoding, while the projector was located outside.

First, a condition cue (i.e., phonological or visuospatial) was presented. Seven memory items consisting of Kanji characters (上, 下, 右, and 左 mean up, down, right, and left, respectively) and a solid gray circle (positioned up, down, right, or left), were presented in order. Finally, a recall number from 1 to 7 was presented. In the phonological strategy, subjects were instructed to read and memorize a Kanji character. In the visuospatial strategy, subjects were instructed to memorize the positions of the solid gray circles. Thus, subjects memorized the directions using different strategies. Subjects answered the direction of the memory item at the serial position corresponding to the recall number by pressing one of four directional buttons with the right index finger. That is, if the recall number was 2 (see **Figure 1**), the subject was instructed to press the up button in the phonological condition because "上" represented "up," and to press the right button in the visuospatial condition.

The condition cue was presented for 500 ms. After 1,500 ms from the disappearance of the condition cue, seven memory items

were presented for 100 ms with a 600-ms interval. After 1,500 ms from the disappearance of the last memory item, the recall number was presented for 2,500 ms ("2" in **Figure 1**). The recall number vanished promptly after the subject pressed a button.

The experiment was composed of four task blocks for each subject. One task block included 56 trials with 28 phonological and 28 visuospatial conditions. Each recall number appeared four times in one block. Condition and recall number were presented pseudo-randomly in each task block.

The answered directions and the duration between the onset of the recall number and the moment of button pressing were recorded automatically to calculate the accuracy rates and reaction times. Here, 0 s denotes the onset of the condition cue. The time periods of −1 to 0 s and 6.2–7.2 s were classified as the baseline period and maintenance period, respectively. Inter-trial intervals were randomly selected between 3 s and 4 s.

2.3. MEG recording

Magnetoencephalography was recorded using a 101-channel helmet-shaped magnetometer system (customized; Elekta-Neuromag) installed at Hokkaido University, Sapporo, Japan. The passband was 0.1–100 Hz and the sampling rate was 600 Hz. The recordings were conducted in a magnetically shielded room. In each recording, the helmet-shaped sensor-array was tilted forward 10°. Three head position indicator coils were attached at subjects' left and right pre-auricular points and nasion, and the head points were obtained using a digitizer according to standard MEG operating procedure (Takase et al., 2022). Each subject was seated in the MEG measurement chair, onto which a table was attached. On the table, a cross-key was placed at a comfortable position for pressing the button with the right hand. To prevent the subject from moving their head to look at the cross-key, the cross-key and right hand were covered with a box so that they were not visible to the subject. The distance between the screen and a subject was approximately 120 cm, and the diameter of the memory items was 7 cm. Therefore, the viewing angle was 3–4°. Each subject was instructed to keep looking straight ahead and not to perform eye movements.

2.4. Behavioral analyses

The accuracy rates and the reaction times were calculated according to each recall number. Two-way repeated analyses of variance (ANOVA), condition (phonological/visuospatial) × recall number (1–7), and *post hoc* analysis were performed on the accuracy rates and the reaction time separately. The significance threshold was set at 5%, and multiple comparisons were corrected using the Bonferroni method.

2.5. MEG preprocessing

Magnetoencephalography data were preprocessed and statistically analyzed using Brainstorm (Tadel et al., 2011), as follows. First, the MEG data from malfunctioning or excessively noisy sensors were removed. Next, mechanical noise or artifact caused by respirations or cardiac beats were eliminated using independent-components analysis. Noise characteristics of each MEG data were checked using Fourier transformation. MEG data recorded by sensors with prominent noise in the 1–40 Hz frequency range were removed. The remaining MEG data were band-pass filtered at 1–40 Hz. The filtered MEG data were extracted for each trial with −1 to 7.2 s (0 s denotes the onset of the condition cue). There were trials of 28 phonological and 28 visuospatial conditions in each of four task blocks. That is, there were 224 trials in total in one experiment. Extracted MEG data of each trial were checked by visual inspection. Data from trials with prominent artifacts were removed. To focus on amplitudes of theta-band (5–7 Hz) rhythm, the template evoked response was calculated and subtracted from the MEG data.

Current dipole moments at 15,002 vertices on the cortex were estimated by unconstrained minimum norm estimation. To estimate the shape of the individual brain, the head points of each subject were co-registered to the template brain and an overlapping-sphere forward model was computed prior to minimum norm estimation. Thus, the time series data of the three orthogonal components were obtained at 15,002 vertices on the estimated individual brain.

2.6. MEG analyses

The theta-band envelopes were computed in each orthogonal component using Hilbert transform. The L2 norm of the obtained envelope was then computed at each vertex. The results were averaged across all trials for each condition. The averaged envelopes were taken as the time series of theta activity in the following analysis. The theta activity data were transformed to standardized theta activity data by calculating the amplitude deviations, X_{std}, which were the change rates of the amplitude against the mean amplitude within the baseline (−1 to 0 s) period. That is,

$$X_{std} = \frac{x - \mu}{\mu} \times 100,$$

where x is the amplitude at each time point, and μ is the mean amplitude within the baseline. 15,002 time series data points for standardized theta activities (%) were obtained for each subject and condition. To focus on memory maintenance, spatio-temporal data in the maintenance period (6.2–7.2 s) were extracted. The data were projected on the template brain to perform group statistics. To reduce the load on the computer, the data were down-sampled to 200 Hz (Costers et al., 2020).

2.7. Modulation of standardized theta activity

To survey the time course of the standardized theta activity overall, the standardized theta activity values for each condition were averaged over 15,002 vertices and all subjects.

2.8. Cluster-based permutation test

Spatio-temporal matrices (15,002 vertices × 200 time-points) of standardized theta activity values per subject of each condition were obtained via the process described above. To extract the difference in standardized theta activities between the phonological condition and the visuospatial condition from the spatio-temporal matrices, we adopted a cluster-based permutation test (Maris and Oostenveld, 2007) to reduce false positives caused by multiple comparisons. However, it has been noted that clusters do not provide precise spatial or temporal information (see Fieldtriptoolbox, 2020), because of the null hypothesis that both conditions come from the same distribution, that is, data were exchangeable. Nevertheless, the cluster provides the characteristics of the difference, which is informative when there is a lack of *a priori* information, as in this study.

Therefore, we selected brain regions of interest as follows. First, we calculated t-values of dependent samples *t*-test over the spatio-temporal matrices (15,002 × 200 samples) between two conditions and selected all samples whose t-value was larger than the threshold (the 97.5th quantile of a T-distribution). Next, we clustered the selected samples in connected sets based on spatial/temporal adjacency; the minimum number of samples was set as 2. We then calculated cluster-level statistics (SumT) by taking the sum of the t-values within each cluster. Finally, we took the largest SumT among the clusters.

In the next step, we performed a permutation test on the largest SumT. That is, we permutated the condition of the spatio-temporal matrices of each subject randomly and performed the above process to calculate the largest SumT. Although there were 2^{19} combinations, we performed 999 random combinations and constructed a histogram of the SumT. We confirmed that the largest SumT of the real combination was included in the largest 5% of the histogram (n = 1,000, alpha-level = 0.05). To ensure higher robustness, we narrowed down the cluster by extracting the samples with more than 100 adjacent vertices in each time point. The narrowed cluster was delineated at each time-point to determine regions of interest. That is, the brain regions of interest were decided by selecting the brain regions using Mindboggle atlas (Klein et al., 2005), which include the vertices in the narrowed cluster.

As we conducted the above process using a two-sided test, SumT could be bilateral: positive clusters with positive SumT values indicate that standardized theta activities of the cluster were larger for the phonological condition than those for visuospatial condition, whereas negative clusters with negative SumT indicate the opposite.

2.9. Effect sizes

It has been reported that a combination of the cluster-based permutation test and effect size evaluation (e.g., Cohen, 1988) is an effective approach for obtaining statistically robust results from large-scale data (Meyer et al., 2021). We therefore constructed spatio-temporal matrices of the standardized theta activities only for the brain regions of interest. Additionally, we calculated Cohen's d of the samples (standardized theta activities) between conditions at each time point. We selected the time points including at least one vertex with a d-value above 0.8. A time span with a large effect (i.e., $d > 0.8$) of the condition on the theta activity was determined for each brain region of interest.

3. Results

3.1. Behavioral results

Regarding accuracy rates, two-way repeated ANOVA revealed a significant main effect of recall number [$F_{(2.834,51.021)} = 13.517$, $p < 0.001$]. There was no main effect of condition [$F_{(1,18)} = 0.147$, $p = 0.706$], and no interaction between condition and recall number [$F_{(4.164,74.961)} = 1.214$, $p = 0.312$]. Regarding reaction time, the results revealed significant main effects of both condition [$F_{(1,18)} = 23.246$, $p < 0.001$, **Figure 2**, inset] and recall number [$F_{(2.258,40.647)} = 30.854$, $p < 0.001$]. There was no interaction between condition and recall number [$F_{(3.538,63.688)} = 1.113$, $p = 0.355$]. For both accuracy rate and reaction time, the degrees of freedom were corrected using the Greenhouse–Geisser method. The *post hoc* analysis showed significant differences between recall numbers in both accuracy rate and reaction time (**Figures 2**, **3**). That is, sequential position effects, a characteristic of sequential memory tasks, were clearly observed. We investigated the relationship between accuracy rate and brain rhythms focusing on the serial position effect adopting similar sequential memory tasks (Yokosawa et al., 2020). Nevertheless, this work aims to validate the differences between memory strategies (i.e., conditions). The main effect of the condition was observed only for reaction time, and there was no interaction with recall number. Therefore, in the subsequent discussion, we focused on the significantly longer reaction time in the phonological condition compared with that in the visuospatial condition.

3.2. Modulation of standardized theta activity

The time courses of the standardized theta activity averaged over 15,002 vertices and grand-averaged among subjects are shown in **Figure 4**. In both conditions, the standardized theta activities peaked after the strategy cue presentation and decreased overall during encoding, exhibiting increases and decreases according to each memory item presentation. During the maintenance period, a distinctive peak was only observed in the phonological condition (indicated by an arrow in **Figure 4**).

3.3. Cluster-based permutation test

Focusing on the distinctive difference between conditions observed in the maintenance period, we conducted a cluster-based permutation test on the standardized theta activities in the maintenance period. The results revealed that the standardized theta activity was significantly greater in the phonological condition

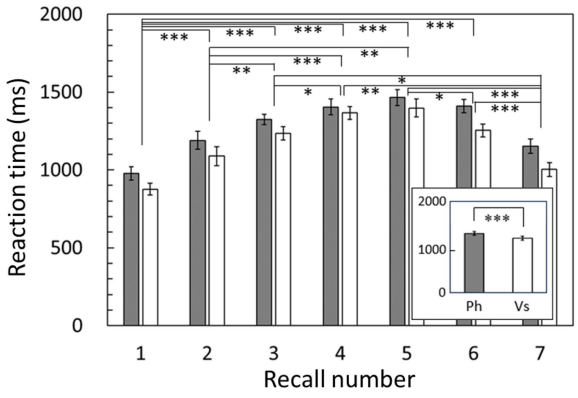


FIGURE 2
Reaction times averaged among subjects for each recall number and compared between conditions (inset). Reaction times in the phonological condition were significantly longer than those in the visuospatial condition ($p < 0.001$). Results of *post hoc* analysis for the recall number are also shown. Error bar: standard error. Ph or gray bars: phonological condition, Vs or white bars: visuospatial condition. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

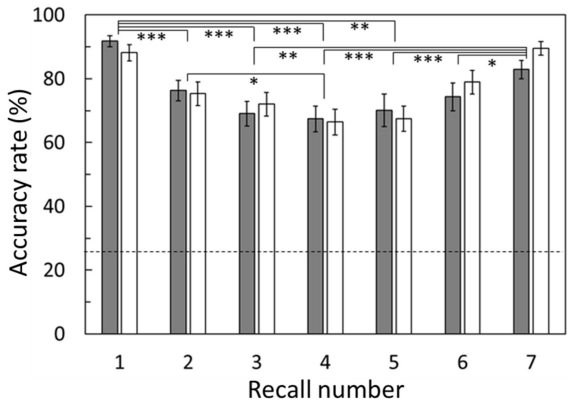


FIGURE 3
Accuracy rates averaged among subjects for each recall number. Accuracy rates exhibited no significant difference between conditions. Clear U-shaped serial position effects were shown in both conditions. Results of *post hoc* analysis for the recall number are also shown. Error bar: standard error. Gray bars: phonological condition, white bars: visuospatial condition. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

compared with the visuospatial condition ($p = 0.024$). One two-dimensional positive cluster was observed in the spatial (15,002 vertices of the whole brain cortices) and temporal (200 time points in 1 s of the maintenance period) matrix with difference of the standardized theta activities between conditions (**Figure 5**). Thus, in the maintenance period, standardized theta activity values in the phonological condition were significantly larger than those in the visuospatial condition. The cluster extended to right prefrontal regions in the early maintenance period and right occipito-parietal regions in the late maintenance period. Given the particularly high reliability of MEG for signal sources in cortices that are close to the scalp, we extracted regions of interest (ROIs) on the lateral surface of the cortices (**Figure 6**).

3.4. Effect sizes

Regions of interest were defined as brain regions that included the vertices, including in the cluster. Right lateral orbitofrontal, pars orbitalis, pars triangularis, superior parietal, inferior parietal, and lateral occipital regions were selected on the basis of Mindboggle atlas (**Figure 6**). In each ROI, effect sizes were calculated at each time-point and vertex. We considered that the effect size was sufficiently large (i.e., valid) when the value of d was greater than 0.8, in accord with a previous study (Cohen, 1988). **Figure 7** summarizes the time periods in which the effect sizes were greater than 0.8. **Figure 8** shows the time traces of the standardized theta activities. The time periods in which each ROI included valid vertices are denoted by gray shadow.

4. Discussion

Behavioral data revealed that the reaction times were significantly longer for the phonological condition than those for the visuospatial condition, although the accuracy rates did not significantly differ between conditions (**Figure 2**). MEG data revealed that theta activity was significantly larger for the phonological condition than that for the visuospatial condition (**Figure 5**). Cluster-based permutation test and effect size data revealed that the difference was large in the prefrontal region in the early maintenance period, whereas it was large in the parietal and occipital regions in the late maintenance period (**Figures 6**, **7**). Theta activity in the prefrontal and parietal regions is considered to reflect executive function in WM, especially that in the brain areas involved in attention control and task setting (Kawasaki et al., 2010; Sobczak-Edmans et al., 2016; Myers et al., 2017). The longer reaction time and greater theta activity associated with executive function suggests that a higher level of cognitive effort is involved in the phonological condition. Additionally, theta activity was larger in the phonological condition in the occipital region, which is considered to function as a visual association area. Positron emission tomography and functional magnetic resonance imaging studies have reported that activity in the occipital region is associated with maintenance of spatial WM (see Wager and Smith, 2003, for a review). In this study, we interpreted the theta activity in sensory areas as supporting the inhibition of irrelevant information (Bahramisharif et al., 2018; Kang et al., 2020). No brain regions were observed in which theta activity was larger in the visuospatial condition.

Below, we discuss the temporal variation in theta activity in each brain region (**Figure 8**). Here, the upward convex temporal variations indicate the transient of the theta-band rhythm. We assumed that these variations represent brain activity. The use of MEG allowed us to discuss such transient temporal variation.

4.1. Prefrontal theta

In the right inferior frontal region and orbitofrontal region, theta activity was larger in the phonological condition compared

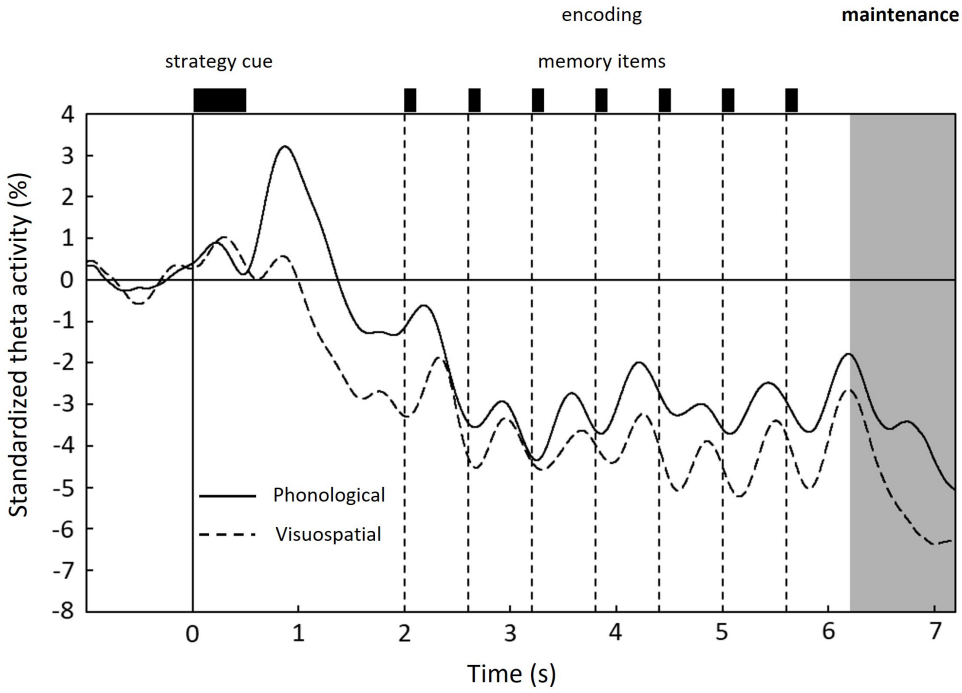


FIGURE 4
Time courses of standardized theta activity in the whole brain averaged over 15,002 vertices and grand-averaged among subjects for each condition: phonological condition (solid line) and visuospatial condition (dashed line). Unique peak during maintenance period is shown only in the phonological condition. 0 s denotes the onset of the strategy cue (see **Figure 1**).

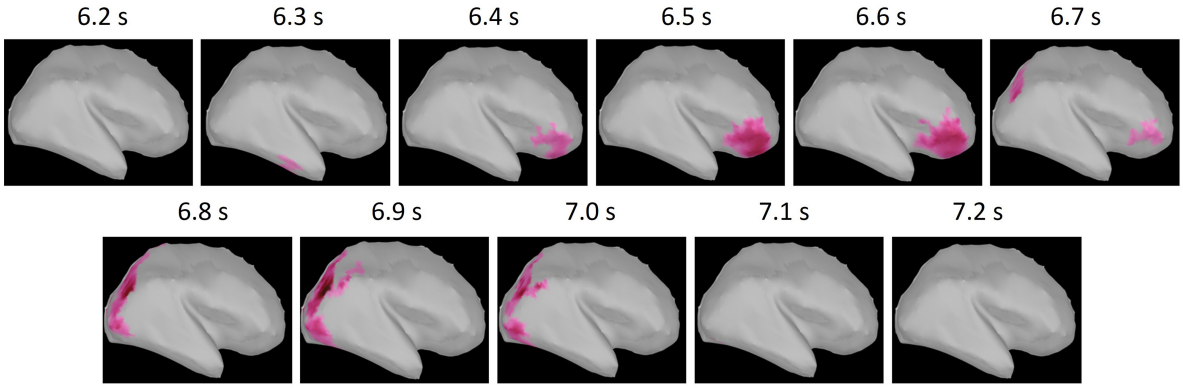


FIGURE 5
The positive (phonological condition >visuospatial condition) cluster obtained from cluster-based permutation test. The *t*-value of each vertex within the cluster is shown by the color maps. The cluster extended to right prefrontal regions in the early maintenance period (6.2–6.7 s) and the right occipito-parietal regions in the late maintenance period (6.8–7.2 s).

with that in the visuospatial condition in the early maintenance period (**Figure 8**). However, peaks (i.e., upward convex variations) in these areas were also observed in both visuospatial and phonological conditions, but were not found in the parietal or occipital regions. Considering that no theta activity specific to each condition was observed in any other brain regions, this may reflect the neural basis of strategy-independent executive function (Allen et al., 2012; Brown et al., 2012). This common peak may reflect rehearsal, particularly attention control, which is known to be important among the executive functions involved in rehearsal (Korsten et al., 2006). Below, we discuss the differentiated theta activity between conditions, referring to the functions of the subdomains of the prefrontal region, the inferior frontal gyrus, and the orbitofrontal cortex.

The ventrolateral prefrontal cortex, including the inferior frontal regions, is responsible for controlling and sustaining internal attention and applying top-down bias (task set) to information (Katsuki and Constantinidis, 2014; Weintraub-Brevda and Chua, 2019). In the phonological condition, reading, i.e., the sound of the word corresponding to the character, must have been preferentially stored in WM storage by applying bias to reading among the multiple properties of the Kanji

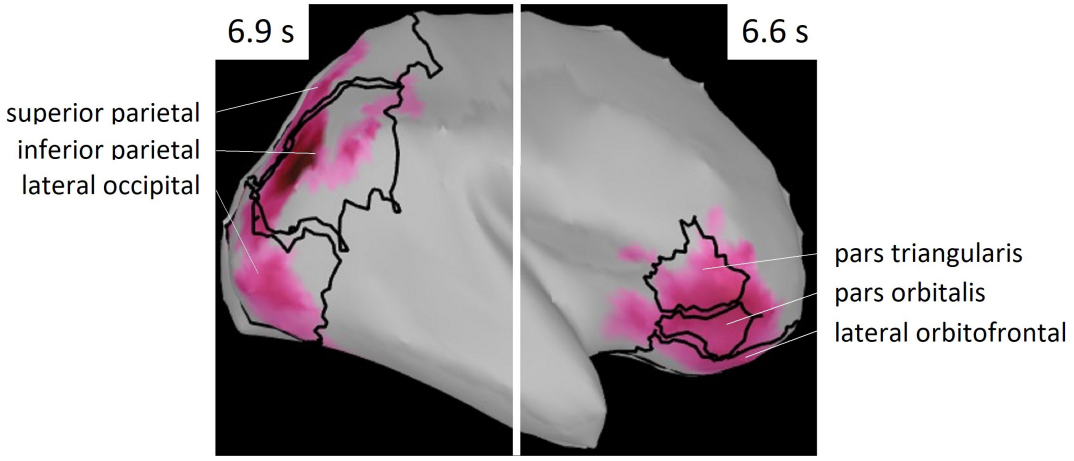


FIGURE 6
ROIs selected from Mindboggle atlas. These areas include vertices of the cluster shown in **Figure 5**.

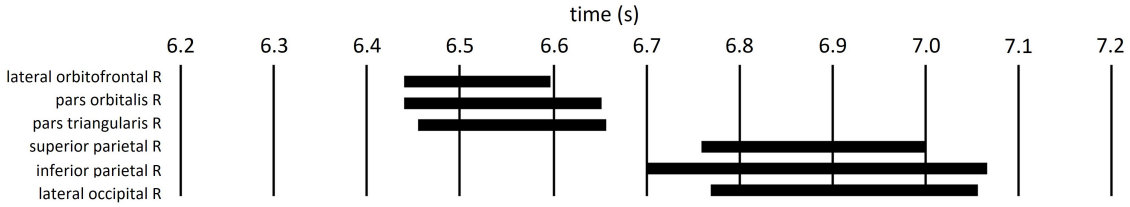


FIGURE 7
Time periods in which the ROI included the vertices with effect sizes greater than 0.8 (black bars).

character, such as shape, reading, and meaning. Thus, this function is already performing attentional selection (external attention) during encoding. During the maintenance period, subjects were asked to covertly rehearse the reading of the Kanji characters in the phonological condition. However, each Kanji character has not only reading, but also shape and meaning information. Thus, it is possible that other types of information were automatically recalled upon rehearsal (Pattamadilok et al., 2017). Some previous findings suggest that the meaning of a word is recalled more automatically than the color or shape of the word (Cohen et al., 1990). Our results indicate that the task set and sustaining internal attention were more strongly required in the phonological condition compared with those in the visuospatial condition (Zhang et al., 2003). In a study of orbitofrontal-injured patients, the patients were not impaired on tests of simple maintenance, but only on tests that required maintenance, monitoring, or manipulation (Barbey et al., 2011). It has also been suggested that the orbitofrontal region is involved in decision-making and manipulation of representations for goal-directed behavior (Kringelbach, 2005 as review). Our results suggest that the monitoring of mental representations is important for efficient rehearsal.

These results suggest that inferior frontal and orbitofrontal regions are involved in attention control for rehearsal that is common to both strategies. In addition, the results suggest that a high level of effort is required for execution to memorize Kanji characters that indicate directions, compared with that for memorizing directions visuospatially.

4.2. Parietal theta

In the right superior and inferior parietal regions, theta activity was greater in the phonological condition compared with that in the visuospatial condition in the late maintenance period (**Figure 8**). The superior parietal region is associated with the manipulation and rearrangement of information. For example, a study of patients with damage to the superior parietal lobule reported that this region was not involved in mere maintenance or recall of information, but was only involved in manipulation and rearrangement of information (Koenigs et al., 2009). Our results suggest that more complex manipulation is required in the phonological condition compared with that required in the visuospatial condition.

In addition, parietal regions, especially the inferior parietal region, are considered to be involved in an attentional functional network together with the prefrontal cortex (Leung and Zhang, 2004; Corbetta and Shulman, 2011; Kuo et al., 2012). Banich et al. (2000) reported activity in prefrontal and parietal regions during the Stroop task. They concluded that the inferior parietal region modulates attention to information that is automatically recalled in conjunction with the properties prioritized by the task set received from the prefrontal cortex and corrects the gap, and reported that

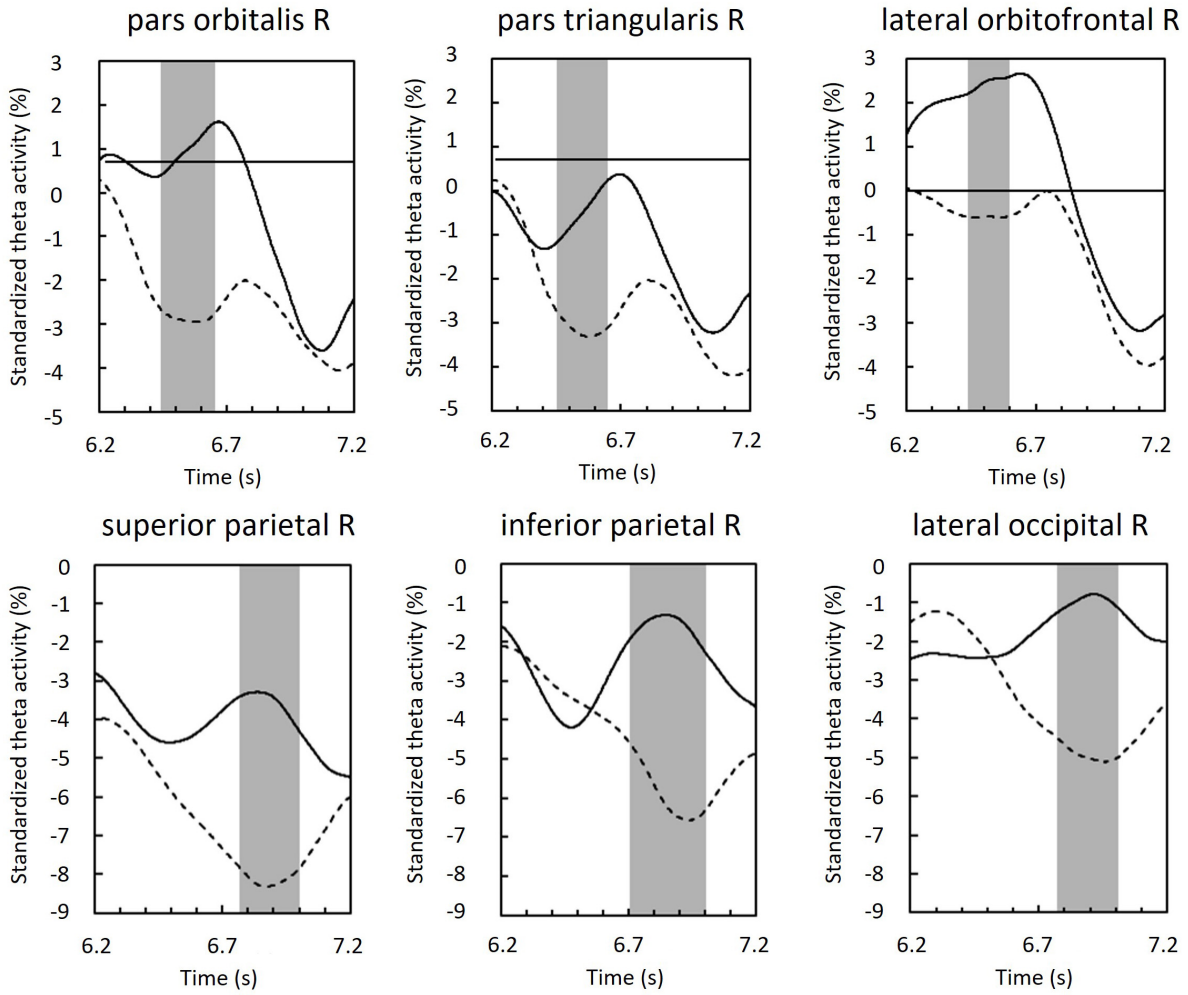


FIGURE 8
Standardized theta activity in the maintenance period of ROIs; those of the phonological condition (solid line) and visuospatial condition (dashed line). Gray shadows show time periods in which the ROI included vertices with effect sizes greater than 0.8 (see also **Figure 7**).

the parietal regions are sensitive to irrelevant information. In a study of cross-modal conflict, the inferior parietal regions, together with the frontal regions, were reported to play an important role in the inhibition of a distractor (Sawamura et al., 2022). The increase in theta activity in the parietal regions can be interpreted as playing a role in controlling the focus of attention, to avoid attending to spatial information, or to inhibit irrelevant information.

Although parietal regions have been reported to perform executive functions together with frontal regions, the present results suggest that parietal regions are active only under conditions in which there is information that interferes with the complex manipulation and rehearsal of mental representations. The present results may provide evidence that rehearsal is regulated by such an attentional functional network.

4.3. Occipital theta

In the right lateral occipital region, theta activity was greater in the phonological condition compared with that in the visuospatial condition in the late maintenance period (**Figure 8**). In a study by Kawasaki et al. (2010) comparing verbal and spatial WM tasks, phase synchronization was found between frontal theta reflecting executive function, and alpha-band rhythm reflecting modality-specific sensory processing (in verbal, left temporal; in spatial, parietal). The authors concluded that alpha activity in sensory areas reflects storage buffers (Kawasaki et al., 2010). In recent studies, during encoding, increases in theta activity in occipito-parietal regions have been commonly observed, and have been shown to reflect sensory processing of visual information and attention to visual stimuli (Dugué et al., 2015; Pisella, 2017; Takase et al., 2022).

It is unsurprising that the auditory cortex was not activated in the phonological condition in our work, because visual stimuli were presented in both the visuospatial and phonological conditions. However, unexpectedly, greater theta activity was observed in a vision-related brain region in the phonological condition compared with that observed in the visuospatial condition. Because the task in the present study required subjects to respond to direction, it may also reflect a gating-like function that permits storage of automatically recalled spatial representations. However, the longer reaction time in the phonological condition suggests

that the spatial representations were recalled on the basis of stored phonological representations after the recall number was presented. Hence, this interpretation may be questionable. Rather, we speculate that theta activity in the phonological condition may reflect inhibition of irrelevant spatial information, which should be automatically recalled during maintenance, along with parietal regions.

5. Limitations

Previous studies have reported modality-specific brain activities in sensory areas using multi-modal WM tasks. We also expected to find a specific neural basis for each strategy, by controlling the strategies for the same visual WM task to memory directions. However, we were not able to differentiate strategy-specific brain activities. This may suggest that such a difference occurs in other memory stages (i.e., encoding or recall) or that a difference is observed in rhythms with other frequency bands. The difference may be observed not in a rhythm with a single frequency band, but in cross-frequency (e.g., gamma-theta) coupling as reported by Bahramisharif et al. (2018).

Another limitation involved in the current study is that the phonological condition was more complex to process. This is likely to have occurred because the identified brain regions were exclusively involved in the phonological condition. Improvements will be needed in future study designs to align the complexities between conditions, including having participants respond in a phonological context. Further studies will be necessary to further elucidate this issue in future.

6. Conclusion

We controlled the working memory strategies (visuospatial and phonological) used by subjects to memorize directions presented visually and sequentially. Source amplitudes of the theta-band (5–7 Hz) rhythm estimated using MEG during the maintenance period were analyzed using cluster-based permutation tests. Brain activity and temporal dynamics were compared between strategies. Theta activity reflecting executive functions (particularly attention control and inhibition control) was characteristically observed in the phonological condition, and these activities were mainly supported by temporal dynamics from the frontal to the parietal regions. Furthermore, the findings suggested that not only attention to relevant information but also inhibition of irrelevant information is important for enhancing information maintenance. This inhibition appeared to occur in parietal regions and sensory areas that represented the modality of irrelevant information.

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 humans were approved by the Ethics Committees of Faculty of Health Sciences at Hokkaido University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

HO was the primary contributor for all aspects of this study, including experimental design and execution, data analysis, and writing the manuscript. KY contributed to experimental design, data analysis, editing and revising the manuscript, and approving the submitted version. Both authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Number JP20H04496 and JST SPRING Grant Number JPMJSP2119. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments

We thank Emeritus Professor Shinya Kuriki, Dr. Atsushi Shimojo, and Mr. Hayato Watanabe of Faculty of Health Sciences, Hokkaido University, Japan and Dr. Jared Boasen of HEC Montréal, Canada for their useful discussion and suggestions. We also thank colleagues, the graduate and undergraduate students at Hokkaido University and Hokkaido Bunkyo University, and Benjamin Knight, M.Sc., from Edanz (https://jp.edanz.com/ac) for editing a draft of this manuscript.

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

References

Albouy, P., Peretz, I., Bermudez, P., Zatorre, R. J., Tillmann, B., and Caclin, A. (2019). Specialized neural dynamics for verbal and tonal memory: fMRI evidence in congenital amusia. *Hum. Brain Mapp.* 40, 855–867. doi: 10.1002/hbm.24416

Allen, C. M., Martin, R. C., and Martin, N. (2012). Relations between short-term memory deficits, semantic processing, and executive function. *Aphasiology* 26, 428–461. doi: 10.1080/02687038.2011.617436

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annu. Rev. Psychol.* 63, 1–29. doi: 10.1146/annurev-psych-120710-100422

Baddeley, A., and Hitch, G. (1974). Working memory. *Psycol. Learn. Motiv.* 8, 47–89. doi: 10.1016/S0079-7421(08)60452-1

Bahramisharif, A., Jensen, O., Jacobs, J., and Lisman, J. (2018). Serial representation of items during working memory maintenance at letter-selective cortical sites. *PLoS Biol.* 16:e2003805. doi: 10.1371/journal.pbio.2003805

Banich, M. T., Milham, M. P., Atchley, R., Cohen, N. J., Webb, A., Wszalek, T., et al. (2000). fMRI studies of stroop tasks reveal unique roles of anterior and posterior brain systems in attentional selection. *J. Cogn. Neurosci.* 12, 1988–2000. doi: 10.1162/08989290051137521

Barbey, A. K., Koenigs, M., and Grafman, J. (2011). Orbitofrontal contributions to human working memory. *Cerebal* 21, 789–795. doi: 10.1093/cercor/bhq153

Brown, L. A., Brockmole, J. R., Gow, A. J., and Deary, I. J. (2012). Processing speed and visuospatial executive function predict visual working memory ability in older adults. *Exp. Aging Res.* 38, 1–19. doi: 10.1080/0361073X.2012.636722

Chiu, H. L., Chan, P. T., Kao, C. C., Chu, H., Chang, P. C., Hsiao, S. T. S., et al. (2018). Effectiveness of executive function training on mental set shifting, working memory and inhibition in healthy older adults: A double-blind randomized controlled trials. *J. Adv. Nurs.* 74, 1099–1113. doi: 10.1111/jan.13519

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, 2nd Edn. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J. D., Dunbar, K., and McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the stroop effect. *Psychol. Rev.* 97, 332–361. doi: 10.1037/0033-295X.97.3.332

Corbetta, M., and Shulman, G. L. (2011). Spatial neglect and attention networks. *Annu. Rev. Neurosci.* 34, 569–599. doi: 10.1146/annurev-neuro-061010-113731

Costers, L., Van Schependom, J., Laton, J., Baijot, J., Sjøgård, M., Wens, V., et al. (2020). Spatiotemporal and spectral dynamics of multi-item working memory as revealed by the n-back task using MEG. *Hum. Brain Mapp.* 41, 2431–2446. doi: 10.1002/hbm.24955

Crivelli-Decker, J., Hsieh, L. T., Clarke, A., and Ranganath, C. (2018). Theta oscillations promote temporal sequence learning. *Neurobiol. Learn. Mem.* 153, 92–103. doi: 10.1016/j.nlm.2018.05.001

Curtis, C. E. (2006). Prefrontal and parietal contributions to spatial working memory. *Neuroscience* 139, 173–180. doi: 10.1016/j.neuroscience.2005.04.070

Deng, Y., Chou, T. L., Ding, G., Peng, D. L., and Booth, J. R. (2011). The involvement of occipital and inferior frontal cortex in the phonological learning of Chinese characters. *J. Cogn. Neurosci.* 23, 1998–2012. doi: 10.1162/jocn.2010.21571

D'Esposito, M., and Postle, B. R. (2015). The cognitive neuroscience of working memory. *Annu. Rev. Psychol.* 66, 115–142. doi: 10.1146/annurev-psych-010814-015031

Diamond, A. (2013). Executive functions. *Annu. Rev. Psychol.* 64, 135–168. doi: 10.1146/annurev-psych-113011-143750

Dugué, L., Marque, P., and VanRullen, R. (2015). Theta oscillations modulate attentional search performance periodically. *J. Cogn. Neurosci.* 27, 945–958. doi: 10.1162/jocn_a_00755

Fegen, D., Buchsbaum, B. R., and D'Esposito, M. (2015). The effect of rehearsal rate and memory load on verbal working memory. *Neuroimage* 105, 120–131. doi: 10.1016/j.neuroimage.2014.10.034

Fieldtriptoolbox (2020). *How NOT to interpret results from a cluster-based permutation test*. Available online at: https://www.fieldtriptoolbox.org/faq/how_not_to_interpret_results_from_a_cluster-based_permutation_test/ (accessed Mar 31, 2023).

Foster, J. J., Suttener, D. W., Serences, J. T., Vogle, E. K., and Awh, E. (2016). The topography of alpha-band activity tracks the content of spatial working memory. *J. Neurophysol.* 115, 168–177. doi: 10.1152/jn.00860.2015

Fuster, J. M., and Bressler, S. L. (2012). Cognit activation: A mechanism enabling temporal integration in working memory. *Trends Cogn. Sci.* 16, 207–218. doi: 10.1016/j.tics.2012.03.005

Goldman-Rakic, P. S. (1995). Cellular basis of working memory review. *Neuron* 14, 477–485. doi: 10.1016/0896-6273(95)90304-6

Kang, C., Li, Y., Novak, D., Zhang, Y., Zhou, Q., and Hu, Y. (2020). Brain networks of maintenance, inhibition and disinhibition during working memory. *IEEE Trans. Neural Syst. Rehabil. Eng.* 28, 1518–1527. doi: 10.1109/TNSRE.2020.2997827

Katsuki, F., and Constantinidis, C. (2014). Bottom-up and top-down attention: Different processes and overlapping neural systems. *Neuroscientist* 20, 509–521. doi: 10.1177/1073858413514136

Kawasaki, M., Kitajo, K., and Yamaguchi, Y. (2010). Dynamic links between theta executive functions and alpha storage buffers in auditory and visual working memory. *Eur. J. Neurosci.* 31, 1683–1689. doi: 10.1111/j.1460-9568.2010.07217.x

Klein, A., Mensh, B., Ghosh, S., Tourville, J., and Hirsch, J. (2005). Mindboggle: Automated brain labeling with multiple atlases. *BMC Med. Imaging* 5:7. doi: 10.1186/1471-2342-5-7

Klimesch, W. (1999). Full-length review EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res. Rev.* 29, 169–195.

Koenigs, M., Barbey, A. K., Postle, B. R., and Grafman, J. (2009). Superior parietal cortex is critical for the manipulation of information in working memory. *J. Neurosci.* 29, 14980–14986. doi: 10.1523/JNEUROSCI.3706-09.2009

Korsten, N. J. H., Fragopanagos, N., Hartley, M., Taylor, N., and Taylor, J. G. (2006). Attention as a controller. *Neural Netw.* 19, 1408–1421. doi: 10.1016/j.neunet.2006.08.008

Kringelbach, M. L. (2005). The human orbitofrontal cortex: linking reward to hedonic experience. *Neuroscience* 6, 691–702. doi: 10.1038/nrn1748

Kuo, B. C., Stokes, M. G., and Nobre, A. C. (2012). Attention modulates maintenance of representations in visual short-term memory. *J. Cogn. Neurosci.* 24, 51–60. doi: 10.1162/jocn_a_00087

Leung, H. C., and Zhang, J. X. (2004). Interference resolution in spatial working memory. *Neuroimage* 23, 1013–1019. doi: 10.1016/j.neuroimage.2004.07.053

Lisman, J., and Idiart, M. (1995). Storage of 7 +/- 2 short-term memories in oscillatory subcycles. *Science* 263, 1512–1515. doi: 10.1126/science.7878473

Magosso, E., Ricci, G., and Ursino, M. (2021). Alpha and theta mechanisms operating in internal-external attention competition. *J. Integr. Neurosci.* 20, 1–19. doi: 10.31083/J.JIN.2021.01.422

Malenínská, K., Rudolfová, V., Šulcová, K., Koudelka, V., Brunovský, M., Horáček, J., et al. (2021). Is short-term memory capacity (7 +/- 2) really predicted by theta to gamma cycle length ratio? *Behav. Brain Res.* 414:113465. doi: 10.1016/j.bbr.2021.113465

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

Meyer, M., Lamers, D., Kayhan, E., Hunnius, S., and Oostenveld, R. (2021). Enhancing reproducibility in developmental EEG research: BIDS, cluster-based permutation tests, and effect sizes. *Dev. Cogn. Neurosci.* 52:101036. doi: 10.1016/j.dcn.2021.101036

Miller, E. K., Lundqvist, M., and Bastos, A. M. (2018). Working memory 2.0. *Neuron* 100, 463–475. doi: 10.1016/j.neuron.2018.09.023

Mizuhara, H., Wang, L. Q., Kobayashi, K., and Yamaguchi, Y. (2004). A long-range cortical network emerging with theta oscillation in a mental task. *Neuroreport* 15, 1233–1238. doi: 10.1097/01.wnr.0000126755.09715.b3

Mogadam, A., Keller, A. E., Taylor, M. J., Lerch, J. P., Anagnostou, E., and Pang, E. (2018). Mental flexibility: An MEG investigation in typically developing children. *Brain Cogn.* 120, 58–66. doi: 10.1016/j.bandc.2017.10.001

Myers, N. E., Stokes, M. G., and Nobre, A. C. (2017). Prioritizing information during working memory: Beyond sustained internal attention. *Trends Cogn. Sci.* 21, 449–461. doi: 10.1016/j.tics.2017.03.010

Parish, G., Hanslmayr, S., and Bowman, H. (2018). The sync/desync model: How a synchronized hippocampus and a desynchronized neocortex code memories. *J. Neurosci.* 38, 3428–3440. doi: 10.1523/JNEUROSCI.2561-17.2018

Pattamadilok, C., Chanoine, V., Pallier, C., Anton, J. L., Nazarian, B., Belin, P., et al. (2017). Automaticity of phonological and semantic processing during visual word recognition. *Neuroimage* 149, 244–255. doi: 10.1016/j.neuroimage.2017.02.003

Pisella, L. (2017). Visual perception is dependent on visuospatial working memory and thus on the posterior parietal cortex. *Ann. Phys. Rehabil.* 60, 141–147. doi: 10.1016/j.rehab.2016.01.002

Sawamura, D., Tanabe, Y., Sakuraba, S., Cui, J., Miura, H., Saito, R., et al. (2022). The impact of visual cross-modal conflict with semantic and nonsemantic distractors on working memory task: A functional near-infrared spectroscopy study. *Medicine* 101:e30330. doi: 10.1097/MD.0000000000030330

Schumacher, R., Halai, A. D., and Lambon Ralph, M. A. (2019). Assessing and mapping language, attention and executive multidimensional deficits in stroke aphasia. *Brain* 142, 3202–3216. doi: 10.1093/brain/awz258

Sobczak-Edmans, M., Ng, T. H. B., Chan, Y. C., Chew, E., Chuang, K. H., and Chen, S. H. A. (2016). Temporal dynamics of visual working memory. *Neuroimage* 124, 1021–1030. doi: 10.1016/j.neuroimage.2015.09.038

Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., and Leahy, R. M. (2011). Brainstorm: A user-friendly application for MEG/EEG analysis. *Comput. Intell. Neurosci.* 2011:879716. doi: 10.1155/2011/879716

Takase, R., Boasen, J., Kuriki, S., Toyomura, A., and Yokosawa, K. (2022). Processing time affects sequential memory performance beginning at the level of visual encoding. *PLoS One* 17:e0265719. doi: 10.1371/journal.pone.0265719

Takase, R., Boasen, J., and Yokosawa, K. (2019). Different roles for theta- and alpha-band brain rhythms during sequential memory. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2019, 1713–1716.

Todd, J. J., and Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature* 428, 751–754. doi: 10.1038/nature02466

Wager, T. D., and Smith, E. E. (2003). Neuroimaging studies of working memory: A meta-analysis. *Cogn. Affect. Behav. Neurosci.* 3, 255–274. doi: 10.3758/CABN.3.4.255

Weintraub-Brevda, R. R., and Chua, E. F. (2019). Transcranial direct current stimulation over the right and left VLPFC leads to differential effects on working and episodic memory. *Brain Cogn.* 132, 98–107. doi: 10.1016/j.bandc.2019.03.005

Wianda, E., and Ross, B. (2019). The roles of alpha oscillation in working memory retention. *Brain Behav.* 9:e01263. doi: 10.1002/brb3.1263

Yokosawa, K., Kimura, K., Takase, R., Murakami, Y., and Boasen, J. (2020). Functional decline of the precuneus associated with mild cognitive impairment: Magnetoencephalographic observations. *PLoS One* 15:e0239577. doi: 10.1371/journal.pone.0239577

Zhang, J. X., Leung, H. C., and Johnson, M. K. (2003). Frontal activations associated with accessing and evaluating information in working memory: An fMRI study. *Neuroimage* 20, 1531–1539. doi: 10.1016/j.neuroimage.2003.07.016

frontiers | Frontiers in Human Neuroscience

TYPE Original Research
PUBLISHED 11 October 2023
DOI 10.3389/fnhum.2023.1208572

OPEN ACCESS

EDITED BY
Tetsuo Kida,
Institute for Developmental Research, Japan

REVIEWED BY
Atsuko Gunji,
Yokohama National University, Japan
Yukiyasu Yaguchi,
Seitoku University, Japan

*CORRESPONDENCE
Kaori Sasaki
✉ kaori.s@iuhw.ac.jp

†PRESENT ADDRESS
Seiichi Kadowaki,
Department of Otolaryngology, Faculty of Medicine, University of Tsukuba, Tsukuba, Japan,
Department of Otolaryngology, Tokyo-Kita Medical Center, Tokyo, Japan

RECEIVED 19 April 2023
ACCEPTED 25 September 2023
PUBLISHED 11 October 2023

CITATION
Sasaki K, Kadowaki S, Iwasaki J, Pijanowska M and Okamoto H (2023) Cognitive neural responses in the semantic comprehension of sound symbolic words and pseudowords. *Front. Hum. Neurosci.* 17:1208572. doi: 10.3389/fnhum.2023.1208572

Cognitive neural responses in the semantic comprehension of sound symbolic words and pseudowords

Kaori Sasaki[1]*, Seiichi Kadowaki[2†], Junya Iwasaki[1], Marta Pijanowska[3,4] and Hidehiko Okamoto[2]

[1]Department of Speech and Hearing Sciences, International University of Health and Welfare, Narita, Japan, [2]Graduate School of Medicine, International University of Health and Welfare, Narita, Japan, [3]Office of Medical Education, International University of Health and Welfare, School of Medicine, Narita, Japan, [4]Graduate School of Humanities and Sociology, University of Tokyo, Tokyo, Japan

Introduction: Sound symbolism is the phenomenon of sounds having non-arbitrary meaning, and it has been demonstrated that pseudowords with sound symbolic elements have similar meaning to lexical words. It is unclear how the impression given by the sound symbolic elements is semantically processed, in contrast to lexical words with definite meanings. In event-related potential (ERP) studies, phonological mapping negativity (PMN) and N400 are often used as measures of phonological and semantic processing, respectively. Therefore, in this study, we analyze PMN and N400 to clarify the differences between existing sound symbolic words (onomatopoeia or ideophones) and pseudowords in terms of semantic and phonological processing.

Methods: An existing sound symbolic word and pseudowords were presented as an auditory stimulus in combination with a picture of an event, and PMN and N400 were measured while the subjects determined whether the sound stimuli and pictures match or mismatch.

Results: In both the existing word and pseudoword tasks, the amplitude of PMN and N400 increased when the picture of an event and the speech sound did not match. Additionally, compared to the existing words, the pseudowords elicited a greater amplitude for PMN and N400. In addition, PMN latency was delayed in the mismatch condition relative to the match condition for both existing sound symbolic words and pseudowords.

Discussion: We concluded that established sound symbolic words and sound symbolic pseudowords undergo similar semantic processing. This finding suggests that sound symbolism pseudowords are not judged on a simple impression level (e.g., spiky/round) or activated by other words with similar spellings (phonological structures) in the lexicon, but are judged on a similar contextual basis as actual words.

KEYWORDS

semantic comprehension, sound symbolism, onomatopoeia, event-related potential, lexical semantic processing

Abbreviations: EEG, electroencephalography; ERPs, event-related potentials; PMN, phonological mapping negativity.

1. Introduction

The arbitrary relationship between form and meaning is one of the essential characteristics of language (Hockett, 1960). Language is said to be arbitrary as there is usually no special reason why a specific form (sound or shape) is used to express a certain meaning. For example, different words, such as "apple" in English and "リンゴ (ringo)" in Japanese, are used to convey the same idea. An exception to this rule, however, is onomatopoeia. Onomatopoeia is the imitation of an actual sound that has become a standard lexical item (e.g., "bow wow" for a dog's bark). As such, onomatopoeia is considered a lexeme with a non-arbitrary connection between form and meaning (Pinker, 1999). This kind of non-arbitrary association of sound and meaning is known as sound symbolism. Sound symbolism is defined as "the direct link between sound and meaning" (Hinton et al., 1994).

Sound symbolism has been demonstrated in experiments in which participants consistently associate certain types of phonemes with specific shapes. The phoneme is the smallest unit of sound used in the word. For instance, plosives such as /t/ and /k/ are often perceived as referring to a straight line or sharp-edged figure, and nasals like /m/,/n/ as referring to a curved line or rounded figure (Köhler, 1947; Ramachandran and Hubbard, 2001). It is believed that the sound symbolism of some phonemes and phoneme clusters is connected to the structural characteristics of the brain or the mechanism by which information is processed in the brain (Sidhu and Pexman, 2018).

While many such sound symbolic elements are consistently associated with certain meanings, they are not usually considered a part of the mental lexicon. The mental lexicon is a set of words that humans retain in their brains, the idea behind a mental lexicon is that the vast majority of information related to a word is stored in long-term memory as a concept (Aitchison, 2012). Usually, using this mental lexicon is necessary to comprehend a word's meaning (Stille et al., 2020). Words are encodings of a specific object, such as an event, concept, or thing, with a linguistic label (Abe et al., 1994). Moreover, each word has a distinct pronunciation, spelling, meaning, and syntactic category, such as noun or verb, which is stored in a mental lexicon. In contrast, sound symbolic elements are not associated with any syntactic category and their semantic content is much vaguer compared to standard lexical items–they only provide impressions, such as round, soft, hard, or sharp (Ramachandran and Hubbard, 2001).

On the other hand, established sound symbolic words, such as onomatopoeia, can be considered full-fledged lexical items. The inventory of such words in Indo-European languages is mostly limited to onomatopoeia that express the sounds made by animals (ex., quack-quack) or sounds made while using certain objects (boom or swish). However, many languages outside of the Indo-European family use sound symbolic words known as ideophones that can describe manners, states and emotions (Osaka, 1999; Akita, 2017). In this study, we focus on Japanese, in which a great variety of such sound symbolic words is frequently used. Japanese onomatopoeia and ideophones both usually incorporate sound symbolic elements with consistent semantic associations (ex./k/: gives the image of hitting a hard surface,/g/: gives the impression of increased strength/weight) (Hamano, 2014), and use identical morphological patterns (often a reduplication of 2-mora elements, as in/fuwa-fuwa/) (Tamori and Schourup, 1999). Onomatopoeia and ideophones are established items of Japanese mental lexicon and in this study, we use the term sound symbolic words to refer to both of these categories.

Phonological mapping negativity (PMN) and the N400 response of event-related potentials (ERPs) have been used as indicators in studies of the processing mechanisms of words and pseudowords. A negative evoked brain response connected to language processing, known as PMN, which has an earlier latency than N400, manifests around between 250 and 300 ms (Desroches et al., 2009). According to van den Brink et al. (2001), PMN represents the stage at which, in relation to the lexical selection process, word-form information resulting from an initial phonological analysis and content information derived from the context interact. However, Newman et al. (2003), and Newman and Connolly (2009) discovered that PMN occurs during the preliminary stage of language processing and reflects phonological processing. Thus, some reports of PMN indicate that it is involved in higher-level language processing, while others indicate that it is only involved in phonological processing, which occurs earlier. The interpretation of PMN depends on whether the task focuses solely on the phonological aspect of the word or on determining if the word is an appropriate choice based on higher-level linguistic characteristics.

N400 is a significant negative wave that peaks around 400 ms after the stimulus is presented (Kutas and Federmeier, 2011). The amplitude of N400 has been proven to increase for sentences with mismatched meanings (e.g., "I like coffee with cream and socks") (Kutas and Hillyard, 1980) and is considered to be a brain potential response associated with semantic processing. Some studies of N400 interpret it as reflecting the process of retrieving words from the mental lexicon, whereas others interpret it as reflecting higher-order processes, like context specific semantic processing of vocabulary. In the former interpretation, Petten and Kutas (1990) showed that the amplitude of N400 increases for low frequency words compared to high frequency words, and Kutas and Federmeier (2000) stated that the N400 reflects the activation of vocabulary stored in long-term memory. These studies indicate that N400 may reflect access to the mental lexicon. On the other hand, Brown and Hagoort (1993) found that N400 reflects not only the process of accessing the mental lexicon but also the higher order semantic processing, as N400 is smaller when the priming effect is involved. Hagoort et al. (2004) also reported that both accessing the mental lexicon and integration into semantic knowledge happened at the same time, beginning 300 ms after language presentation. They reported that N400 is caused by the top-down influence of semantic information. In previous studies, the interpretation of PMN and N400 depended on the nature of the task and stimulus words. In their report of N400 in sound symbolic pseudowords, Deacon et al. (2004) showed that non-words derived from existing words show the same level of semantic activation as existing words when participants listen to them. On the other hand, Westbury (2005), in a study using sound symbolic pseudowords, reported that the lexical access stage, which is a preliminary stage of semantic processing, is activated, indicating that the effect of sound symbols is pre-semantic.

In past studies, the interpretation of PMN and N400 was dependent on the nature of the task and stimulus words. Moreover, previous studies on the sound symbolism of pseudowords mostly

involved judgment tasks that use simple pictures of round or angular shapes with no deeper meaning (Kovic et al., 2010; Asano et al., 2015; Sučević et al., 2015). In contrast, the task setting in this study was designed to reflect semantic processing by asking the participants to determine whether the depicted situation or object and matched or mismatched the sound stimuli (existing sound symbolic words or pseudowords). In the sound symbolic word task, there is always a correct word for the situation. On the other hand, the sound-symbolic pseudoword task is a non-word that does not correspond to any of the tasks. Sound symbolic pseudowords are non-existing words created for this study and, unlike established onomatopoeia and ideophones, cannot be considered part of the mental lexicon. Based on previous reports, it is possible that the processing of sound symbolic pseudowords also involves semantic processing; however, it is not clear whether the process is similar to that of established lexemes contained in the mental lexicon.

The goal of this study is to investigate the differences between the semantic processing of established sound symbolic words and pseudowords. Based on the results of previous studies, we suppose that the amplitude of both the PMN and the N400 will be larger in the mismatch condition than in the match condition when assessment of whether the language is consistent with more complex situations requires semantic processing. If so, the PMN appears to reflect neural processing that includes not only phonological processing, which is the precursor to language processing, but also linguistic processing (e.g., word selection). The phonological structure of sound symbolic pseudowords used the common patterns observed in existing Japanese sound symbolic words, and it included sound symbolic phonemes typical in Japanese. The phonological structure of sound symbolic pseudowords may provoke semantic neural processing. Therefore, it would support the hypothesis that sound symbolic pseudowords are semantically processed in the same way as the words if the amplitude of both the PMN and the N400 increased more in the mismatch condition than in the match condition in the sound symbolic pseudoword task, similarly to the sound symbolic words.

In the present study, we aim to clarify the processes reflected by PMN and N400 and to identify the differences in the semantic processing for existing sound symbolic words and sound symbolic pseudowords. The results obtained will give us some clues regarding the neural process of phonological and linguistic features of sound symbolism.

2. Materials and methods

Event-related potentials (ERPs) were measured and analyzed while sound symbolic word and pseudoword comprehension tasks were carried out in order to investigate the specifics of neural processing during lexical processing when a picture of an event and a speech sound are consistent (match) or inconsistent (mismatch).

2.1. Participants

A total of 30 healthy university students, ranging in age from 19 to 22, took part in the study. Participants were paid volunteers recruited at the International University of Health and Welfare. They were all native Japanese speakers and had no medical or mental health issues in the past. They had normal or corrected-to-normal vision, and they had normal hearing. The International University of Health and Welfare's Ethical Review Committee gave its approval to this study. The experiment was explained to participants orally and in writing, and they voluntarily signed an informed consent form. This study was conducted in accordance with the ethical principles of the Declaration of Helsinki. Three participants were excluded from the study because their correct response rate to the task was less than 80%, and seven participants were excluded because the difference between their correct response rate to the sound symbolic word task and the sound symbolic pseudo-word task was more than 5%. In addition, two had many artifacts (less than 80% valid epochs) and one had incomplete task results. We used data from a total of 17 subjects for the final analysis.

2.2. Stimulus

In both sound symbolic word and pseudoword tasks, pictures and sounds were presented. In both task types, match conditions—where the appropriate sound for the picture was presented—and mismatch conditions—where a sound unrelated to the picture was presented—were used.

There were 160 stimulus word sounds, 80 each of sound symbolic words and pseudowords. The Japanese Onomatopoeia Dictionary (Ono, 2007) was used to choose existing sound symbolic words. The most prevalent type, the two-mora reduplication sound symbolic word (e.g., /fuwafuwa/, meaning "soft and puffy") (Tamori and Schourup, 1999), was used as the sound stimulus. We chose words that expresses various receptive senses like touch and hearing, and the meaning is relatively clear and not context-dependent. Since sound symbolic words used in Japanese to express feelings often change their meaning depending on the situation (like/moyamoya/: thoughts and memories to be fuzzy, but also the feeling of unease or pent-up lust, depending on context), we limited our stimuli to sound symbolic words that express clear sensory states or manners. Established sound symbolic words that were used as stimulus words were modified to create sound symbolic pseudowords. The first and third syllables were not altered during creation; the second and fourth syllable were. Additionally, in keeping with Hamano's (2014) research, we changed the second syllable to a sound symbolic phoneme that elicits the same semantic impression, for example, roundness or sharpness, to preserve the word's overall semantic characteristics (e.g., /fuwafuwa/versus/fuhafuha/) (**Supplementary Data Sheet 1**).

A Text to speech program (Azure, Text to speech, Microsoft Inc., Redmond, WA, USA) was used to synthesize the voice of the stimulus words. All stimulus words were controlled using version: 1.0, voice type: male voice, prosody rate: 50%, pitch: 0%, used language: Japanese. We manipulated the reading speed (between 0.5 and 1.5) to fit the length of a sample to around 500 ms. In order to use the generated speech samples as stimuli for the task, the length of the speech sample was then adjusted to exactly

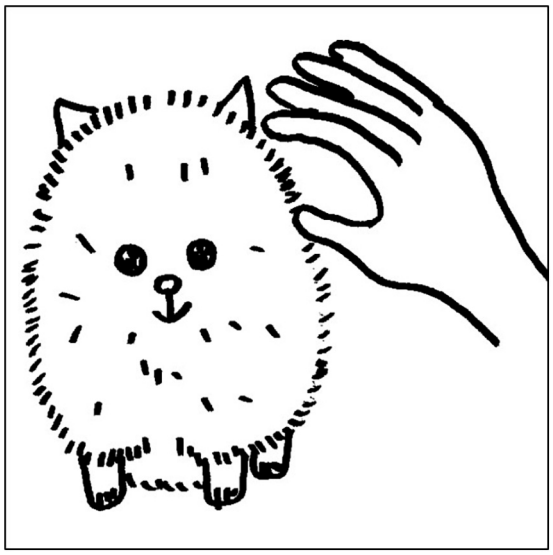


FIGURE 1
An example of stimulus illustration. Existing Japanese sound symbolic word–onomatopoeia:/fuwa fuwa/, which means "fluffy" in English, the corresponding sound symbolic pseudoword:/fuha fuha/. In this illustration,/fuwa- fuwa/indicates the dog's hair is soft.

500 ms in Audacity 3.2, an open-source audio file editing program.[1] Fade in/out effects were added to the first and last 10 ms of each speech sample.

Both the sound symbolic word and pseudoword tasks made use of the same stimulus pictures. The meaning of established lexemes was verified in a Japanese onomatopoeia dictionary, and corresponding black-and-white pictures were created to act as the stimulus pictures (**Figure 1**).

There were 160 conventional sound symbolic word tasks, 80 of which were in the match condition and contained the appropriate stimulus word and picture combination. The other 80 tasks were in the mismatch condition and contained the inappropriate stimulus word and picture combination. The sound symbolic pseudoword tasks were created by replacing the existing word with a corresponding pseudoword with alternate phonemes for both the match and mismatch conditions. Consequently, there were 160 sound symbolic pseudoword tasks, 80 of which were in the match condition and 80 in the mismatch condition. Both sound-symbolic words and sound symbolic pseudowords tasks are presented twice for 40 words, for a total of 80 tasks. All of the auditory and visual stimuli were the same in the match and mismatch conditions but in different combination. The sound symbolic word and pseudoword tasks' match and mismatch conditions added up to 320 tasks in total. In the sound symbolic word and sound symbolic pseudoword tasks, the stimulus words were presented randomly in each task.

2.3. Procedure

Participants in the experiment were seated in a soundproofed, electrically shielded room. While the participants were deciding whether the picture on the display matched the sound coming from their earphones, their brain waves were monitored. Button-pushing tasks were given to participants to ensure they were paying close attention during the experiment. The participants were monitored while the tasks were carried out to make sure the participant was paying attention. Participants practiced on several words in both the sound symbolic task and pseudoword task prior to the actual experiment. Participants began the actual experiment when they understood the content of the task.

The experimental design is schematically represented in **Figure 2**. Stimulus pictures were shown on a 23.0″ LCD screen with a resolution of 629 × 629 pixels (FlexScan EV2316, EIZO Inc., CA, USA, JP, resolution: 1,920 × 1,080), Subject-to-screen distance: range 50–85 cm, average 70 cm, and audio was played through ER-3A insert earphones (Etymotic Research Inc., IL, USA). The experiment started with the display of a stimulus picture. Next, 550 ms after the stimulus picture was displayed, a stimulus word sound (sound symbolic words or pseudowords) was presented for 500 ms. After the audio presentation, a white screen was displayed for 100 ms. Following this, a message prompting the user to interact ("Press 1 for a match, 4 for a mismatch") was displayed in MS Gothic font size 80. Participants immediately responded by pressing the 1 or 4 key on the keyboard. The message disappeared once the participant chose a response. The next task started after 2,000 ms of a white screen. The Medical Try System Multi Trigger System Ver. 2 (MTS0410, Medical Try System, Co., Ltd., Japan) was used to control the task presentation, and the match and mismatch conditions were distributed at random.

During the experiment, participants were told to avoid blinking, swallowing saliva, or unnecessary movement. In addition, the sound symbolic word and sound symbolic pseudoword tasks were counterbalanced by randomly administering the sound symbolic word and sound symbolic pseudoword tasks to each subject. A break was always included between the pseudoword task segment and the existing word task segment. Additionally, breaks were also taken during the segments upon the participant's request. The experiment took approximately 30–40 min to complete.

2.4. Acquiring EEG data

A unipolar recording of EEG signals was made at the Cz position. Because PMN and N400 are greatest at the parietal site (Dumay et al., 2001: D'Arcy et al., 2004; Kutas and Federmeier, 2011), PMN and N400 were measured at Cz in this experiment. Electrodes placed on the mastoids served as a reference, and the grounding electrode was positioned on the forehead. EEG signals were captured on a Neurofax EEG-1200 system (Nihon Koden, Co., Ltd., Japan). Before starting the task, it was determined that the contact impedance between the skin and electrode was less than 10 kΩ. The signal was converted to a digital signal at 1 kHz after being bandpass filtered between 0.3 and 30 Hz. The data were divided into epochs of 700 ms, after the speech stimulus presented at 100 ms. The baseline was set based on the 100 ms prior to the presentation of the speech stimulus. Reactions associated with blinking, eye movements, etc. were rejected offline based on electrooculography. Upon visual inspection, epochs with excessive noise were removed from the analysis. The total number of data points in each condition averaged 69.7 epochs (range: 79–65).

1 https://www.audacityteam.org

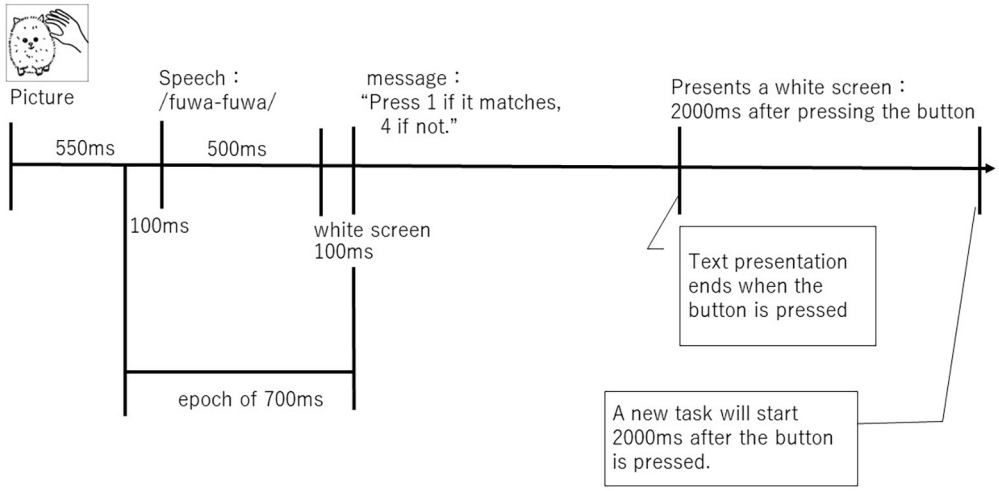


FIGURE 2
A schematic figure of audio-visual stimulation.

2.5. Statistical analysis

In this study, PMN and N400 components are of particular interest. Therefore, the maximum amplitude between 250 and 350 ms after the sound onset was defined as PMN amplitude (Lee et al., 2012) and its latency was defined as PMN latency. The average amplitude between 350 and 500 ms after the sound onset was defined as N400 amplitude (D'Arcy et al., 2004; Asano et al., 2015; Manfredi et al., 2017). We analyzed PMN and N400 using Matlab R2022b (MathWorks, Inc., USA). For the existing word and pseudoword tasks, as well as for each of the amplitudes and emergent latencies, PMN was subjected to a two-way ANOVA (two factors: existing word task/pseudoword task x match/mismatch). The N400 amplitude was similarly analyzed by a two-way ANOVA using sound type (existing word vs. pseudoword) and audio-visual matching (match vs. mismatch) as factors. The statistical significance for all two-way ANOVA was expressed as a p-value of less than 0.05. Effect size was calculated using partial eta squared (η_p^2) and generalized eta squared (η_G^2) (Bakeman, 2005).

In order to more clearly analyze the differences between sound symbolic words and pseudowords, planned comparisons were conducted using paired t-tests between matching and mismatching conditions in each existing word and pseudoword task, as well as between the existing word and pseudoword task in each matching and mismatching condition. Cohen's d was used to examine effect size, with 0.2 small, 0.5 medium, and 0.8 large (Cohen, 1988). We used IBM SPSS Statistics Version 29 for Windows (IBM Corp., Armonk, NY, USA).

3. Results

3.1. Behavioral results

The average score for the sound symbolic word task for the participants was 151, and the average score percentage was 95%. On the other hand, the mean score for the sound-symbolic pseudoword task was 149, and the mean percentage score was 93%.

3.2. Electrophysiological results

Figure 3 displays the mean Cz brain wave reading for the match and mismatch conditions of established sound symbolic words and pseudowords from 100 ms prior to speech presentation to 600 ms following presentation.

A significant main effect of sound type (existing words vs. pseudoword) was observed with the maximum amplitude in the PMN (250–350 ms) component of the ERP [$F_{(1,\ 16)} = 11.89$, $p = 0.003$, $\eta_p^2 = 0.426$, $\eta_G^2 = 0.030$] (**Supplementary Data Sheet 2**). The main effect of audio-visual matching (match vs. mismatch) was also observed [$F_{(1,\ 16)} = 10.07$, $p = 0.006$, $\eta_p^2 = 0.386$, $\eta_G^2 = 0.030$]. No significant interaction was found between sound type and audio-visual matching [$F_{(1,\ 16)} = 0.21$, $p = 0.653$, $\eta_p^2 = 0.0013$, $\eta_G^2 < 0.001$]. A significant main effect of audio-visual matching was observed for the latency at which the PMN peak appeared [$F_{(1,\ 16)} = 19.46$, $p < 0.001$, $\eta_p^2 = 0.549$, $\eta_G^2 = 0.085$] (**Supplementary Data Sheet 3**). There was no significant difference in the sound type [$F_{(1,\ 16)} = 0.03$, $p = 0.875$, $\eta_p^2 = 0.002$, $\eta_G^2 < 0.001$] and no significant interaction between sound type and audio-visual matching [$F_{(1,\ 16)} = 0.42$, $p = 0.529$, $\eta_p^2 = 0.025$, $\eta_G^2 = 0.002$]. The mean amplitude in the N400 range (350–500 ms) showed significant main effects for sound type [$F_{(1,\ 16)} = 16.24$, $p = 0.001$, $\eta_p^2 = 0.504$, $\eta_G^2 = 0.064$] and audio-visual matching [$F_{(1,\ 16)} = 36.68$, $p < 0.001$, $\eta_p^2 = 0.696$, $\eta_G^2 = 0.070$] (**Supplementary Data Sheet 4**). There was no significant interaction between sound type and audiovisual agreement [$F_{(1,\ 16)} = 0.84$, $p = 0.374$, $\eta_p^2 = 0.050$, $\eta_G^2 = 0.002$]. The amplitudes of PMN and N400 showed significant differences for sound type and audio-visual matching, but no significant interaction between them. The PMN latencies showed a significant difference for audio-visual matching, but no significant difference for sound type and no significant interaction between them.

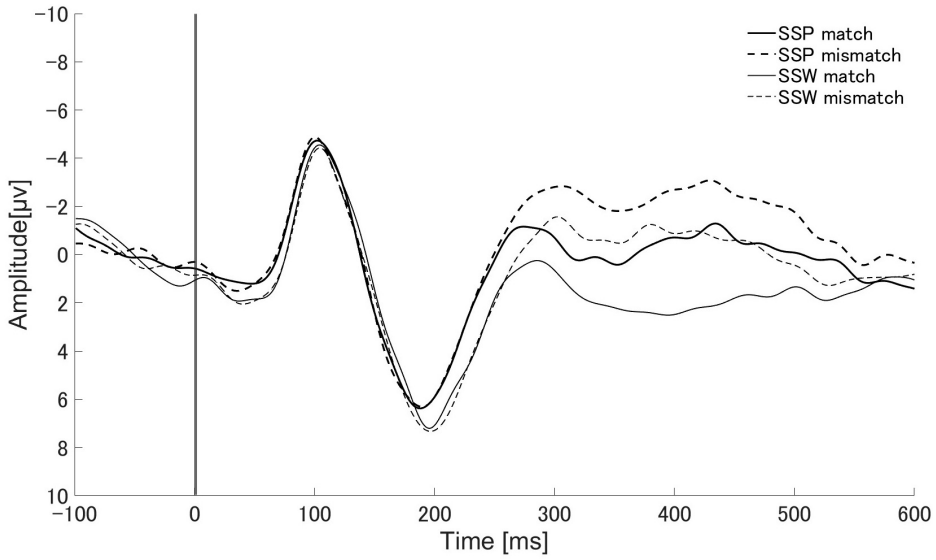


FIGURE 3
Auditory evoked responses elicited by sound symbolic words (SSW: gray lines) and sound symbolic pseudowords (SSP: black lines). Solid and dotted lines represent audio-visual matching and non-matching conditions, respectively. SSP match: match condition of sound symbolic pseudowords task. SSP mismatch: mismatch condition of sound symbolic pseudowords task. SSW match: match condition of sound symbolic words task. SSW mismatch: mismatch condition of sound symbolic words task.

There were significant differences in the amplitudes of PMN between the match and mismatch conditions for the sound symbolic words [t (16) = 2.93, p = 0.010, Cohen's d = 0.711] and pseudowords, [t (16) = 2.55, p = 0.021, d = 0.619]. Furthermore, there were significant PMN amplitude differences between the sound symbolic word and pseudoword tasks for the match condition [t (16) = 3.07, p = 0.007, d = 0.744] and mismatch conditions [t (16) = 2.73, p = 0.015, d = 0.663] (**Figure 4**).

Similar results were obtained for N400. We found significant N400 amplitude differences between the match and mismatch conditions for sound symbolic words [t (16) = 6.00, $p < 0.001$, d = 1.455] and pseudowords [t (16) = 2.87, p = 0.011, d = 0.697], and between the sound symbolic word and pseudoword tasks for the match [t (16) = 4.66, $p < 0.001$, d = 1.129] and mismatch conditions [t (16) = 2.25, p = 0.039, d = 0.546] (**Figure 5**).

There were significant differences in PMN latency between the match and mismatch conditions for sound symbolic words [t (16) = −3.64, p = 0.002, d = −0.882] and pseudowords [t (16) = −2.58, p = 0.020, d = −0.626], but no significant difference in PMN latency between the sound symbolic word and pseudoword tasks for the match [t (16) = −0.49, p = 0.632, d = −0.119] nor mismatch conditions [t (16) = 0.13, p = 0.902 d = 0.030] (**Figure 6**).

4. Discussion

This study investigated the semantic processing of both existing sound symbolic words and sound symbolic pseudowords with similar phonological structures using ERPs elicited while making semantic assessments of whether the sound stimuli are a match or mismatch with a picture of an event. During the task of determining whether pictures of events and speech sounds match or mismatch PMN and N400 were measured, and the semantic processing of sound symbolic words and pseudowords was examined and compared. The results confirm the hypothesis that PMN reflects not only phonological but also word processing, while N400 reflects contextual semantic processing. Furthermore, both the sound-symbolic pseudowords (non-words) and the sound-symbolic words (existing words) showed an amplitude increase in the mismatch condition compared to the match condition, suggesting that pseudowords, like existing words, may undergo semantic processing.

Connolly et al. (2001) discovered that when phonetic features match between the target word and a candidate word, less phonological processing work is needed, and the PMN is reduced; when they do not match, more thorough phonological analysis is needed, and a larger PMN is generated. This study presented pictures presented prior to sounds, so pictures are prime for target words. The PMN results of the present study confirmed the findings of Connolly et al. (2001) by demonstrating an increase in amplitude in the mismatch condition compared to the match condition for conventional sound symbolic words. Thus, the results of this study support the hypothesis of van den Brink et al. (2001) and Kujala et al. (2004) that PMNs reflect a continuum of processing at the acoustic, phonological, and word levels, suggesting that PMNs do not represent phonological processing alone, but rather a process that reflects access to the mental lexicon.

As for the N400 following the PMN, it has been reported that when both contextual and lexical information for a word is available, the contextual information influences the N400, decreasing its amplitude in words that match the context (van Berkum et al., 1999). Therefore, Kutas and Federmeier (2000) describe the N400 as reflecting "integration into the context." Deacon et al. (2004) stated that it is generated by orthographic/phonological analysis and is affected by semantic information in a "top-down" manner. In this study, in the existing

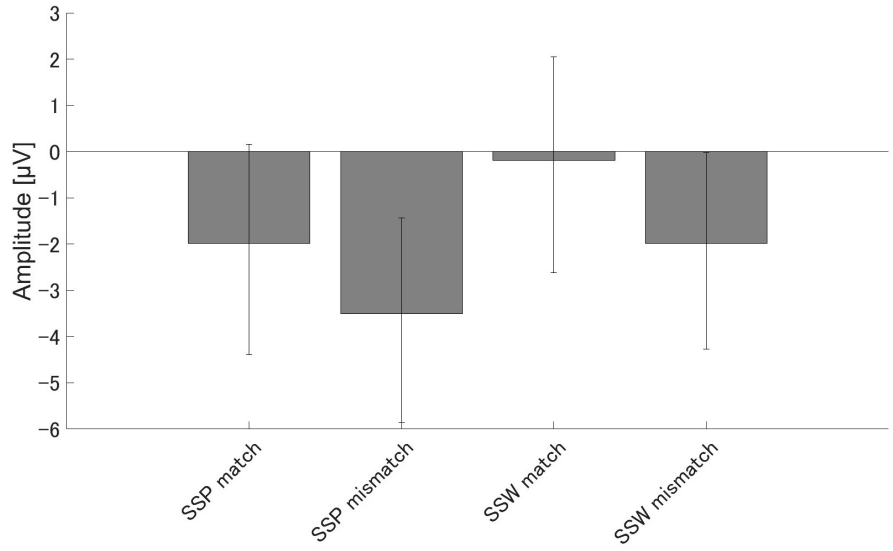


FIGURE 4
Group means (N = 17) of PMN amplitudes. Group means (N = 17) of PMN, the maximum amplitude elicited 250–350 ms after speech presentation. The error bars denote 95% confidence intervals. SSP, Sound symbolic pseudowords, SSW, sound symbolic words.

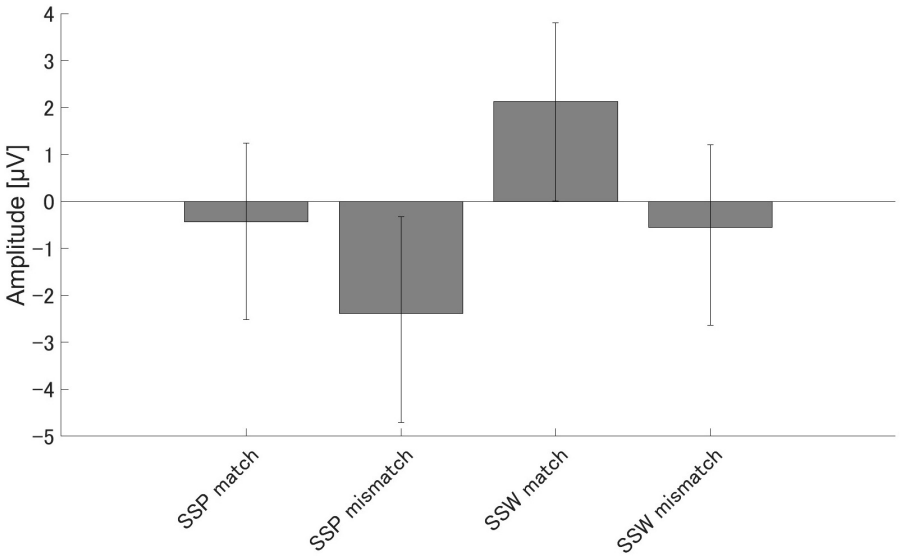


FIGURE 5
Group means (N = 17) of N400 amplitudes. Group means (N = 17) of N400 amplitudes elicited 350–500 ms after speech sound presentation. The error bars denote 95% confidence intervals.

word task, words are presented in both the match and mismatch conditions, and match/mismatch judgments must be integrated into the context of the pictures. Therefore, the difference in N400 between the match and mismatch conditions in this task may reflect not only access to the mental lexicon, but also semantic processing (integration into the context). The results obtained in this study give clues as to what processes PMN and N400 reflect: PMN may reflect the phonological and lexical processing, while N400 may reflect a higher-order processing, the process of integration into context.

As a word is being presented auditorily, several candidate words (cohort) that start with the same phoneme are activated in the listener's mental lexicon. As the phonemes are perceived one at a time, the candidate vocabulary is reduced to the final pertinent term (Marslen-Wilson, 1987). This process has also been supported by brain wave evaluation, which discovered that for the semantic processing of vocabulary, the responses to the phoneme of the words change over time. In the McMurray et al. (2022) experiment, after target words were presented auditorily, a word phonologically similar (cohort) to the target word, and a word unrelated to the target word was presented in written form. The task was to choose which one is closer to the phonetically presented word. The results showed a characteristic pattern in which the target word and its phonologically similar competitor word (cohort) are

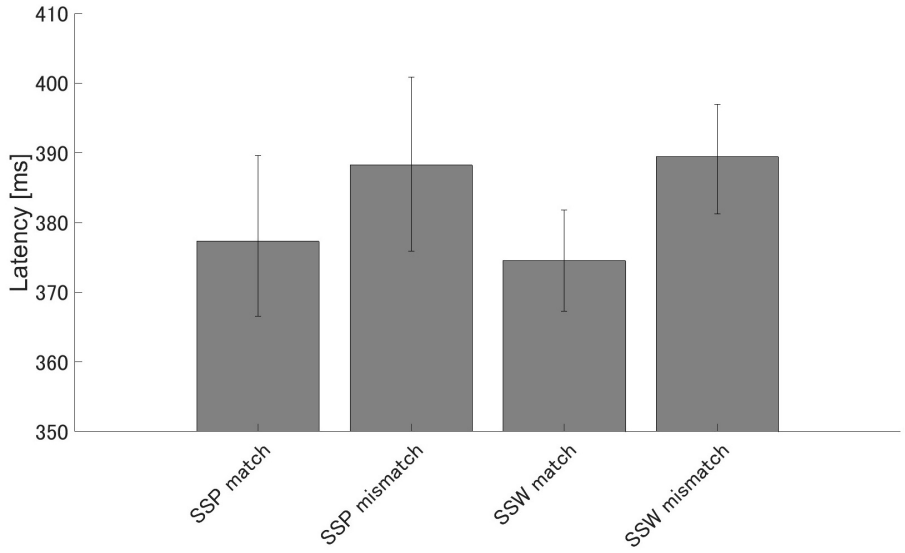


FIGURE 6
Group means (N = 17) of PMN latencies. Group means (N = 17) of PMN latencies that elicited maximum amplitude 250–350 ms after speech presentation. The error bars denote 95% confidence intervals. SSP: sound symbolic pseudowords, SSW: sound symbolic words.

active immediately after word onset (at levels that were greater than unrelated items), and at 100 ms activation levels were the same. However, 500 ms after word presentation, the response to the competing cohort is suppressed and the target is selected. If we interpret the findings of this study through the lens of the processing of words presented auditorily, PMN seems to reflect the activation of word candidates (cohorts) from the point of speech input and the process of lexical judgments in the search for the corresponding word. The subsequent N400 reflects the higher-level semantic processing of contextual integration.

Although sound symbolic pseudowords are non-words, both PMN and N400 increased in amplitude in the mismatch condition relative to the matching condition, same as in the case of existing sound symbolic words. This suggests that even though sound symbolic pseudowords are non-words, they are processed similarly to standard words in terms of semantics. Access to meaning through automatic activation of phonological information has been reported for non-words with phonological features similar to those of real words (Kujala et al., 2004). The results of this study show that this is also the case for physiological brain responses. This experiment included a situation picture and word/pseudoword matching task. The higher processing effort required for the mismatch condition compared to the match condition for the sound-symbol pseudowords suggests that word-level processing, such as retrieving word candidates, was also performed for non-word sound symbolic pseudowords. Furthermore, the increase in N400 amplitude in the mismatch condition for the sound symbolic pseudowords is consistent with Deacon et al. (2004). Therefore, it has been suggested that the sound-symbolic pseudowords are indicative of processing at a higher level, that is, semantic processing, including integration into the context. In the difference between the match and mismatch conditions for sound symbol pseudowords indicates that sound symbol pseudowords are not judged on a simple impression level (such as spiky/round) or activated by other words with similar spellings (phonological structures) in the lexicon, but are judged on the similar context basis as actual words.

Language is the process of ascribing a label to a situation or object. It emerges as a result of semantic integration of auditory and visual information. Attentional processes are also involved (Wang et al., 2017; McCormick et al., 2018). Multiple integration processes in various cortical areas are involved in semantic integration, which is modulated by attention (Xi et al., 2019). According to McCormick et al. (2021), greater attention may be needed for the adjusting and controlling of information processing for discrepancies between visual and auditory information when attention is focused on auditory information. In this task, the concept that a word meant was presented as an illustration, and if the word matched the presented illustration, it was processed quickly and with low processing load due to the top-down effect of context. However, if the word did not match, the participants had to reinterpret the picture and search again for a possible match with the scene/event, which would have required more attention due to the processing load of adapting and suppressing information processing. As shown in **Figure 1**, in addition to representing the softness of the dog's fur, the picture could also represent the act of touching the dog or the feelings of the person touching the dog. It has been noted that higher-level processing of elements like syntax and contextual meaning is also carried out during the semantic processing of a word as sounds are entered one at a time and the candidate vocabulary is narrowed down; thus contextual-level semantic information has a significant impact on semantic processing (Dahan and Tanenhaus, 2004; Costa et al., 2009; McMurray et al., 2022). We assume that in the existing word task and pseudoword task, the increased amplitude and latency delay in the mismatch condition may have reflected the search for alternative interpretation of what the stimulus picture may represent, the top-down influence of different possibilities, and the load in narrowing the vocabulary. In addition, the present results showed that the amplitude of both the PMN

and the N400 was increased for sound-symbolic pseudowords (non-words) compared to sound-symbolic words (words). Sound-symbolic pseudowords are not words that are stored in the mental lexicon. Therefore, the expectation of this result was that sound-symbolic pseudowords would require more extensive processing (attention). In this study, we were able to show the differences between the semantic processing of existing sound symbolic words and pseudowords in terms of the physiological responses of the brain. The similarities between sound symbolic pseudowords and conventional words suggest that both are processed temporally from the input of phonemes to the contextual semantic processing. The difference was that the processing load for sound symbolic pseudowords was greater than that for sound symbolic words. In order to clarify this difference, it will be necessary to identify the brain regions involved in the semantic processing of both sound symbolic words and pseudowords, which is an issue for the future. Although fMRI at high spatial resolution or other methods could be used to search for the region of the brain responsible for semantic processing of sound symbolic pseudowords, it is challenging to distinguish PMN from N400 due to inferior temporal resolution. Therefore, a more detailed analysis using a combination of various neuroimaging techniques is needed. In addition, Japanese is a language that is rich in onomatopoeia and ideophones and commonly makes use of sound symbolism. This could be why sound symbolic words are processed is a way similar to the semantic processing of other words of the language. Thus, a cross-linguistic approach is necessary for future studies in order to clarify the semantic processing of sound symbolism.

5. Conclusion

In this study, the semantic processing of sound symbolic words included in the mental lexicon is compared to the semantic processing of sound symbolic pseudowords, which have no concrete meaning stored in the lexicon, but use sounds that are consistently associated with a certain meaning. It also clarifies the processes reflected by PMN and N400, which are used in many studies of lexical semantic processing. The findings of this study suggest that PMNs may reflect not only phonological processing, but also further processing up to the point of accessing vocabulary, whereas N400 may reflect final semantic judgments. Furthermore, similar responses toward sound symbolic words and pseudowords were observed in the match and mismatch conditions, indicating that both of them undergo semantic processing.

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 humans were approved by the International University of Health and Welfare's Ethical Review Committee. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

KS and HO contributed to conception and design of the study. KS and SK performed the experiments, collected the data, and performed the statistical analysis. KS wrote the first draft of the manuscript. KS, SK, JI, MP and HO wrote sections of the manuscript. MP edited English language. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

This work was supported by the Japan Society for the Promotion of Science (JSPS) Grants in Aids for Scientific Research (KAKENHI) Grant Number 21K13620.

Acknowledgments

We gratefully acknowledge our participants for their diligent cooperation and the work of present members of our laboratory.

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnhum.2023.1208572/full#supplementary-material

References

Abe, J., Momouti, Y., and Kaneko, Y. (1994). *Human language information processing- Cognitive Science of Language Understanding*. Tokyo: Science Inc.

Aitchison, J. (2012). *Words in the Mind: An Introduction to the Mental Lexicon*, 4th Edn. New Jersey: Wiley-Blackwell.

Akita, K. (2017). "Do foreign languages have onomatopoeia?," in *The Mystery of Onomatopoeia- From Pikachu to Mohammed- Iwanami Science Library 261*, ed. H. Kubozono (Tokyo: Iwanamisyoten).

Asano, M., Imai, M., Kita, S., Kitajo, K., Okada, H., and Thierry, G. (2015). Sound symbolism scaffolds language development in preverbal infants. *Cortex* 63, 196–205. doi: 10.1016/j.cortex.2014.08.025

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behav. Res. Methods* 37, 379–384. doi: 10.3758/BF03192707

Brown, C., and Hagoort, P. (1993). The processing nature of the n400: evidence from masked priming. *J. Cogn. Neurosci.* 5, 34–44. doi: 10.1162/jocn.1993.5.1.34

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York, NY: Routledge Academic.

Connolly, J. F., Service, E., D'Arcy, R. C., Kujala, A., and Alho, K. (2001). Phonological aspects of word recognition as revealed by high-resolution spatio-temporal brain mapping. *Neuroreport* 12, 237–243. doi: 10.1097/00001756-200102120-00012

Costa, A., Strijkers, K., Martin, C., and Thierry, G. (2009). The time course of word retrieval revealed by event-related brain potentials during overt speech. *Proc. Natl. Acad. Sci. U. S. A.* 106, 21442–21446. doi: 10.1073/pnas.0908921106

Dahan, D., and Tanenhaus, M. K. (2004). Continuous mapping from sound to meaning in spoken-language comprehension: immediate effects of verb-based thematic constraints. *J. Exp. Psychol.* 30, 498–513. doi: 10.1037/0278-7393.30.2.498

D'Arcy, R. C., Connolly, J. F., Service, E., Hawco, C. S., and Houlihan, M. E. (2004). Separating phonological and semantic processing in auditory sentence processing: a high-resolution event-related brain potential study. *Hum. Brain Mapp.* 22, 40–51. doi: 10.1002/hbm.20008

Deacon, D., Dynowska, A., Ritter, W., and Grose-Fifer, J. (2004). Repetition and semantic priming of nonwords: Implications for theories of N400 and word recognition. *Psychophysiology* 41, 60–74. doi: 10.1111/1469-8986.00120

Desroches, A. S., Newman, R. L., and Joanisse, M. F. (2009). Investigating the time course of spoken word recognition: electrophysiological evidence for the influences of phonological similarity. *J. Cogn. Neurosci.* 21, 1893–1906. doi: 10.1162/jocn.2008.21142

Dumay, M., Benraïss, A., Barriol, B., Colin, C., Radeau, M., and Besson, M. (2001). Behavioral and electrophysiological study of phonological priming between bisyllabic spoken words. *J. Cogn. Neurosci.* 13, 121–143. doi: 10.1162/089892901564117

Hagoort, P., Hald, L., Bastiaansen, M., and Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science* 304, 438–441. doi: 10.1126/science.1095455

Hamano, S. (2014). *Onomatopoeia in Japanese: Sound Symbolism and Structure*. Tokyo: Kurosio Publishers.

Hinton, L., Nichols, J., and Ohala, J. (1994). *Sound symbolism*. Cambridge: Cambridge University Press.

Hockett, C. F. (1960). The origin of speech. *Sci. Am.* 203, 89–96.

Köhler, W. (1947). *Gestalt psychology*, 2nd Edn. New York: Liveright Publishing.

Kovic, V., Plunkett, K., and Westermann, G. (2010). The shape of words in the brain. *Cognition* 114, 19–28. doi: 10.1016/j.cognition.2009.08.016

Kujala, A., Alho, K., Service, E., Ilmoniemi, R. J., and Connolly, J. F. (2004). Activation in the anterior left auditory cortex associated with phonological analysis of speech input: Localization of the phonological mismatch negativity response with MEG. *Brain Res. Cogn. Brain Res.* 21, 106–113. doi: 10.1016/j.cogbrainres.2004.05.011

Kutas, M., and Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends Cogn. Sci.* 4, 463–470. doi: 10.1016/s1364-6613(00)01560-6

Kutas, M., and Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annu. Rev. Psychol.* 62, 621–647. doi: 10.1146/annurev.psych.093008.131123

Kutas, M., and Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207, 203–205. doi: 10.1126/science.7350657

Lee, J. Y., Harkrider, A. W., and Hedrick, M. S. (2012). Electrophysiological and behavioral measures of phonological processing of auditory nonsense V-CV-VCV stimuli. *Neuropsychologia* 50, 666–673. doi: 10.1016/j.neuropsychologia.2012.01.003

Manfredi, M., Cohn, N., and Kutas, M. (2017). When a hit sounds like a kiss: An electrophysiological exploration of semantic processing in visual narrative. *Brain Lang.* 169, 28–38. doi: 10.1016/j.bandl.2017.02.001

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition* 25, 71–102. doi: 10.1016/0010-0277(87)90005-9

McCormick, K., Lacey, S., Stilla, R., Nygaard, L. C., and Sathian, K. (2018). Neural basis of the crossmodal correspondence between auditory pitch and visuospatial elevation. *Neuropsychologia* 112, 19–30. doi: 10.1016/j.neuropsychologia.2018.02.029

McCormick, K., Lacey, S., Stilla, R., Nygaard, L. C., and Sathian, K. (2021). Neural basis of the sound-symbolic crossmodal correspondence between auditory pseudowords and visual shapes. *Multisens. Res.* 35, 29–78. doi: 10.1163/22134808-bja10060

McMurray, B., Sarrett, M. E., Chiu, S., Black, A. K., Wang, A., Canale, R., et al. (2022). Decoding the temporal dynamics of spoken word and nonword processing from EEG. *Neuroimage* 260:119457. doi: 10.1016/j.neuroimage.2022.119457

Newman, R. L., and Connolly, J. F. (2009). Electrophysiological markers of pre-lexical speech processing: evidence for bottom-up and top-down effects on spoken word processing. *Biol. Psychol.* 80, 114–121. doi: 10.1016/j.biopsycho.2008.04.008

Newman, R. L., Connolly, J. F., Service, E., and McIvor, K. (2003). Influence of phonological expectations during a phoneme deletion task: evidence from event-related brain potentials. *Psychophysiology* 40, 640–647. doi: 10.1111/1469-8986.00065

Ono, M. (2007). *Imitative words and Mimetic words 4500 Japanese Onomatopoeia Dictionary*. Tokyo: SHOGAKUKAN.

Osaka, M. (1999). "Multilingual comparison of onomatopoeia and ideophone," in *Exploration of the language of sensibility- Where is the heart in onomatopoeia and ideophone?*, ed. N. Osaka (Tokyo: Shinyousya).

Petten, C. V., and Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Mem.Cogn.* 18, 380–393. doi: 10.3758/bf03197127

Pinker, S. (1999). *Words and Rules: the ingredients of language*. New York: Basic Books.

Ramachandran, V., and Hubbard, E. (2001). Synaesthesia – A window into perception, thought and language. *J. Conscious. Stud.* 8, 3–34.

Sidhu, D. M., and Pexman, P. M. (2018). Five mechanisms of sound symbolic association. *Psychon. Bull. Rev.* 25, 1619–1643.

Stille, C. M., Bekolay, T., Blouw, P., and Kröger, B. J. (2020). Modeling the mental lexicon as part of long-term and working memory and simulating lexical access in a naming task including semantic and phonological cues. *Front. Psychol.* 11:1594. doi: 10.3389/fpsyg.2020.01594

Sučević, J., Savić, A. M., Popović, M. B., Styles, S., and Ković, V. (2015). Balloons and bavoons versus spikes and shikes: ERPs reveal shared neural processes for shape-sound-meaning congruence in words, and shape-sound congruence in pseudowords. *Brain Lang.* 145, 11–22. doi: 10.1016/j.bandl.2015.03.011

Tamori, I., and Schourup, L. (1999). *Onomatopoeia - form and meaning- Japanese-English Subject Matter Study Series*. Tokyo: Kuroshio Publishing.

van Berkum, J. J., Hagoort, P., and Brown, C. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *J. Cogn. Neurosci.* 11, 657–671. doi: 10.1162/089892999563724

van den Brink, D., Brown, C. M., and Hagoort, P. (2001). Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. *J. Cogn. Neurosci.* 13, 967–985. doi: 10.1162/089892901753165872

Wang, H., Zhang, G., and Liu, B. (2017). Influence of auditory spatial attention on cross-modal semantic priming effect: Evidence from N400 effect. *Exp. Brain Res.* 235, 331–339. doi: 10.1007/s00221-016-4792-4

Westbury, C. (2005). Implicit sound symbolism in lexical access: Evidence from an interference task. *Brain Lang.* 93, 10–19. doi: 10.1016/j.bandl.2004.07.006

Xi, Y., Li, Q., Gao, N., He, S., and Tang, X. (2019). Cortical network underlying audiovisual semantic integration and modulation of attention: An fMRI and graph-based study. *PLoS One* 14:e0221185. doi: 10.1371/journal.pone.0221185

frontiers | Frontiers in Human Neuroscience

TYPE Original Research
PUBLISHED 12 October 2023
DOI 10.3389/fnhum.2023.1203100

OPEN ACCESS

EDITED BY
Tetsuo Kida,
Institute for Developmental Research, Japan

REVIEWED BY
Dalila Burin,
Trinity College Dublin, Ireland
Yann Coello,
Université Lille Nord de France, France

*CORRESPONDENCE
Tsukasa Kimura
✉ kimura@ai.sanken.osaka-u.ac.jp

RECEIVED 10 April 2023
ACCEPTED 21 September 2023
PUBLISHED 12 October 2023

CITATION
Kimura T and Katayama J (2023) Visual stimuli in the peripersonal space facilitate the spatial prediction of tactile events—A comparison between approach and nearness effects. *Front. Hum. Neurosci.* 17:1203100. doi: 10.3389/fnhum.2023.1203100

Visual stimuli in the peripersonal space facilitate the spatial prediction of tactile events—A comparison between approach and nearness effects

Tsukasa Kimura[1]* and Jun'ichi Katayama[2,3]

[1]The Institute of Scientific and Industrial Research, Osaka University, Ibaraki, Japan, [2]Department of Psychological Science, Kwansei Gakuin University, Nishinomiya, Japan, [3]Center for Applied Psychological Science (CAPS), Kwansei Gakuin University, Nishinomiya, Japan

Previous studies reported that an object in one's peripersonal space (PPS) attracts attention and facilitates subsequent processing of stimuli. Recent studies showed that visual stimuli approaching the body facilitated the spatial prediction of subsequent tactile events, even if these stimuli were task-irrelevant. However, it is unclear whether the approach is important for facilitating this prediction or if the simple existence of stimuli within the PPS is what matters. The present study aimed to scrutinize the predictive function of visuo–tactile interaction in the PPS by examining the effects of visual stimuli approaching the hand and of visual stimuli near the hand. For this purpose, we examined electroencephalograms (EEGs) during a simple reaction time task for tactile stimuli when visual stimuli were presented approaching the hand or were presented near the hand, and we analyzed event-related spectral perturbation (ERSP) as an index of prediction and event-related brain potentials (ERPs) as an index of attention and prediction error. The tactile stimulus was presented to the left (or right) wrist with a high probability (80%) and to the opposite wrist with a low probability (20%). In the approach condition, three visual stimuli were presented approaching the hand to which the high-probability tactile stimulus was presented; in the near condition, three visual stimuli were presented repeatedly near the hand with the high-probability tactile stimulus. Beta-band activity at the C3 and C4 electrodes, around the primary somatosensory area, was suppressed before the onset of the tactile stimulus, and this suppression was larger in the approach condition than in the near condition. The P3 amplitude for high-probability stimuli in the approach condition was larger than that in the near condition. These results revealed that the approach of visual stimuli facilitates spatial prediction and processing of subsequent tactile stimuli compared to situations in which visual stimuli just exist within the PPS. This study indicated that approaching visual stimuli facilitates the prediction of subsequent tactile events, even if they are task-irrelevant.

KEYWORDS

peripersonal space, prediction, multimodal interaction, event-related desynchronization (ERD), event-related spectral perturbation (ERSP), time-frequency analysis, event-related brain potential (ERP)

1. Introduction

Primates, including humans, maintain a special representation of the space surrounding the body called peripersonal space (PPS; Brain, 1941; Rizzolatti et al., 1981). Many decades after it was originally conceived of and named, the PPS is still a hot topic in various research areas and has gained increasing attention, especially with systematic reviews in the last 10 years (Cléry et al., 2015; de Vignemont and Iannetti, 2015; Van der Stoep et al., 2015; Hunley and Lourenco, 2018). Two achievements of these reviews are that they have categorized the PPS as having not only a defensive but also a non-defensive function and evaluated the effect of multisensory stimulus processing in the PPS.

The defensive function has been a classical explanation of the importance of the PPS. The approach of a dangerous object (e.g., a knife) could hurt the body; then, it is necessary to predict physical contact before the object touches us and to decide whether to defend ourselves or flee. Therefore, attention to objects in the PPS is important for physical defense, and the PPS is interpreted as a function safety margin (e.g., Cooke and Graziano, 2003; Graziano and Cooke, 2006; Sambo and Iannetti, 2013). In parallel with this interpretation, it is known that even non-dangerous and non-affective stimuli attract attention within the PPS (e.g., Graziano and Gross, 1995; Reed et al., 2006). This non-defensive function has been interpreted as functioning to focus attention on useful objects (e.g., food and tools) within close proximity (e.g., Rizzolatti et al., 1997; Brozzoli et al., 2014). Recent studies have assumed that defense is not the only important function of the PSS but also that non-defense functions and the interaction of each function also matter (e.g., de Vignemont and Iannetti, 2015). The defensive function is important for self-protection, and this is interpreted by an automatic process by which attention is attracted to a specific threat stimulus, whereas the non-defensive function is the active process of allocating attention to useful information around us and preparing for present and future benefits, e.g., perceiving reward prospects in one's surroundings and estimating reward values in changing social contexts (Coello et al., 2018; Gigliotti et al., 2021, 2023). In addition, the benefits from the non-defensive function could be leveraged when the defense is needed (e.g., we can eat the food we get and also throw it at menacing animals as a decoy). Therefore, the importance of the non-defensive function is still drawing attention and discussion in PSS studies on how non-affective information attracts our attention.

When objects are present in the PPS, they are often in the process of approaching the body. If we can see or hear them approaching, this visual and auditory information in the PPS is processed at the bimodal or trimodal neurons, and processing elicits larger neural activity than information presented outside of the PPS (e.g., Rizzolatti et al., 1981; Graziano et al., 1999). Moreover, the combination of these visual (or auditory) stimuli within the PPS and a tactile stimulus facilitates a response to the tactile stimulus (e.g., Gray and Tan, 2002; Makin et al., 2007; Canzoneri et al., 2012). Taking together the approach of an object to the PPS and a multisensory interaction, it is possible that this interaction could be used to predict subsequent contact (tactile event) since the approaching object might subsequently contact the body. It is necessary to keep in mind that such visuo–tactile interaction in the PPS could vary flexibly, e.g., the facilitation effect of multisensory integration in the PPS could be extended to a wider space depending on the meaning of the visual stimuli and the social context (Geers and Coello, 2023). Therefore, the relationship between multisensory integration and interactions with the environment should be considered for each situation. In this study, we focused only on the PPS effect of the approach of non-affective and task-irrelevant visual stimuli and a subsequent tactile stimulus.

Recent studies reported that the approach of non-affective and task-irrelevant visual stimuli facilitates the prediction of a subsequent tactile stimulus. In research by Kimura (2021), visual stimuli were presented, followed by a tactile stimulus to the hand, and these visual stimuli were presented either sequentially approaching the hand or at a fixed distance from the hand (not near each hand). To confirm the event-related desynchronization (ERD) for a subsequent tactile event, event-related spectral perturbation (ERSP) of the electroencephalogram (EEG) was examined. The results showed that ERD in the beta band occurred strongly at the electrodes near the primary somatosensory area, which processes the tactile stimulus when the visual stimuli approached the hand (tactile stimulus presentation site). ERD is one of the analytical indicators of EEG (e.g., Pfurtscheller and Lopes da Silva, 1999; Engel et al., 2001), and ERD of the beta-band activity in the primary somatosensory area reflects the intensity of prediction of subsequent tactile events (e.g., van Ede et al., 2010). Therefore, this study shows that the approach of visual stimuli within the PPS facilitates the spatial prediction of a subsequent tactile stimulus. Moreover, in previous studies examining ERPs, when a tactile stimulus was presented to the opposite hand from the one approached, the amplitude of event-related brain potentials (ERPs) reflecting prediction error increased (Kimura and Katayama, 2015). This phenomenon occurs not only in spatial prediction but also in temporal prediction (Kimura and Katayama, 2017) and in response to the type of stimulus (Kimura and Katayama, 2018).

These results indicate that the approach of visual stimuli in the PPS facilitates the spatial prediction of a subsequent tactile event. However, these studies did not examine the effect of the presentation of visual stimuli immediately near the hand. Therefore, it is unclear whether the approach of stimuli is important for the prediction of subsequent tactile events or if being located within the PPS (i.e., the presentation of visual stimuli near the hand) is sufficient. Considering the safety margin, it is possible to predict subsequent contact at an early stage, before contact, by using the approach of visual stimuli. Furthermore, the facilitation effect of bimodal sensory stimulation on multisensory interactions varies flexibly and could occur even in spaces away from the body; thus, it might be possible to predict the contact of a tactile stimulus from the beginning of the approach of the visual stimuli (e.g., Serino et al., 2015a,b; Bertoni et al., 2021; Geers and Coello, 2023). Considering finite cognitive resources, it is possible to allocate attentional resources just before contact (i.e., near the hand) to predict a tactile stimulus. The present study aimed to scrutinize the predictive function of visuo–tactile interaction in the PPS by examining the effects of visual stimuli approaching the hand and of visual stimuli simply existing near the hand.

We focused on ERDs in the beta-band activity as an index of the spatial prediction caused by the approach of visual stimuli and P3 in ERP as an index of attention to a subsequent tactile stimulus and the prediction error between spatial prediction from this approach and the deviation of the tactile stimulus from this prediction. Even non-dangerous and non-affective stimuli attract attention within the PPS (e.g., Graziano and Gross, 1995; Reed et al., 2006). In addition, the sequential approach of visual stimuli is processed as useful information for the spatial prediction of a subsequent tactile stimulus (e.g., Kimura and Katayama, 2015; Kimura, 2021), and it is possible that this facilitation effect might not be generated by visual stimuli simply existing in close proximity to the body (e.g., Serino et al., 2015a,b; Bertoni et al., 2021; Geers and Coello, 2023). Therefore, we hypothesized that sequentially approaching visual stimuli facilitates the spatial prediction of a subsequent tactile stimulus due to a non-defensive function of the PSS to direct attention to useful information around us. Based on this hypothesis, we predicted that ERDs in the beta band would occur in each condition before the presentation of the tactile stimulus if participants can predict the tactile stimulus (e.g., van Ede et al., 2010) and that the ERD in the approach condition would be larger than that in the near condition if the approach of visual stimuli better facilitates prediction compared with the nearness of the presentation (e.g., Kimura, 2021). P3 reflects attention to the target stimulus (e.g., Duncan-Johnson and Donchin, 1977; Donchin, 1981; Katayama and Polich, 1996a), and its amplitude increases in response to the deviation of a tactile stimulus from spatial prediction due to the approach of visual stimuli (e.g., Kimura and Katayama, 2015). In the present study, participants were asked to perform a simple reaction time task in response to a tactile stimulus, in which it was necessary to allocate attention rapidly to the tactile stimulus. We predicted that the P3 amplitude elicited by the high-probability stimulus in the approach condition would be larger than that elicited by the high-probability stimulus in the near condition if the approach of a visual stimulus induces attention to the direction from which a tactile stimulus approaches, making it easier to predict and quickly allocate attention to the tactile stimulus. In addition, we predicted that the P3 amplitude elicited by the low-probability stimulus in the approach condition would be larger than that elicited by the low-probability stimulus in the near condition if the approach of visual stimuli better facilitates the prediction of a tactile stimulus compared with the near presentation of visual stimuli. In addition, we examined contingent negative variation (CNV; Walter et al., 1964) in ERP before the presentation of tactile stimuli to ensure that the temporal prediction of tactile stimuli did not differ between conditions. CNV reflects the temporal prediction of a subsequent stimulus, including tactile stimuli (Kimura and Katayama, 2015; Kimura, 2021). We predicted that the amplitude of CNV would not differ between conditions if participants could predict the timing of the presentation of the tactile stimulus in both conditions, as was found in previous studies (Kimura and Katayama, 2015; Kimura, 2021).

2. Materials and methods

2.1. Participants

Twelve undergraduate and graduate students (nine female students and three male students; 18–24 years of age) participated in the experiment, and all were newly recruited for this experiment. All participants were right-handed, had a normal or corrected-to-normal vision, and had tactile sensitivity without hindrance to perform the task, according to their self-report. This experiment was approved by the Kwansei Gakuin University (KGU) Research Ethics Review Board under the KGU Regulations for Research with Human Participants. Written informed consent was obtained from all participants, and their rights as experimental subjects were protected. In this study, this sample size was determined with reference to Kimura and Katayama (2015) by a similar experimental paradigm rather than by prior calculation. Therefore, we conducted a *post-hoc* power analysis to consider the power of the result. The most important analysis of this study was examining the influence of visual stimuli approaching the hand and visual stimuli close to the hand on the prediction of a subsequent tactile stimulus. The analysis of ERD corresponded to this predictive effect; thus, in our results section, we report a *post-hoc* power analysis of the results of ERD analysis.

2.2. Stimulus and equipment

2.2.1. Visual and tactile stimuli

Figure 1 shows the positioning of the visual and tactile stimuli. The stimuli were set according to the previous studies (Kimura and Katayama, 2015, 2017, 2018). Participants were seated and put their hands and forearms on a desk in front of them. Their hands were 32.0 cm apart. Visual stimuli were presented by three white light-emitting diodes (LEDs; square with 0.8 cm sides). These were placed between the arms on the desk (at equal distances of 8.0 cm intervals). The intensity and duration of visual stimuli were 25 cd and 200 ms. Somatosensory stimuli were presented by an electrical stimulus generator (Nihon Kohden Corporation, SEN-7203, Japan), electric isolators (Nihon Kohden Corporation, SS-203J, Japan), and Ag/AgCl electrodes (diameter of 1.0 cm) on participants' forearms. The anode electrode was placed on the participants' wrists, and the cathode electrode was 3.0 cm from the anode toward the elbow. The electrical stimulus was a single block pulse with a 0.2-ms duration. The intensities of stimuli were three times as high as the sensory threshold for each participant. This intensity was also used in a previous study of ERP elicited by tactile stimuli (e.g., Kimura and Katayama, 2015), and it never caused pain. The average intensity of the stimuli across all participants was 3.3 mA. The presentation of visual and tactile stimuli was controlled with MATLAB R2010b (MathWorks, Inc.) and Psychtoolbox (Kleiner et al., 2007) installed on a desktop computer (Precision T5500, DELL).

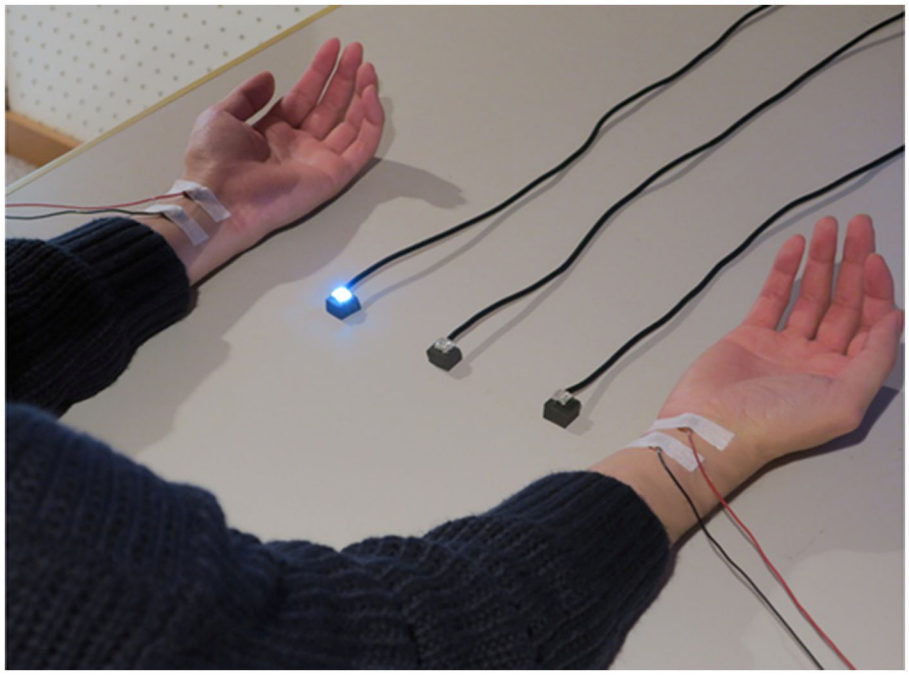


FIGURE 1
The positions of visual stimuli and tactile stimuli.

2.2.2. Recording of EEG

EEG data were recorded by NuAmps (Compumedics Neuroscan, USA) and an electrode cap (Easycap GmbH, Germany) using Ag/AgCl electrodes at 30 sites (Fp1, Fp2, F7, F3, Fz, F4, F8, FT7, FC3, FCz, FC4, FT8, T7, C3, Cz, C4, T8, TP7, CP3, CPz, CP4, TP8, P7, P3, Pz, P4, P8, O1, Oz, and O2) according to the modified 10–20 System. In addition, electrodes were also placed on both earlobes (A1 and A2). The reference electrode was on the tip of the nose, and the ground electrode site was AFz. The data from all channels were recorded using SCAN software (Compumedics Neuroscan, USA). The electrode impedances were kept below 5 kΩ. A bandpass filter of 0.1–200 Hz was used for recording. The sampling rate was 1000 Hz.

2.2.3. Procedure

Figure 2 shows the experimental procedure. The procedure was set according to the previous studies (Kimura and Katayama, 2015; Kimura, 2021). Each trial was composed of three visual stimuli and one tactile stimulus. The stimulus onset asynchrony (SOA) was set to 1000 ms. The interval between trials was either 1000 or 1200 ms at random with equal probability. Each block was composed of 84 trials [high-probability tactile stimuli: 64 trials; low-probability tactile stimuli: 16 trials; no tactile stimuli (catch trial): 4 trials], which took 7 min. Two blocks were presented for each condition (overall: four blocks per participant). The interval between blocks was 2 min, and after the second block, the participants rested for 10 min and then started the remaining two blocks. The order of conditions was randomized between participants.

The difference between the two conditions was the pattern of visual stimuli, and these patterns were presented in separate blocks. In the approach condition, LEDs flashed sequentially toward the hand where the high-probability tactile stimulus was presented (i.e., if the high-probability tactile stimulus was set at the left wrist, the LEDs flashed sequentially right, center, and left), and the subsequent tactile stimulus was presented to the left (or right) wrist with a high probability (80%) and to the opposite wrist with a low probability (20%). In the near condition, the LED flashed three times near the hand where the high-probability tactile stimulus was presented (i.e., if the high-probability tactile stimulus was set at the left wrist, the left LEDs flashed three times), and the subsequent tactile stimulus was presented to the left (or right) wrist. The participants were required to gaze at the center LED in order to control their eye movements and not to move their eyes and bodies more than necessary in each condition. Moreover, the participants were instructed to respond by pressing a button with the left (or right) foot whenever the tactile stimuli were presented and to not respond when tactile stimuli were not presented (i.e., the catch trials). Half of the participants used the left foot, and the other half used the right foot. Finally, they were told at the start of each block which hand would be presented with the high-/low-probability stimuli.

2.2.4. Data analysis

The mean reaction times (RTs) for the tactile stimuli were calculated. Based on previous studies, trials with an incorrect response or with RTs shorter than 200 ms or longer than 1500 ms were discarded from analysis (Kimura and Katayama, 2015, 2017, 2018). After the rejection, a two-way repeated-measures ANOVA on RTs was conducted with the two conditions (approach

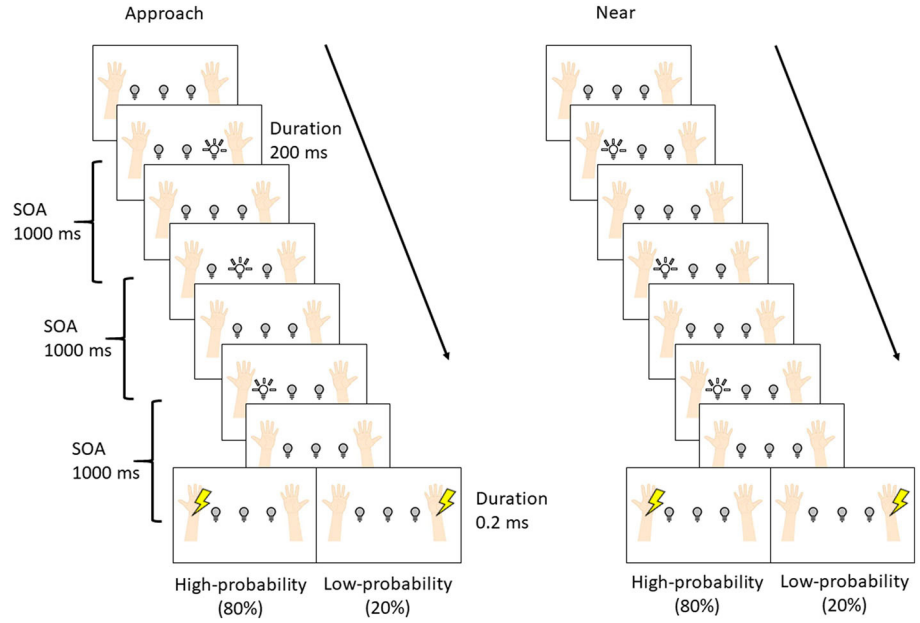


FIGURE 2
The procedure of the experiment.

and near condition) and two probabilities (high probability and low probability).

The EEG data were analyzed based on the method of Kimura (2021). The EEGLAB toolbox (Delorme and Makeig, 2004) and ERPLAB toolbox (Lopez-Calderon and Luck, 2014) on MATLAB were used for this analysis. Artifacts derived from eye movements and eye blinks were rejected using an automatic EEG artifact detector based on the joint use of spatial and temporal features (ADJUST) of the EEGLAB toolbox (Mognon et al., 2011). In the time-frequency analysis, the EEG data epoch was 1800 ms (including a 900 ms prestimulus of the tactile stimulus). Epochs in which the EEG signal variation exceeded ±100 μV were rejected. After artifact rejection, EEG data were transformed by the Morlet wavelet transformation function applied in a Hanning-tapered window in EEGLAB. The settings were as follows: epoch time limits: −900 to 900 ms, using 400-time points; frequency limits: 8–30 Hz; baseline limits: −900 to −500 ms; wavelet cycles: 3–0.5. The processed data were output from −691.88 to 690.88 ms (400 time points) and from 8 to 30 Hz (22 frequency points). The beta-band (14.29–30 Hz) ERSPs for the time range −300 to 0 ms at the electrodes of C3 and C4 (i.e., the neighboring electrodes for the primary somatosensory area) were averaged in each block. In addition, these electrodes were distinguished by the prediction of a tactile stimulus. C3 (C4) is ipsilateral and C4 (C3) is contralateral when the block with the high-probability tactile stimulus is presented to the left (right) hand. The averaged beta-band ERSP for ipsilateral and contralateral was calculated in each condition. After this processing, the numbers of the remaining trials were 146–160 (0–8.65% rejected) for the approach condition and 149–160 (0–6.87% rejected) for the near condition. To check the ERD, one-sample *t*-tests of beta-band ERSPs were conducted with all combinations between conditions and lateralities (ipsilateral and contralateral). If ERD occurred in all combinations, two-way repeated-measures ANOVAs of ERSPs were conducted with the two conditions and two lateralities.

The ERPs were analyzed based on the method of Kimura and Katayama (2015). The data were digitally low-pass filtered at 30 Hz (6 dB/octave) using an IIR Butterworth analog simulation filter. Artifacts derived from eye movements and eye blinks were rejected using ADJUST. To extract P3, the EEG epoch was set at 1000 ms (including a 200 ms prestimulus of the tactile stimulus). The epoch in which the EEG signal variation exceeded ± 100 μV was excluded from averaging. After artifact rejection, the numbers of remaining trials ranged from 118 to 128 (0–7.81% of trials were rejected) for the high-probability stimuli and 28–32 (0–12.5% rejected) for the low-probability stimuli in the approach condition, and 121–128 (0–5.47% rejected) for the high-probability stimuli and 28–32 (0–12.5% rejected) for the low-probability stimuli in the near condition.

To analyze the P3, the time range of P3 was defined to be 230–290 ms for high-probability stimuli and 260–350 ms for low-probability stimuli. These time ranges were decided by peak latencies of the grand averaged waves for each probability used in the analysis. The mean P3 amplitudes at Pz, where the P3 was elicited at maximum amplitude, were analyzed. Two-way repeated-measures ANOVAs of P3 were conducted with the two conditions and two probabilities. In addition, to investigate CNV, the EEG epoch was set at 1400 ms (the baseline was a – 200 to 0 ms prestimulus of the third visual stimulus). The signal processing and rejection were the same as in the method for P3. After artifact rejection, the numbers of remaining trials were 146–160 (0–8.65% rejected) for the approach condition and 142–160 (0–11.25% rejected) for the near condition. The mean CNV amplitude was obtained from a latency window of 500–1000 ms. The appropriate

latency window was defined based on the observation of the resultant ERP waveforms. The mean CNV amplitudes at Cz, where the CNV was elicited at maximum amplitude, were compared between conditions by a paired *t*-test.

The normality of the data for ERD, P3, CNV, and RT was checked by the Shapiro–Wilk test using the R function shapiro.test. These ANOVAs were conducted by applying Greenhouse–Geisser corrections to the degrees of freedom when appropriate (Greenhouse and Geisser, 1959). *Post-hoc* comparisons were made using Shaffer's modified sequentially rejective multiple-test procedure, which extends the Bonferroni *t*-tests in a stepwise fashion (Shaffer, 1986). The effect sizes for ANOVAs were indicated in terms of partial eta squared (η_p^2), and *t*-tests were calculated by computing Cohen's *d*. The significance level was set at a *p*-value of < 0.05 for all statistical analyses.

3. Results

3.1. Normality test for each data

To confirm the normality of values of ERD, P3, CNV, and RT, Shapiro–Wilk tests were conducted using the R function shapiro.test. The results revealed that these distributions of the data did not significantly differ from a normal distribution ($ps > 0.05$).

3.2. *Post-hoc* power analysis

To confirm the statistical power of this study, a *post-hoc* power analysis was conducted using G*Power 3.1.9.4 (Faul et al., 2007). The analysis revealed that the statistical power for the result of ERD was 1-β = 0.99 [12 participants, $\alpha = 0.05$, $\eta_p^2 = 0.38$ ($f = 0.78$)].

3.3. Reaction times

Figure 3 shows the mean RTs for the tactile stimuli for each condition. Averaged RTs of all participants were 310 ms (SE = 15.33), 361 ms (SE = 20.44), 337 ms (SE = 18.00), and 372 ms (SE = 23.15) for the approach-high-probability, approach-low-probability, near-high-probability, and near-low-probability stimuli. The results of the ANOVA revealed that the main effect of conditions [$F_{(1,11)} = 8.72$, $p = 0.013$, $\eta_p^2 = 0.44$] and probabilities [$F_{(1,11)} = 27.42$, $p < 0.001$, $\eta_p^2 = 0.71$] were significant. In addition, the interaction of conditions and probabilities was significant [$F_{(1,11)} = 12.84$, $p = 0.004$, $\eta_p^2 = 0.54$]. *Post-hoc* comparisons indicated that the RT in the approach condition was shorter than in the near condition for high-probability stimuli ($p = 0.002$), and low-probability stimuli did not show a significantly different condition effect ($p = 0.171$). In addition, RT to high-probability stimuli was shorter than to low-probability stimuli in each condition ($ps < 0.001$).

3.4. Event-related spectral perturbations

Figure 4A shows the ERSPs in each condition and each laterality and Figure 4B the averaged beta-band ERSPs at the time range of −300 to 0 ms in all conditions and lateralities. The results of the one-sample *t*-test revealed that the beta-band ERSPs were smaller than zero in all conditions and lateralities [$ts(11) > 3.30$, $ps < 0.07$, $ds > 1.14$]; therefore, ERD occurred in all conditions and lateralities. The results of the ANOVA revealed that the main effect of the condition was significant [$F_{(1,11)} = 6.83$, $p = 0.024$, $\eta_p^2 = 0.38$] and that the ERD of the approach condition was larger than that of the near condition. The main effect of laterality [$F_{(1,11)} = 1.21$, $p = 0.295$, $\eta_p^2 = 0.20$] and the interaction [$F_{(1,11)} = 0.23$, $p = 0.618$, $\eta_p^2 = 0.02$] was not significant.

3.5. Event-related brain potentials

Figure 5A shows the grand averages for ERPs elicited by tactile stimuli in each condition and probability from Pz. The positive deflection in the high-probability stimuli showed peak latency at approximately 260 ms and the positive deflection in the low-probability stimuli showed peak latency at approximately 320 ms. Figure 5B shows the topographic map at the time range and mean amplitude of P3. The ANOVA for the mean amplitude of P3 revealed a significant main effect of probabilities [$F_{(1,11)} = 5.18$, $p = 0.044$, $\eta_p^2 = 0.32$]; the P3 mean amplitude elicited by the low-probability stimuli was larger than that elicited by the high-probability stimuli. The main effect of conditions was not significant [$F_{(1,11)} = 3.99$, $p = 0.071$, $\eta_p^2 = 0.26$]. In addition, the interaction of conditions and probabilities was significant [$F_{(1,11)} = 8.73$, $p = 0.013$, $\eta_p^2 = 0.44$]. *Post-hoc* comparisons indicated that the P3 mean amplitude in the approach condition was larger than in the near condition for high-probability stimuli ($p < 0.001$), and low-probability stimuli did not show a significantly different condition effect ($p = 0.992$). In addition, the probability effect of the P3 mean amplitude shown in the near condition for low-probability stimuli was larger than that elicited by the high-probability stimuli ($p = 0.005$), and this effect did not show in the approach condition ($p = 0.321$).

Figure 6 shows the grand average CNV elicited in all trials at Cz, where the CNV was elicited at maximum amplitude and the topographic map at the time range of CNV (5000–1000 ms). The results of the paired *t*-test revealed no significant difference between conditions [$t(11) = 0.32$, $p = 0.754$, $d = 0.12$].

4. Discussion

This study aimed to scrutinize the predictive function of visuo–tactile interaction in the PSS by examining the effect of visual stimuli approaching the hand and visual stimuli close to the hand. For this purpose, ERDs, ERPs, and RTs were compared between the approach condition and the near condition.

RTs to high-probability stimuli were shorter than those to low-probability stimuli in each condition, and RTs to high-probability stimuli in the approach condition were shorter than those to the high-probability stimuli in the near condition. In addition, the amplitude of CNV did not differ between the conditions. These results suggest that the participants could predict the timing of the presentation of the tactile stimulus in both conditions and that the spatial prediction of the high-probability tactile stimulus is facilitated in the approach condition.

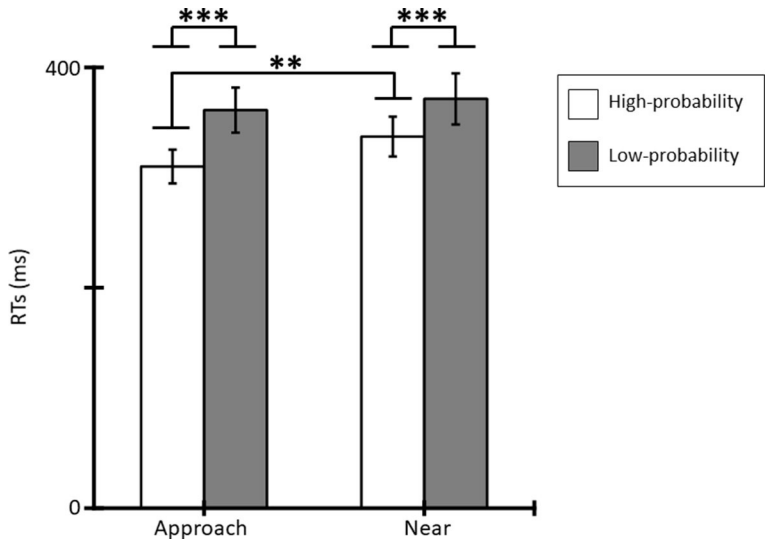


FIGURE 3
Mean RTs (ms) for tactile stimuli and standard errors (SE) of RTs in each condition (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$).

The beta-band ERSPs were suppressed approximately 300 ms before the presentation of the tactile stimulus in all conditions and lateralities. More importantly, the beta-band ERD of the approach condition was larger than that of the near condition. The beta-band ERD before the presentation of a tactile stimulus reflects the intensity of the prediction of a subsequent tactile stimulus (van Ede et al., 2010), and the approach of visual stimuli facilitates this prediction (Kimura, 2021). The only difference between conditions was the method of presentation of visual stimuli. Therefore, this result suggests that the approach of visual stimuli is more important for the prediction of a subsequent tactile stimulus compared with the presentation of visual stimuli near the hand. The tactile stimulus in this study did not cause pain, and the visual stimuli were merely flashes of LED lights; thus, it is possible that these results were caused by a non-defensive function of the PPS. Visual stimuli approaching the body offer useful information for the spatial prediction of subsequent tactile stimuli (e.g., Kimura and Katayama, 2015; Kimura, 2021), and the non-defensive function of the PPS allocate more attention to useful information in the PPS (e.g., Rizzolatti et al., 1997; Brozzoli et al., 2014). Moreover, previous studies reported that an object approaching the PPS activates bimodal and trimodal neurons and then promotes stimulus processing (e.g., Rizzolatti et al., 1981; Graziano et al., 1999). The distance at which these neuronal activities (and the resulting facilitation effect of stimulus processing within the PPS) occur varies flexibly (e.g., Serino et al., 2015a,b; Bertoni et al., 2021; Geers and Coello, 2023). Taken together, the results of this study suggest that visual stimuli approaching the hand might increase the activity of these neurons, facilitating the spatial prediction of a subsequent tactile stimulus within the PPS, even when the stimuli are non-threatening, while the mere presence of visual stimuli in close proximity to the hand might not.

The P3 amplitude elicited by the high-probability stimuli in the approach condition was larger than that in the near condition. P3 reflects attention to a stimulus, and this attention is allocated based on the importance and meaning of the stimulus (e.g., Duncan-Johnson and Donchin, 1977; Donchin, 1981; Katayama and Polich, 1996b). In this study, participants performed a simple reaction time task in response to a tactile stimulus and were required to allocate attention quickly to the stimuli presented. Therefore, it is considered that approaching visual stimuli attracts attention to the space that is approached and that this attention is allocated to a high-probability tactile stimulus that is presented there. As a result, P3 amplitude was increased for high-probability stimuli in the approach condition. Moreover, the P3 amplitude elicited by the low-probability stimuli did not differ between conditions. In the near condition, this amplitude was larger than that elicited by the high-probability stimuli. The results of RT, ERD, and P3 elicited by the high-probability stimuli suggest that the approach condition attracted attention to the presentation location of the high-probability tactile stimulus and facilitated the prediction of the stimulus, whereas the degree of spatial deviance of the low-probability tactile stimulus did not differ between conditions because the distance from one wrist to the other wrist at which the high-probability stimulus was presented is comparable. Therefore, it is possible that spatial prediction errors did not differ between conditions. Moreover, this result may also have been caused by the ceiling effect. In previous studies using similar conditions as the approach condition, the P3 amplitude elicited by low-probability stimuli was comparable to the results in this study (e.g., Kimura and Katayama, 2015). Therefore, the P3 amplitude elicited by low-probability stimuli may not increase further in experiments using this paradigm.

In summary, the present study indicates that the approach of visual stimuli facilitates the spatial prediction and processing of a subsequent tactile stimulus compared to situations in which visual stimuli only exist, without moving, within the PPS. In this study, visual stimuli were task-irrelevant, and the subsequent tactile stimulus also produced no pain; thus, the approach of the visual stimuli within the PPS is considered to relate to an automatic predictive function that facilitates the prediction of subsequent tactile events, even if the visual stimuli are non-affective.

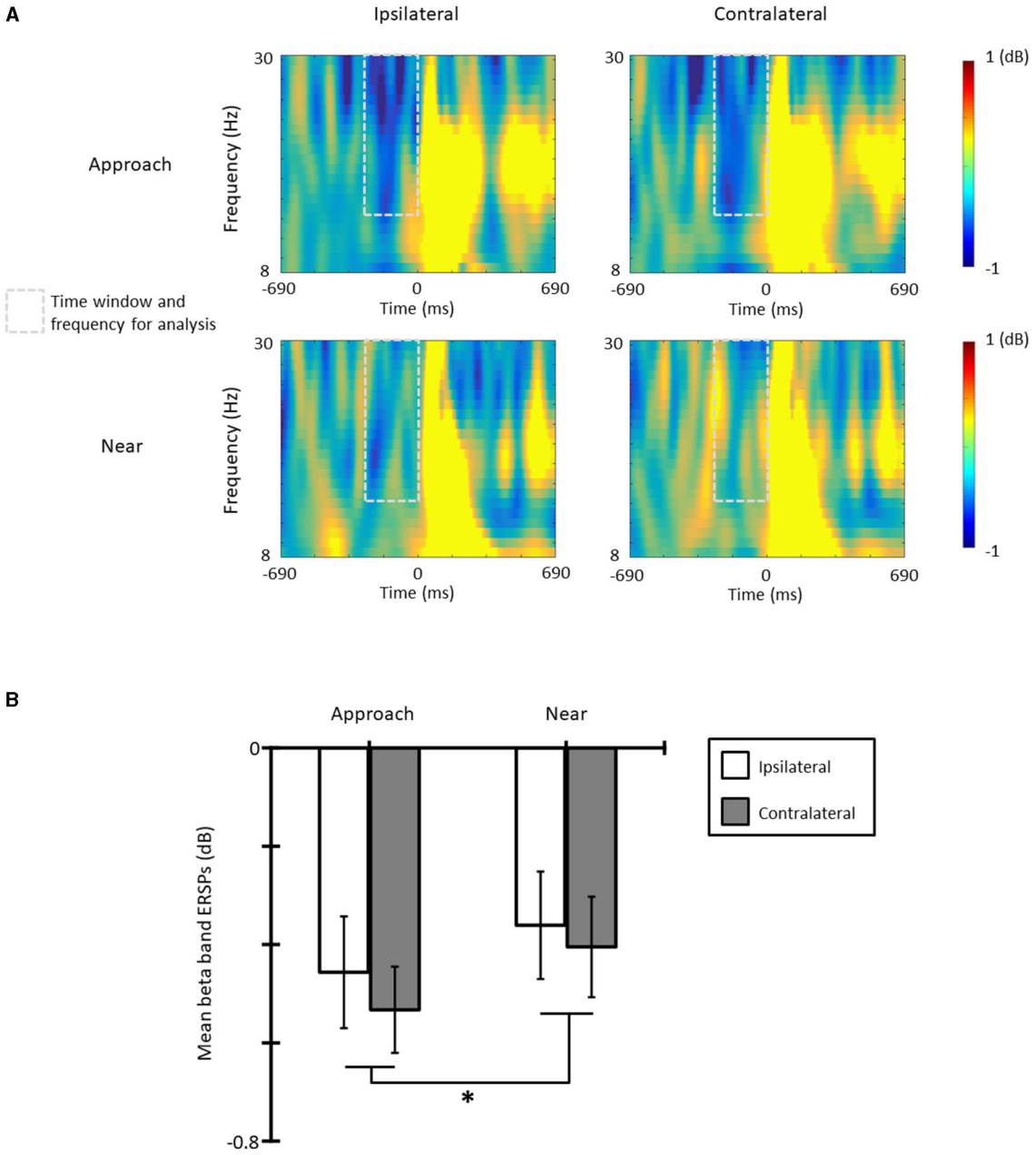


FIGURE 4
(A) The beta band event-related spectral perturbations (ERSPs) in each condition and laterality, and **(B)** the mean beta band ERSPs at the time range of −300 to 0 ms (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$). The error bars indicate the standard errors (SE).

Finally, it is necessary to consider, in a future study, whether this effect influences other predictions for visuo-tactile processing. The approach of visual stimuli facilitates not only spatial prediction but also temporal prediction manipulated by the presentation timing of a tactile stimulus (Kimura and Katayama, 2017) and the prediction of stimulus type manipulated by tactile stimulus features (Kimura and Katayama, 2018). However, it is unclear whether this effect is caused by the approach of visual stimuli or the presentation of visual stimuli within the PPS. Therefore, whether this effect affects the overall prediction for visuo-tactile processing should be investigated in future research.

5. Conclusion

The results of this study revealed that the approach of visual stimuli better facilitates spatial prediction and processing of a subsequent tactile stimulus compared to situations in which visual stimuli just exist within the PPS. This study extended our understanding of attentional processing within the PPS and indicated that visual stimuli within the PPS might be related to an automatic predictive function that facilitates the prediction of subsequent tactile events, rather than only having a defensive function.

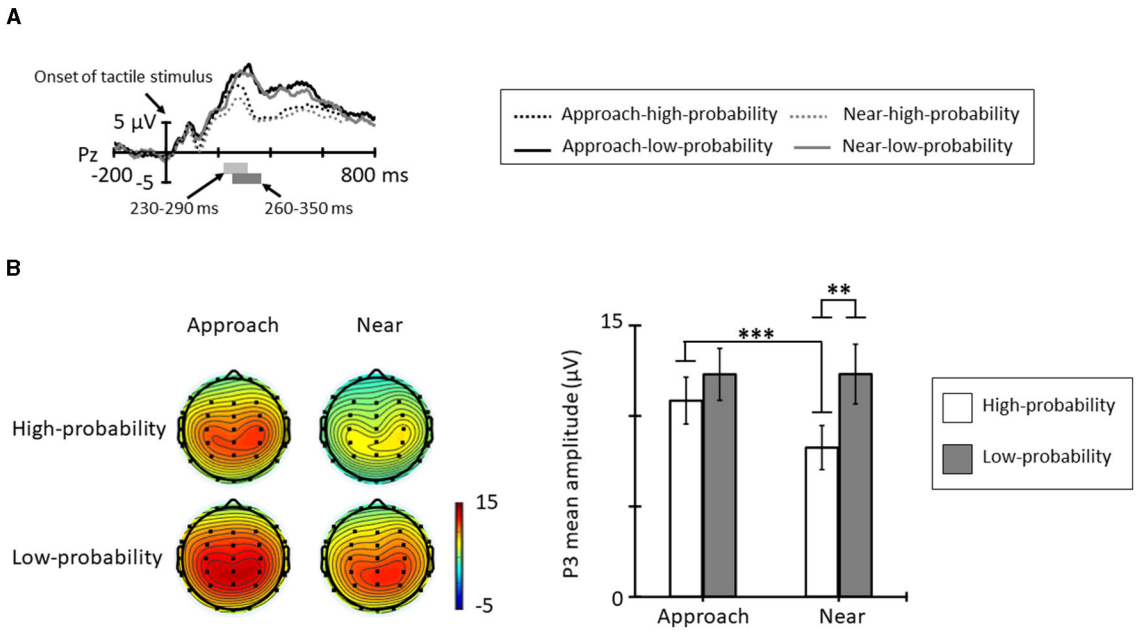


FIGURE 5
(A) Grand average ERP waveforms elicited by tactile stimuli for each condition and probability at Pz. The light gray area denotes the time range of P3 for high-probability stimuli (230–290 ms), and the dark gray area denotes the time range of P3 for low-probability stimuli (260–350 ms). **(B)** The topographic maps and mean amplitudes of P3 for these time ranges (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$). The error bars indicate the standard errors (SE).

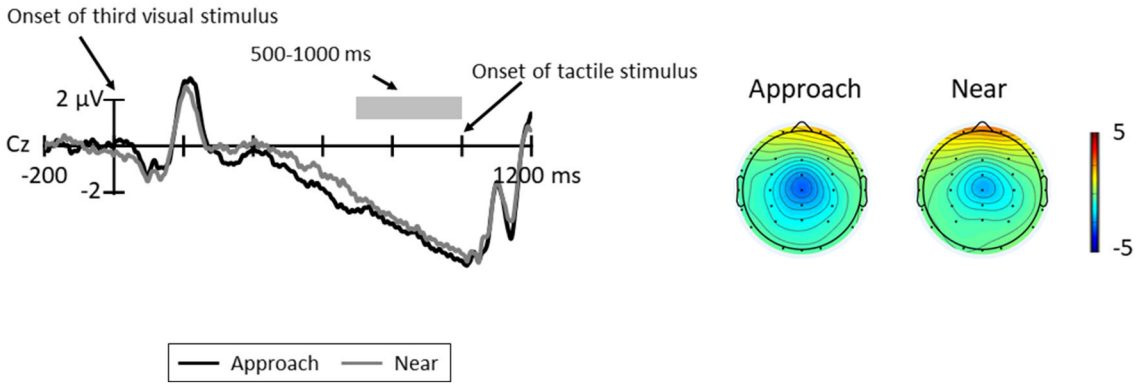


FIGURE 6
Grand average ERP waveforms between the third visual stimuli and tactile stimuli for each condition at Cz; the gray area indicates the time range for CNV (500–1000 ms). The topographic maps indicate the mean amplitudes of CNV for these time ranges.

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 study involving human participants was reviewed and approved by the Kwansei Gakuin University (KGU) Research Ethics Review Board under the KGU Regulations for Research with Human Participants. The patients/participants provided their written informed consent to participate in this study.

Author contributions

TK and JK contributed to the conception and design of the study. TK contributed to the data acquisition

and wrote the first draft of the manuscript. Both authors contributed to the statistical analysis, revising and reading the manuscript, and finally approved the submitted version.

Funding

TK declares that this study received funding from a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (no. 21K13751). The funding source was not involved in the study design; collection, analysis, and interpretation of data; writing of the article; or the decision to submit the article for publication.

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

References

Bertoni, T., Magosso, E., and Serino, A. (2021). From statistical regularities in multisensory inputs to peripersonal space representation and body ownership: insights from a neural network model. *Eur. J. Neurosci.* 53, 611–636. doi: 10.1111/ejn.14981

Brain, W. R. (1941). Visual disorientation with special reference to lesions of the right cerebral hemisphere. *Brain J. Neurol.* 64, 244–272. doi: 10.1093/brain/64.4.244

Brozzoli, C., Ehrsson, H. H., and Farnè, A. (2014). Multisensory representation of the space near the hand: from perception to action and interindividual interactions. *The Neurosci.* 20, 122–135. doi: 10.1177/1073858413511153

Canzoneri, E., Magosso, E., and Serino, A. (2012). Dynamic sounds capture the boundaries of peripersonal space representation in humans. *PloS ONE* 7, e44306. doi: 10.1371/journal.pone.0044306

Cléry, J., Guipponi, O., Wardak, C., and Ben Hamed, S. (2015). Neuronal bases of peripersonal and extrapersonal spaces, their plasticity and their dynamics: knowns and unknowns. *Neuropsychologia* 70, 313–326. doi: 10.1016/j.neuropsychologia.2014.10.022

Coello, Y., Quesque, F., Gigliotti, M. F., Ott, L., and Bruyelle, J. L. (2018). Idiosyncratic representation of peripersonal space depends on the success of one's own motor actions, but also the successful actions of others!. *PloS ONE* 13, e0196874. doi: 10.1371/journal.pone.0196874

Cooke, D. F., and Graziano, M. S. (2003). Defensive movements evoked by air puff in monkeys. *J. Neurophysiol.* 90, 3317–3329. doi: 10.1152/jn.00513.2003

de Vignemont, D., and Iannetti, F. (2015). How many peripersonal spaces?. *Neuropsychologia* 70, 327–334. doi: 10.1016/j.neuropsychologia.2014.11.018

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Method.* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009

Donchin, E. (1981). Surprise! … surprise? *Psychophysiology* 18, 493–513. doi: 10.1111/j.1469-8986.1981.tb01815.x

Duncan-Johnson, C. C., and Donchin, E. (1977). On quantifying surprise: the variation of event-related potentials with subjective probability. *Psychophysiology* 14, 456–467. doi: 10.1111/j.1469-8986.1977.tb01312.x

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

Faul, F., Erdfelder, E., Lang, A. G., and Buchner, A. (2007). G* Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/BF03193146

Geers, L., and Coello, Y. (2023). The relationship between action, social and multisensory spaces. *Sci. Rep.* 13, 202. doi: 10.1038/s41598-023-27514-6

Gigliotti, M. F., Bartolo, A., and Coello, Y. (2023). Paying attention to the outcome of others' actions has dissociated effects on observer's peripersonal space representation and exploitation. *Sci. Rep.* 13, 10178. doi: 10.1038/s41598-023-37189-8

Gigliotti, M. F., Soares Coelho, P., Coutinho, J., and Coello, Y. (2021). Peripersonal space in social context is modulated by action reward, but differently in males and females. *Psychol. Res.* 85, 181–194. doi: 10.1007/s00426-019-01242-x

Gray, R., and Tan, H. Z. (2002). Dynamic and predictive links between touch and vision. *Exp. Brain Res.* 145, 50–55. doi: 10.1007/s00221-002-1085-x

Graziano, M. S., and Cooke, D. F. (2006). Parieto-frontal interactions, personal space, and defensive behavior. *Neuropsychologia* 44, 845–859. doi: 10.1016/j.neuropsychologia.2005.09.009

Graziano, M. S., Reiss, L. A., and Gross, C. G. (1999). A neuronal representation of the location of nearby sounds. *Nature* 397, 428–430. doi: 10.1038/17115

Graziano, M. S. A., and Gross, C. G. (1995). "The representation of extrapersonal space: a possible role for bimodal, visual-tactile neurons," in *The Cognitive Neurosciences*, ed M. S. Gazzaniga (London: The MIT Press), 1021–1034.

Greenhouse, S. W., and Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika* 24, 95–112. doi: 10.1007/BF02289823

Hunley, S. B., and Lourenco, S. F. (2018). What is peripersonal space? An examination of unresolved empirical issues and emerging findings. Wiley interdisciplinary reviews. *Cognit. Sci.* 9, e1472. doi: 10.1002/wcs.1472

Katayama, J., and Polich, J. (1996a). P300 from one-, two-, and three-stimulus auditory paradigms. *Int. J. Psychophysiol.* 23, 33–40. doi: 10.1016/0167-8760(96)00030-X

Katayama, J., and Polich, J. (1996b). P300, probability, and the three-tone paradigm. *Electroencephal. Clin. Neurophysiol.* 100, 555–562. doi: 10.1016/S0168-5597(96)95171-0

Kimura, T. (2021). Approach of visual stimuli facilitates the prediction of tactile events and suppresses beta band oscillations around the primary somatosensory area. *Neuroreport* 32, 631–635. doi: 10.1097/WNR.0000000000001643

Kimura, T., and Katayama, J. (2015). Approach of visual stimuli modulates spatial expectations for subsequent somatosensory stimuli. *Int. J. Psychophysiol.* 96, 176–182. doi: 10.1016/j.ijpsycho.2015.04.002

Kimura, T., and Katayama, J. (2017). Visual stimuli approaching toward the body influence temporal expectations about subsequent somatosensory stimuli. *Brain Res.* 1664, 95–101. doi: 10.1016/j.brainres.2017.03.030

Kimura, T., and Katayama, J. (2018). The approach of visual stimuli influences expectations about stimulus types for subsequent somatosensory stimuli. *Exp. Brain Res.* 236, 1563–1571. doi: 10.1007/s00221-018-5244-0

Kleiner, M., Brainard, D., and Pelli, D. (2007). What's new in psychtoolbox-3? *Perception* 36, 1. doi: 10.1177/03010066070360S101

Lopez-Calderon, J., and Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Front. Hum. Neurosci.* 8, 213. doi: 10.3389/fnhum.2014.00213

Makin, T. R., Holmes, N. P., and Zohary, E. (2007). Is that near my hand? Multisensory representation of peripersonal space in human intraparietal sulcus. *J. Neurosci.* 27, 731–740. doi: 10.1523/JNEUROSCI.3653-06.2007

Mognon, A., Jovicich, J., Bruzzone, L., and Buiatti, M. (2011). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology* 48, 229–240. doi: 10.1111/j.1469-8986.2010.01061.x

Pfurtscheller, G., and Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin. Neurophysiol.* 110, 1842–1857. doi: 10.1016/S1388-2457(99)00141-8

Reed, C. L., Grubb, J. D., and Steele, C. (2006). Hands up: attentional prioritization of space near the hand. *J. Exp. Psychol. Hum. Percep. Perf.* 32, 166–177. doi: 10.1037/0096-1523.32.1.166

Rizzolatti, G., Fadiga, L., Fogassi, L., and Gallese, V. (1997). The space around us. *Science* 277, 190–191. doi: 10.1126/science.277.5323.190

Rizzolatti, G., Scandolara, C., Matelli, M., and Gentilucci, M. (1981). Afferent properties of periarcuate neurons in macaque monkeys. II visual responses. *Behav. Brain Res.* 2, 147–163. doi: 10.1016/0166-4328(81)90053-X

Sambo, C. F., and Iannetti, G. D. (2013). Better safe than sorry? The safety margin surrounding the body is increased by anxiety. *J. Neurosci.* 33, 14225–14230. doi: 10.1523/JNEUROSCI.0706-13.2013

Serino, A., Canzoneri, E., Marzolla, M., di Pellegrino, G., and Magosso, E. (2015a). Extending peripersonal space representation without tool-use: evidence from a combined behavioral-computational approach. *Front. Behav. Neurosci.* 9, 4. doi: 10.3389/fnbeh.2015.00004

Serino, A., Noel, J. P., Galli, G., Canzoneri, E., Marmaroli, P., Lissek, H., et al. (2015b). Body part-centered and full body-centered peripersonal space representations. *Sci. Rep.* 5, 18603. doi: 10.1038/srep18603

Shaffer, J. P. (1986). Modified sequentially rejective multiple test procedures. *J. Am. Stat. Assoc.* 81, 826–831. doi: 10.1080/01621459.1986.10478341

Van der Stoep, N., Nijboer, T. C., Van der Stigchel, S., and Spence, C. (2015). Multisensory interactions in the depth plane in front and rear space: a review. *Neuropsychologia* 70, 335–349. doi: 10.1016/j.neuropsychologia.2014.12.007

van Ede, F., Jensen, O., and Maris, E. (2010). Tactile expectation modulates pre-stimulus beta-band oscillations in human sensorimotor cortex. *NeuroImage* 51, 867–876. doi: 10.1016/j.neuroimage.2010.02.053

Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., and Winter, A. L. (1964). Contingent negative variation: an electric sign of sensorimotor association and expectancy in the human brain. *Nature* 203, 380–384. doi: 10.1038/203380a0

frontiers | Frontiers in Human Neuroscience

TYPE Original Research
PUBLISHED 01 November 2023
DOI 10.3389/fnhum.2023.1252686

OPEN ACCESS

EDITED BY
José Manuel Reales,
National University of Distance Education (UNED), Spain

REVIEWED BY
Chunlin Li,
Capital Medical University, China
Jun'ichi Katayama,
Kwansei Gakuin University, Japan
Kazuhiro Sugawara,
Sapporo Medical University, Japan

*CORRESPONDENCE
Tetsuo Kida
tkida@inst-hsc.jp

RECEIVED 04 July 2023
ACCEPTED 16 October 2023
PUBLISHED 01 November 2023

CITATION
Kida T, Kaneda T and Nishihira Y (2023) ERP evidence of attentional somatosensory processing and stimulus-response coupling under different hand and arm postures. *Front. Hum. Neurosci.* 17:1252686. doi: 10.3389/fnhum.2023.1252686

ERP evidence of attentional somatosensory processing and stimulus-response coupling under different hand and arm postures

Tetsuo Kida[1]*, Takeshi Kaneda[2] and Yoshiaki Nishihira[3]

[1]Higher Brain Function Unit, Department of Functioning and Disability, Institute for Developmental Research, Aichi Developmental Disability Center, Kasugai, Japan, [2]Faculty of Education, Hakuoh University, Oyama, Japan, [3]Graduate School of Comprehensive Human Sciences, University of Tsukuba, Tsukuba, Japan

We investigated (1) the effects of divided and focused attention on event-related brain potentials (ERPs) elicited by somatosensory stimulation under different response modes, (2) the effects of hand position (closely-placed vs. separated hands) and arm posture (crossed vs. uncrossed forearms) on the attentional modulation of somatosensory ERPs, and (3) changes in the coupling of stimulus- and response-related processes by somatosensory attention using a single-trial analysis of P300 latency and reaction times. Electrocutaneous stimulation was presented randomly to the thumb or middle finger of the left or right hand at random interstimulus intervals (700–900ms). Subjects attended unilaterally or bilaterally to stimuli in order to detect target stimuli by a motor response or counting. The effects of unilaterally-focused attention were also tested under different hand and arm positions. The amplitude of N140 in the divided attention condition was intermediate between unilaterally attended and unattended stimuli in the unilaterally-focused attention condition in both the mental counting and motor response tasks. Attended infrequent (target) stimuli elicited greater P300 in the unilaterally attention condition than in the divided attention condition. P300 latency was longer in the divided attention condition than in the unilaterally-focused attention condition in the motor response task, but remained unchanged in the counting task. Closely locating the hands had no impact, whereas crossing the forearms decreased the attentional enhancement in N140 amplitude. In contrast, these two manipulations uniformly decreased P300 amplitude and increased P300 latency. The correlation between single-trial P300 latency and RT was decreased by crossed forearms, but not by divided attention or closely-placed hands. Therefore, the present results indicate that focused and divided attention differently affected middle latency and late processing, and that hand position and arm posture also differently affected attentional processes and stimulus–response coupling.

KEYWORDS

touch, somatosensation, attention, resource, cognition, anatomical space, physical space, P3

Introduction

Humans often face situations that require attention to both hands, such as typing, driving, cooking, playing sports, and playing an instrument. In some of these activities, the focus of attention is sometimes directed to the body part under different positions, such as crossing the forearms and placing the hands closely or separately, with and without overt motor behavior.

Such unhabitual postures of body parts have been reported to affect various perceptual processes (Eimer et al., 2001; Kennett et al., 2001; Eimer et al., 2004). Previous studies demonstrated that attention facilitated behavioral performance in individual sensory modalities. Furthermore, attentional increases were noted in the amplitude of early and middle latency components of event-related brain potentials (ERPs) elicited by somatosensory stimulation (Desmedt and Robertson, 1977; Garcia-Larrea et al., 1995; Eimer et al., 2002; van Velzen et al., 2002; Valeriani et al., 2003; Kida et al., 2004b,c; Forster and Eimer, 2005; Kida et al., 2006; Gherri and Eimer, 2008; Press et al., 2008; Adler et al., 2009; Keil et al., 2017; Novicic and Savic, 2023; Savic et al., 2023). Auditory spatial selective attention exerts two types of effects on N1 amplitude: the superimposition of another negativity (processing negativity, PN, or its negative difference between attended and unattended channels, Nd) and the enhancement of N1 itself (Hillyard et al., 1973; Naatanen et al., 1978; Naatanen, 2000), and visual spatial attention also exerts both of these effects on amplitudes in the N1-P2 latency range (Johannes et al., 1995; Hillyard et al., 1998; Hillyard and Anllo-Vento, 1998). Regarding somatosensory spatial attention, previous studies demonstrated that an increase in N140 amplitude by selective spatial attention was caused by the superimposition of PN (Michie et al., 1987; Garcia-Larrea et al., 1995; Kida et al., 2004b), whereas others reported an enhancement of the exogenous component (Josiassen et al., 1982). Somatosensory Nd has been used to extract attentional modulations under different conditions (Eimer and Driver, 2000; Eimer et al., 2001; Eimer and Forster, 2003a). The modality-non-specific, late ERP component, P300, has been associated with the amount of attentional resource (Wickens et al., 1983; Kramer et al., 1985; Kida et al., 2004a, 2012a; Reuter et al., 2013; Munoz et al., 2014; Reuter et al., 2014; Akaiwa et al., 2022) as well as subjective probability and stimulus uncertainty (Johnson, 1986; Kok, 2001; Polich, 2020). However, the mechanisms by which somatosensory ERPs are modulated when attention is divided between different hands with and without overt motor responses and those by which the attentional modulation of ERPs is affected by hand position and arm posture have not yet been elucidated.

Previous studies in the auditory modality demonstrated that the amplitude of N1 in the divided attention condition was intermediate between those elicited by the attended and unattended channels during focused attention (Hink et al., 1977, 1978; Parasuraman, 1978). Regarding the distribution of attention, a gradient was detected in visual (Mangun and Hillyard, 1988; Wijers et al., 1989; Heinze et al., 1994), auditory (Teder-Salejarvi and Hillyard, 1998; Teder-Salejarvi et al., 1999), and somatosensory ERPs (Heed and Roder, 2010). Some studies reported somatosensory-specific findings on attentional selectivity and gradients using ERPs (Eimer and Forster, 2003b; Forster and Eimer, 2004) and MEG (Kida et al., 2018). Moreover, psychophysical studies showed that visual (LaBerge, 1983; Downing and Pinker, 1985; Shulman et al., 1985, 1986), auditory (Mondor and Zatorre, 1995; Rorden and Driver, 2001), and somatosensory attention (Craig, 1985; Evans et al., 1992; Rinker and Craig, 1994; Lakatos and Shepard, 1997) had a gradient. In addition to the early component, the amplitude of P300 has been considered to reflect the amount of the attentional resource, and the correlation of single-trial P300 latency with reaction times has been associated with the allocation of resources (Kida et al., 2004a, 2012b). Therefore, different ERP components may provide useful information on the effects of divided attention within the somatosensory modality at different stages. In consideration of divided attention to different body parts, somatosensory attention to a stimulus is closely associated with attention to an action regarding target locations (Gherri and Eimer, 2010), i.e., the target body parts, in contrast to other sensory modalities. Therefore, further studies are needed to establish whether focused and divided attention exert the same effects on somatosensory processing in both a motor (overt) response task and mental (covert) task.

The position of the hands and posture of the arms have been reported to affect behavioral performance and cortical activation regarding attentional processing (Eimer et al., 2001; Kennett et al., 2001; Eimer et al., 2004; Gherri and Forster, 2012a). Previous studies demonstrated that somatosensory ERPs were markedly affected by hand position and arm posture, with the attentional effect being smaller for crossed forearms than for uncrossed forearms (Eimer et al., 2001; Gherri and Forster, 2012a). Another ERP study employed a cue-target attention task to show that posterior late directing attention positivity (LDAP) elicited during the cue-target interval and the attentional enhancement of somatosensory N140 amplitude increased when the hands were wide apart (Eimer et al., 2004). Neural modulations by manipulating hand position and arm posture may be associated with an interaction or incongruence between anatomical and external spaces where body parts receive somatosensory inputs. A psychophysical study indicated that the gradient of somatosensory attention depended on the physical space, but not the anatomical space (Lakatos and Shepard, 1997), whereas ERP studies showed that ERP modulations by crossing the forearms were caused by an incongruency between different spatial coordinates (Eimer et al., 2001; Gherri and Forster, 2012a). Physical space is a three-dimensional extent in the physical world whereas anatomical space or reference frame is the extent based on the body of the perceiver. That is, the difference between the two spatial codes is whether these are based on the physical (external) world or our body (internal world). More concretely, anatomical space codes the location of a somatosensory stimulus according to a somatotopic map where specific body locations are determined by the position of the stimulated cutaneous receptors and their cortical representation (Gherri and Forster, 2012a). Hence, anatomical codes are independent of the position of the body in physical or external space. We here use these two terms to describe what kinds of spatial reference frame attention is coordinated in. These terms have been used in a number of previous studies (Eimer et al., 2001, 2004; Gillmeister and Forster, 2012; Gherri and Forster, 2012a,b). Modality-non-specific late ERP P300 may be sensitive to changes in hand position and arm posture, which yield an interaction between different sensory coordinates. The amplitude of P300 has been implicated in the allocation of modality-non-specific attentional resources (Wickens et al., 1983; Kramer et al., 1985; Kok, 2001; Kida et al., 2004a, 2012a,b) as well as post-stimulus uncertainty (Sutton et al., 1965; Johnson, 1986; Polich, 2007). Therefore, manipulations of hand position and arm posture may affect post-stimulus uncertainty or resource allocation, resulting in changes in P300 and behavioral measures as well as their association.

A classical technique, the adaptive correlation filter (Woody, 1967), has been used to estimate single-trial P300 latency in association with behavioral reaction times. The correlation of single-trial P300 latency with reaction times has been successfully used to examine the coupling and decoupling of stimulus- and response-related processes under various task conditions; e.g.,

speed-*vs*-accuracy task instructions (Kutas et al., 1977; Pfefferbaum et al., 1983) and dual-task performance (Kida et al., 2012b). These studies reported diverse findings, which may be explained by the task-dependent coupling modes of stimulus- and response-related processes. Therefore, this technique may effectively detect changes in stimulus–response coupling caused by focused vs. divided attention or by manipulating hand position and arm posture.

In a series of three experiments, we examined the effects of focused versus divided attention on somatosensory ERPs, and also the impact of hand position and arm posture on the effects of focused attention. The aims of the present study were to clarify (1) whether somatosensory attention divided between the hands produced the same pattern of modulation of ERPs as reported in other modalities, (2) whether the patterns of the attentional modulation of ERPs were similar between mental (covert) and motor (overt) target detection tasks, (3) whether the effects of somatosensory selective attention were based on anatomical or physical spaces or their interaction, (4) whether the modality-non-specific late component (P300) was more sensitive to changes in hand position and arm posture than that of an earlier component (N140), and (5) the mechanisms by which these factors, including the type of attention, hand position, and arm posture, affect stimulus–response coupling assessed by the correlation of single-trial P300 latency with reaction times.

Materials and methods

The present study consisted of 3 experiments (Figure 1). Experiment 1 examined ERP modulations by focused and divided attention in a mental counting task. Experiment 2 investigated ERP modulations by focused and divided attention in a motor response task, and also ERP modulations by focused attention in different hand positions (closely-placed vs. separated hands). Experiments 1 and 2 were conducted in the uncrossed forearm position. Experiment 3 examined ERP modulations by focused attention in uncrossed and crossed forearm postures.

Subjects

Ten right-handed healthy adults (1 female, 9 males), aged 22–30 years old, participated in experiment 1. Ten adults (1 female, 9 males), aged 23–30 years old, participated in experiment 2. Ten adults (2 female, 8 males), aged 23–30 years old, participated in experiment 3. In the present study, the inclusion criterion was an age of 20–40 years, while exclusion criteria were a history of neurological and psychiatric diseases, neurological surgery, and substance abuse. Some subjects participated in two or three experiments in a random order with at least a one-month interval between experiments. The present study was approved by the Ethics Committee, Graduate School of Comprehensive Human Sciences at the University of Tsukuba.

Stimulation

Electrocutaneous stimuli (square wave, constant current pulse) of a 0.2-ms duration were presented to the left thumb (40%, standard) or middle finger (10%, deviant) and right thumb (40%, standard) or middle finger (10%, deviant) in a random order through ring electrodes attached to the first (anode) and second (cathode) interphalangeal spaces. The stimulus intensity was adjusted to approximately 2.5-fold the subject's sensory threshold and was never reported as painful and uncomfortable. Interstimulus intervals varied randomly between 700 and 900 ms for 11 steps (mean 800 ms).

Task condition

Subjects were seated comfortably in a chair in an electrically-shielded room, placed their hands with the palms down on a wooden board, and performed several attention conditions in each experiment. They were instructed to look at a crosshair 1.5 m in front of them and not to look at their hands during performance of the task. In all experiments, each condition consisted of 4 runs of approximately 200–300 stimuli, resulting in 1000 stimuli (400 left and 400 right standard stimuli and 100 left and 100 right deviant stimuli). The order of conditions was randomized among subjects. The interval between conditions was approximately 3 min, and that between runs was about 1 min. In all experiments, a control condition was performed, where subjects were instructed to relax and look at a crosshair 1.5 m in front of them and had no task. The control condition allowed us to assess whether ERPs were facilitated on the attended side or were suppressed on the unattended side (Alho et al., 1987; Desmedt and Tomberg, 1989, 1991; Garcia-Larrea et al., 1995; Kida et al., 2004b).

Experiment 1 (focused vs. divided attention in a mental counting task)

Each subject performed 4 different conditions including a control condition (Figure 1). In the attend-right condition, they silently counted the number of infrequent deviant stimuli (targets) presented to the right middle finger. In the attend-left condition, they silently counted the number of infrequent deviant stimuli presented to the left middle finger. The attend-right and attend-left conditions were regarded as the focused (or unilateral) attention conditions. In the divided attention condition, they counted the number of infrequent deviant stimuli presented to the right and left hands. The thumbs of the left and right hands were located separately at a distance of 50 cm. Subjects were asked to report the number after the termination of each run. The number of target stimuli slightly differed (difference of 0–6) among the runs to prevent subjects from assuming the number without counting; however, each condition consisted of the same number of target (deviant) stimuli.

Experiment 2 (focused vs. divided attention and effect of hand position in a motor response task)

This experiment was the same as experiment 1, except for the type of target detection and manipulation of hand position. Each subject performed 5 different conditions where the direction of attention and hand position varied, plus a control condition (a total of 6 task conditions). Subjects were instructed to press a button with the index finger as fast as possible when they detected infrequent deviant stimuli presented to the left or right middle finger, respectively, in two focused attention (attend-left or right) conditions with the same hand position as that in experiment 1 (i.e., with a 50-cm inter-hand distance). In the divided attention condition, they responded to infrequent deviant stimuli to the

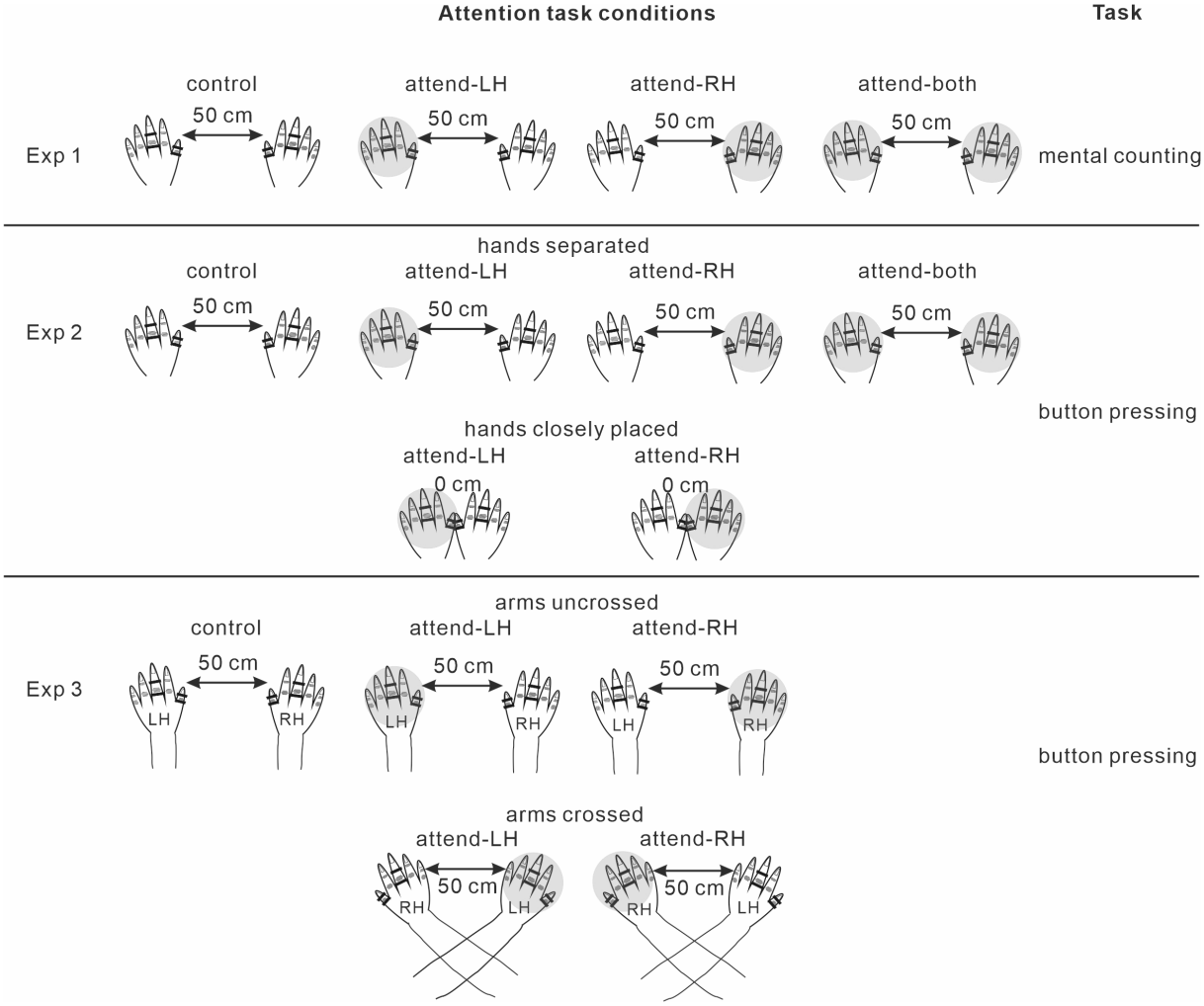


FIGURE 1
Attention task conditions in each experiment. Shaded is the to-be-attended hand. LH, left hand; RH, right hand.

right and left hands by pressing a button. In the other 2 conditions, the thumbs were located close to each other such that they were almost in contact, and subjects performed the same task as focused attention conditions. Button pressing was conducted with compatible mapping within the somatosensory modality, irrespective of hand position (i.e., motor response with the right hand to the right-hand stimulus, and motor response with the left hand to the left-hand stimulus, irrespective of hand position).

Experiment 3 (effects of crossing forearms)

Each subject performed 4 attention conditions plus a control condition (a total of 5 conditions). Attend-right and attend-left conditions were performed with the hands located separately at a distance of approximately 50 cm, as in experiments 1 and 2 (uncrossed-forearms condition). In the other 2 conditions, attend-right and attend-left conditions were performed with the forearms crossed (crossed-forearms condition). Subjects were instructed to detect infrequent stimuli on the designated side by button pressing, which was conducted with compatible mapping within the somatosensory modality, irrespective of forearm postures.

Recordings and analysis

Electroencephalograms (0.5–100 Hz) were recorded at a sampling rate of 500 Hz with Ag/AgCl electrodes from 5 locations on the scalp: Fz, Cz, Pz, C3, and C4 (SYNAFIT, Nihon Denki San-ei Corp., Japan). All the electrodes were referenced to the average of earlobes. Impedance was carefully balanced and maintained below 5 kohm. Electrooculograms (EOG) were recorded bipolarly from the right outer canthus and suborbital region to monitor eye movements or blinks. The analysis time was 600 ms, including a 50-ms prestimulus baseline. Trials exceeding ±80 μV (EOG and EEG amplitudes) were automatically excluded from averaging, and trials with eye blinks and eye movements were excluded manually. Trials with omission and commission errors were also excluded from further analyzes. Grand-averaged waveforms were filtered using a low-pass Butterworth filter with a cut-off frequency of 40 Hz.

Count accuracy

Count accuracy (CA) in experiment 1 was computed in each run using the following equation: $CA = 100–100*(abs(correct\ count$

– *subject's count)/correct count)*, and then averaged across 4 runs in each condition. A one-way repeated measures analysis of variance (ANOVA) was performed on CA with attention (3 levels; Attend-right, attend-left, and both) as a factor.

Reaction times and response accuracy

The reaction time (RT) was measured between 100 and 550 ms after the onset of target stimuli in the button-pressing task in experiments 2 and 3. Response accuracy (RA) was computed using the following equation: *RA = 100*((target number-missed target number)/target number)*. A two-way ANOVA was performed separately on RT and RA with attention (3 levels; focused/hands-separated, focused/closely-spaced hands, and divided) and the stimulus hand (2 levels; left and right) as factors in experiment 2. In experiment 3, a two-way ANOVA was performed separately on RT and RA with forearm posture (crossed and uncrossed) and the stimulus hand (left and right) to compare the effects of forearm posture on behavioral measures.

Statistical analysis of ERPs

N140 was observed for both standard and deviant stimuli, in contrast to only P300 for target stimuli. The peak amplitude and latency of N140 at Fz and the contralateral central site (the average between C3 for right-hand stimulation and C4 for left-hand stimulation) and P300 at Pz were measured within time windows of 110–180 and 250–550 ms, respectively. Time windows were selected based on previous studies that investigated these ERP components (Kida et al., 2004a,b, 2012a,b). N140 is considered to be generated in frontal areas, such as the anterior cingulate cortex (Tanaka et al., 2008), supplementary motor area (Allison et al., 1992), and second somatosensory cortex (Tarkka et al., 1996; Valeriani et al., 2000), and, thus, is recorded maximally at the frontal and central midline electrodes. Furthermore, in consideration of the involvement of the post-central region, some studies examined N140 at the contralateral central electrode (Eimer and Driver, 2000; Kennett et al., 2001; Eimer et al., 2003b; Forster and Eimer, 2004; Gherri et al., 2023). Some source modeling studies reported the partial generation of N140 in the parietal cortex, such as SI (Valeriani et al., 2001). Based on these findings and the distribution of data in the present study, we focused on data obtained from the Fz and contralateral central electrodes. P300 generally showed a broad distribution, with the maximal amplitude being obtained at the parietal electrode in all modalities. In our experience, somatosensory P300 showed the maximal amplitude at the central (Cz) and parietal (Pz) electrodes (Kida et al., 2003a,b,c, 2004a, 2012b). Therefore, we focused on P300 data obtained from Pz in the present study. In each of experiments 1 and 2, for the amplitudes of N140, a two-way ANOVA was performed with condition (control, focused, unattended, and divided) and stimulus type (standard vs. deviant) as factors. In experiment 2, we also performed a three-way ANOVA of N140 amplitude with condition (attended vs. unattended), hand position (closely-placed vs. separated), and stimulus type (standard vs. deviant). In experiment 3, we performed a three-way ANOVA of N140 amplitude with condition (attended vs. unattended), forearm posture (crossed vs. uncrossed), and stimulus type (standard vs. deviant). P300 was observed for attended deviant (target) stimuli in the focused attention and divided attention conditions, whereas unattended deviant stimuli in the focused attention condition, identical deviant stimuli in the control condition, and standard stimuli did not elicit P300. Therefore, a paired *t*-test was performed to compare the amplitude or latency of P300 between focused attention and divided attention conditions, between hand positions (closely-placed vs. separated), and between forearm postures (crossed vs. uncrossed). In ANOVA, if the assumption of sphericity was violated in Mauchly's test, the degree of freedom was corrected using the Greenhouse–Geisser correction coefficient epsilon, and the value of p was then recalculated. The significance level was set at $p < 0.05$. A multiple comparison test with Šidák correction was used for the post-hoc analysis. Therefore, the reported value of p in the multiple comparison test was based on an adjusted value computed backward; i.e., the adjusted value of p, *p(adjusted)*, was computed using the following equation: *p(adjusted) = 1-(1-p(unadjusted)^c)*, where *c* is the comparison number. The Šidák correction was applied to multiple comparisons, including behavioral and ERP data. Partial-eta squared (η_p^2) was computed as an effect size measure in ANOVA. Cohen's *d* was also reported as an effect size measure in the paired *t*-test.

Analysis of single-trial P300 latency and RT

We used an adaptive correlation filter method to estimate single-trial P300 latency (Woody, 1967; Kutas et al., 1977; Kida et al., 2012b). Data measured at Pz in response to target stimuli were used in this analysis. We followed our previous analysis procedure for this technique (Kida et al., 2012b). We selected only the trials showing a cross-correlation coefficient, $R > 0.80$, in the final template, i.e., we regarded these trials as good estimates. The correlation coefficient between single-trial P300 latency and RT was examined in each condition. Cohen's *q* was computed as an effect size measure for the significance of differences between two correlations.

Results

Experiment 1 (focused vs. divided attention in a mental counting task)

N140 amplitude was increased more by focused attention than by the other conditions examined. Furthermore, N140 amplitude in the divided attention condition was intermediate between the focused and unattended conditions. P300 amplitude was lower in the divided attention condition than in the focused attention condition. Detailed information is provided below.

Behavioral data

CA was slightly lower when attention was divided between hands than when it was focused on one hand (Table 1) (F (2, 18) = 1.7, $\eta_p^2 = 0.16$).

TABLE 1 Means (±SE) of count accuracy (CA) in the silent counting task in experiment 1.

	LH target	RH target	Divided attention
CA (%)	98.6 (0.8)	98.7 (0.6)	97.3 (1.1)

LH, left hand; RH, right hand.

ERPs

N140 and P300 components were detected in all subjects (Figure 2). In the two-way ANOVA of N140 amplitude with attention (control, attended, unattended, and divided) and stimulus type (standard vs. deviant), there were significant main effects of the attention condition (F (3, 27) = 7.5, $p < 0.001$, $\eta_p^2 = 0.69$ at Fz; F (3, 27) = 9.7, $p < 0.001$, $\eta_p^2 = 0.52$ at the contralateral central site) and stimulus type (F (1, 9) = 20.3, $p < 0.001$, $\eta_p^2 = 0.69$ at Fz; F (1, 9) = 13.6, $p < 0.005$, $\eta_p^2 = 0.60$ at the contralateral central site). N140 amplitude was higher in the focused attention condition than in the control ($p < 0.001$), unattended ($p < 0.05$), and divided attention conditions ($p < 0.05$) (Figure 3). N140 amplitude in the divided attention condition was intermediate between the focused attention and unattended conditions. The analysis of the effect of stimulus type showed that deviant stimuli elicited larger N140 than standard stimuli. The interaction between the stimulus type and attention condition was not significant for N140 amplitude. P300 amplitude was significantly lower in the divided attention condition than in the focused attention condition ($t = 4.1$, $p < 0.005$), whereas its latency was not significantly changed ($t = 1.1$) (Figure 3).

Experiment 2 (focused vs. divided attention and closely-placed vs. separated hands in a motor response task)

The effects of focused and divided attention on ERP amplitude were similar to those in experiment 1 (Figure 4). P300 latency was

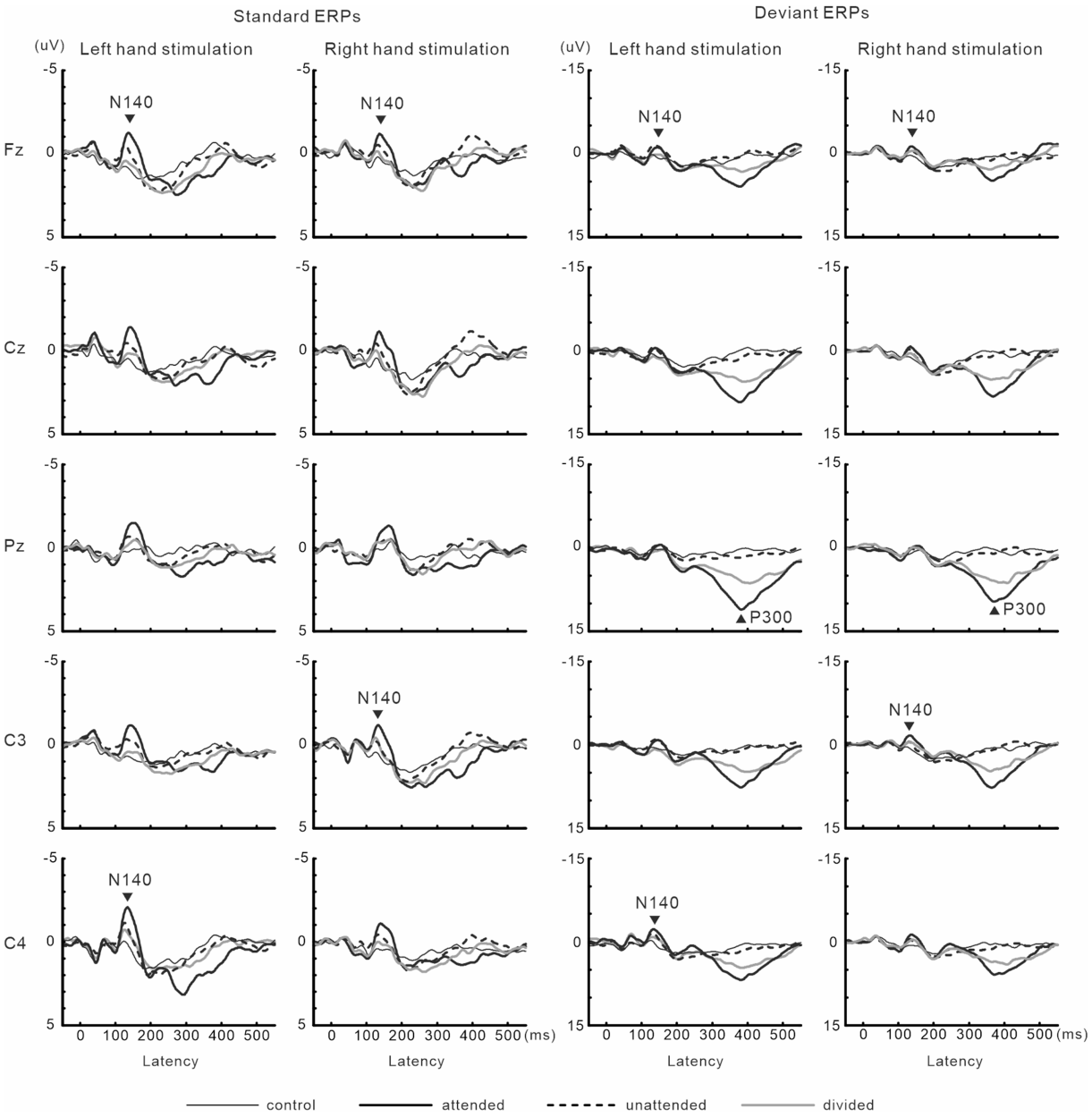


FIGURE 2
Grand-averaged waveforms of ERPs elicited by standard and target stimuli in experiment 1.

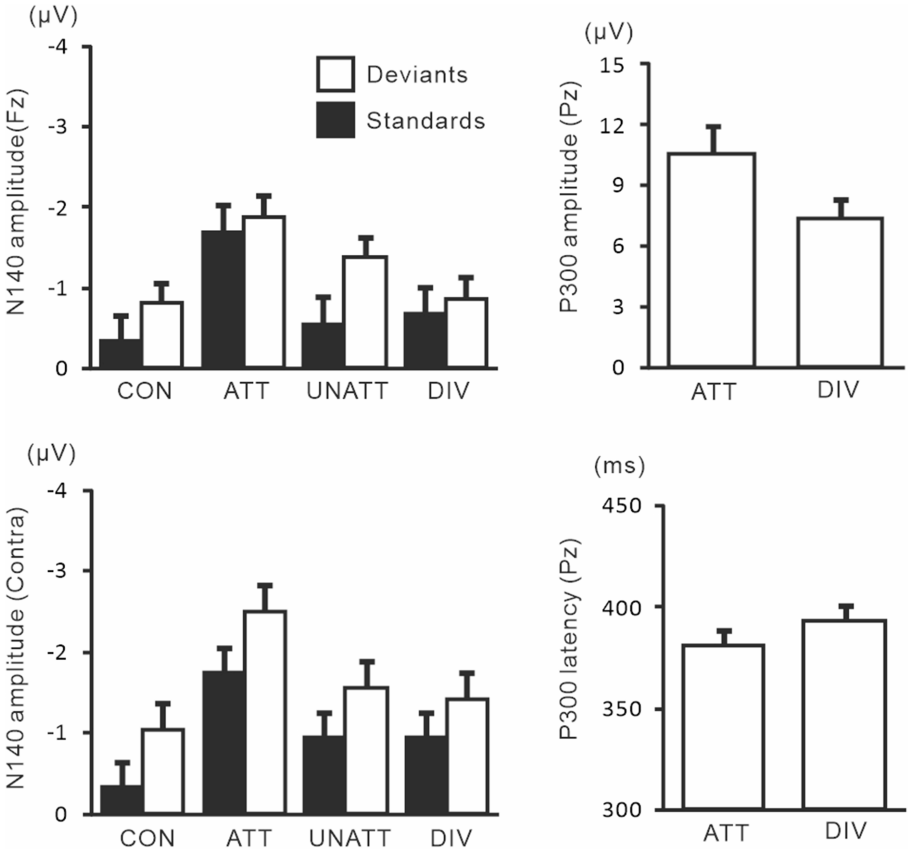


FIGURE 3
Mean values of ERP amplitudes and latencies across subjects in experiment 1. ERP modulations are shown for comparison between unilaterally focused vs. divided attention conditions in a mental counting task. The amplitudes and latencies averaged across left-and right-hand stimuli are shown here because there was no main effect of the stimulus hand and no interaction, including the stimulus hand factor. CON, control condition; ATT, attended stimuli; UNATT, unattended stimuli; DIV, divided attention condition; Contra, contralateral central electrode to stimulation.

longer in the divided attention condition than in the focused attention condition. Hand position did not affect N140, but changed P300 amplitude and latency. Detailed information is provided below.

Behavioral data

The two-way ANOVA of RT showed a significant main effect of attention (F (2, 9) = 4.7, $p < 0.05$, $\eta_p^2 = 0.34$), such that RT was significantly longer when attention was divided between the hands than when it was focused on one hand ($p < 0.05$) (Table 2). The two-way ANOVA of RA showed a significant main effect of attention (F (2, 18) = 7.2, $p < 0.01$, $\eta_p^2 = 0.45$), such that RA was lower when attention was divided between the hands than when it was focused on one hand ($p < 0.05$) or when the hands were closely placed than when separated ($p < 0.005$) (Table 2). There was no significant interaction for RT or RA.

ERPs

In the two-way ANOVA of N140 amplitude with attention (control, attended, unattended, and divided) and stimulus type (standard vs. deviant), there were significant main effects of the attention condition (F (3, 27) = 20.9, $p < 0.001$, $\eta_p^2 = 0.69$ at Fz; F (3, 27) = 11.4, $p < 0.001$, $\eta_p^2 = 0.56$ at the contralateral central site) and stimulus type (F (1, 9) =8.1, $p < 0.01$, $\eta_p^2 = 0.47$ at Fz; F (1, 9) = 14.1, $p < 0.05$, $\eta_p^2 = 0.61$ at the contralateral central site). N140 amplitude was higher in the focused attention condition than in the control ($p < 0.001$ at Fz and $p < 0.05$ at the contralateral central site), unattended ($p < 0.05$ at Fz and p < 0.005 at the contralateral central site), and divided attention conditions ($p < 0.05$) (Figure 5). N140 amplitude in the divided attention condition was intermediate between the focused and unattended conditions. The analysis of the stimulus type effect showed that deviant stimuli elicited larger N140 than standard stimuli. Regarding N140 amplitude, there was a significant interaction between the stimulus type and attention condition (F (3, 27) = 3.4, $p < 0.05$, $\eta_p^2 = 0.27$ at the contralateral central site), such that the amplitude was higher for focused attention and divided attention standard stimuli than for unattended standard stimuli, whereas it was higher for focused attention deviant (target) stimuli than deviant stimuli in the control condition and unattended deviant stimuli. P300 amplitude was significantly lower ($t = 3.7$, $p < 0.005$) and latency was longer ($t = 3.9$, $p < 0.005$) in the divided attention condition than in the focused attention condition.

Figure 6 shows grand-averaged ERP waveforms in the hands closely-placed and separated conditions. We performed a 3-way ANOVA of N140 amplitude with attention (attended and unattended), hand position (closely-placed vs. separated), and the stimulus type (standard and deviant) to examine the impact of

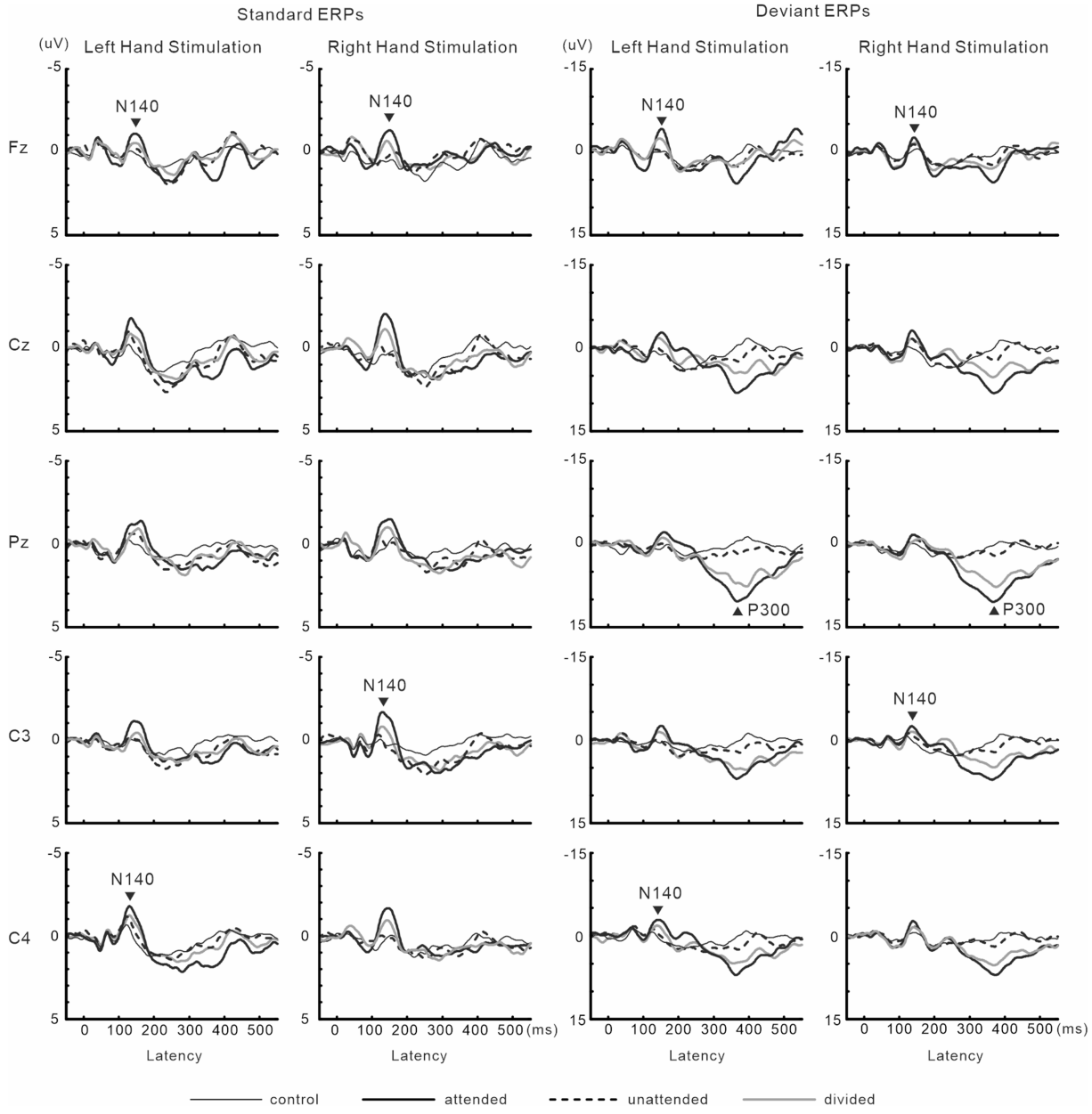


FIGURE 4
Grand-averaged waveforms of ERPs elicited by standard and target stimuli in experiment 2. Data in the hands closely-placed condition are not displayed here, but are shown in Figure 6.

hand position on somatosensory processing and attention effect. There were significant main effects of the attention condition (F (1, 9) = 31.8, $p < 0.001$, $\eta_p^2 = 0.78$ at Fz; F (1, 9) = 15.3, $p < 0.005$, $\eta_p^2 = 0.63$ at the contralateral central site) and stimulus type (F (1, 9) = 11.5, $p < 0.01$, $\eta_p^2 = 0.56$ at Fz; F (1, 9) = 20.2, $p < 0.005$, $\eta_p^2 = 0.69$ at the contralateral central site) (Figure 7). There was also a significant main effect of hand position at the contralateral central site (F (1, 9) = 8.2, $p < 0.05$, $\eta_p^2 = 0.47$), but no significant interactions, including hand position. P300 amplitude was significantly lower ($t = 2.3$, $p < 0.05$) and latency was significantly longer ($t = 2.4$, $p < 0.05$) when the hands were closely placed than when they were wide apart.

Experiment 3 (crossed vs. uncrossed forearms)

Figure 8 shows grand-averaged ERP waveforms in experiment 3. The effect size of focused attention on N140 amplitude was changed by crossing the forearms. Forearm posture affected P300 amplitude and latency. Detailed information is provided below.

Behavioral data

Regarding RT, there was a main effect of forearm posture (F (1, 9) = 7.3, $p < 0.05$, $\eta_p^2 = 0.34$), with RT being longer when the forearms were crossed than when they were uncrossed independent of the

TABLE 2 Means (±SE) of the reaction time (RT) and response accuracy (RA) in a button-pressing task in experiment 2.

	Focused (unilateral) attention				Divided attention	
	Hands separated		Hands closely placed			
	LH target	RH target	LH target	RH target	LH target	RH target
RT (ms)	376.3 (14.2)	371.7 (14.8)	374.5 (14.5)	369.1 (15.6)	394.7 (13.8)*	393.3 (11.4)*
RA (%)	95.7 (1.2)	96.9 (0.9)	90.7 (1.4)*	92.2 (1.9)*	90.3 (2.6)*	87.9 (3.3)*

*$p < 0.05$, vs. Focused attention/Hands separated.

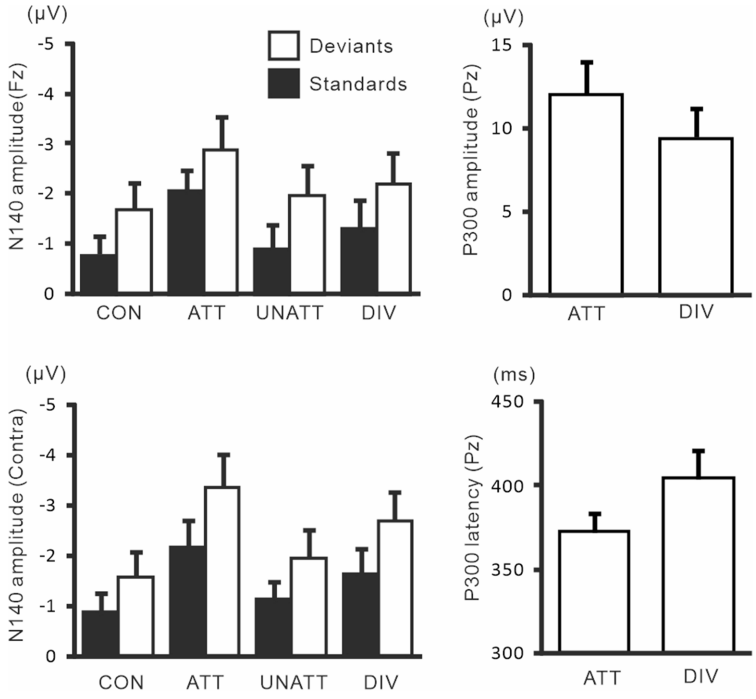


FIGURE 5
Mean values of ERP amplitudes and latencies across subjects in experiment 2. ERP modulations are shown for comparison between unilaterally focused vs. divided attention conditions in a motor response task. Amplitudes and latencies averaged across left-and right-hand stimuli are shown.

stimulus hand (Table 3). Concerning RA, there was a main effect of forearm posture (F (1, 9) = 18.2, $p < 0.005$, $\eta_p^2 = 0.67$), with RA being lower when the forearms were crossed than when they were uncrossed independent of the stimulus hand. No interaction was found for RA.

ERPs

Figure 9 shows the amplitudes and latencies of ERPs in Experiment 3. In the 3-way ANOVA of N140 amplitude, there was a significant main effect of the attention condition (F (1, 9) = 21.9, $p < 0.05$, $\eta_p^2 = 0.71$ at Fz; F (1, 9) = 18.5, $p < 0.01$, $\eta_p^2 = 0.67$ at the contralateral central site), with amplitude being higher when attention was directed to one hand than with unattended stimuli. The stimulus type effect was also significant (F (1, 9) = 11.8, $p < 0.01$, $\eta_p^2 = 0.57$ at Fz; F (1, 9) = 12.8, $p < 0.01$, $\eta_p^2 = 0.59$ at the contralateral central site). The interaction including forearm posture was not significant on N140 amplitude at both the Fz and contralateral central sites; however, low-level multivariate ANOVAs showed that the effect size of attention was the highest for standard stimuli with uncrossed forearms (F (1, 9) = 20.4, $p < 0.005$; $\eta_p^2 = 0.69$), intermediate for standard stimuli with crossed forearms (F (1, 9) = 10.4, $p < 0.01$; $\eta_p^2 = 0.55$), and the lowest (but generally a moderate effect) for deviant stimuli with uncrossed (F (1, 9) = 7.5, $p < 0.05$; $\eta_p^2 = 0.46$) and crossed forearms (F (1, 9) = 6.9, $p < 0.05$; $\eta_p^2 = 0.44$). P300 amplitude was lower ($t = 2.6$, $p < 0.05$) and its latency was longer ($t = 3.5$, $p < 0.01$) when the forearms were crossed than when they were uncrossed.

Single-trial P300 latency and RT

In the analysis of single-trial P300 latency using the ACF technique, 38.6 and 39.9% of all trials were identified as good estimates in experiments 2 and 3, respectively. In experiment 2, the two-way ANOVA (3 attention conditions [unilaterally-focused/hands-separated, unilateral focused/hands closely-placed, and divided] and 2 target sides [left-and right-hand targets]) of single-trial P300 latency showed a main effect of the attention condition (F (2, 18) = 4.4, $p < 0.05$; $\eta_p^2 = 0.33$), similar to the analysis of the averaged waveform,

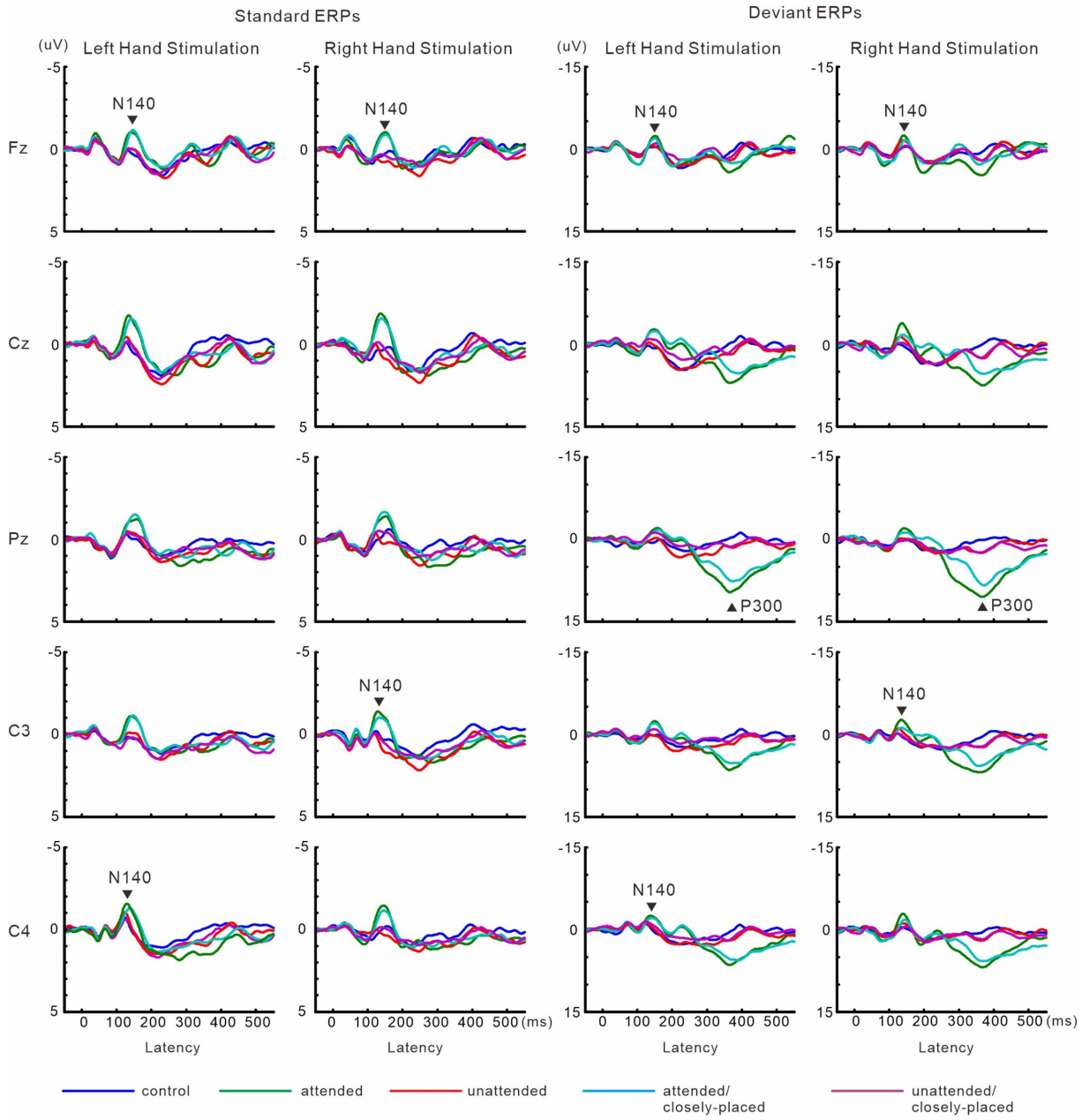


FIGURE 6
Grand-averaged waveforms of ERPs elicited by standard and target stimuli in experiment 2. Data for attended and unattended stimuli in the hands-separated condition, which are identical to that for attended and unattended stimuli in Figure 4, are shown here for comparison with the hands closely-placed condition.

TABLE 3 Means (±SE) of RT and RA in the button-pressing task in experiment 3.

	Forearms uncrossed		Forearms crossed	
	LH target	RH target	LH target	RH target
RT (ms)	369.7 (14.6)	363.5 (15.4)	377.1 (15.1)*	376.9 (16.3)*
RA (%)	95.6 (1.2)	96.1 (1.3)	90.2 (2.0)*	90.6 (2.2)*

*$p < 0.05$, vs. Forearms uncrossed.

with latency being longer with hands closely-placed than separated (Table 4). Divided attention also resulted in slightly longer latency (Table 4) than unilaterally-focused attention. This was the same pattern as that observed in the results on averaged waveforms. The single-trial analysis of RT and P300 latency showed that RT preceded P300 latency in 46.0, 36.9, and 36.6% of good trials in the focused/hands-separated, focused/hands closely-placed, and divided attention conditions, respectively, for left-hand targets, and in 54.6, 43.1, and 39.2%, respectively, for right-hand targets. Therefore, right-hand targets had more trials with RT preceding P300 latency than left-hand targets in all conditions tested, and also hands closely-placed and divided attention decreased the preceding ratio of RT (or the delayed ratio of P300), resulting in a decrease in the difference between left-and right-hand targets. The variabilities (SD) of single-trial P300 latency and RT were not significant in ANOVA with the attention condition and target side. The correlation of single-trial P300 latency

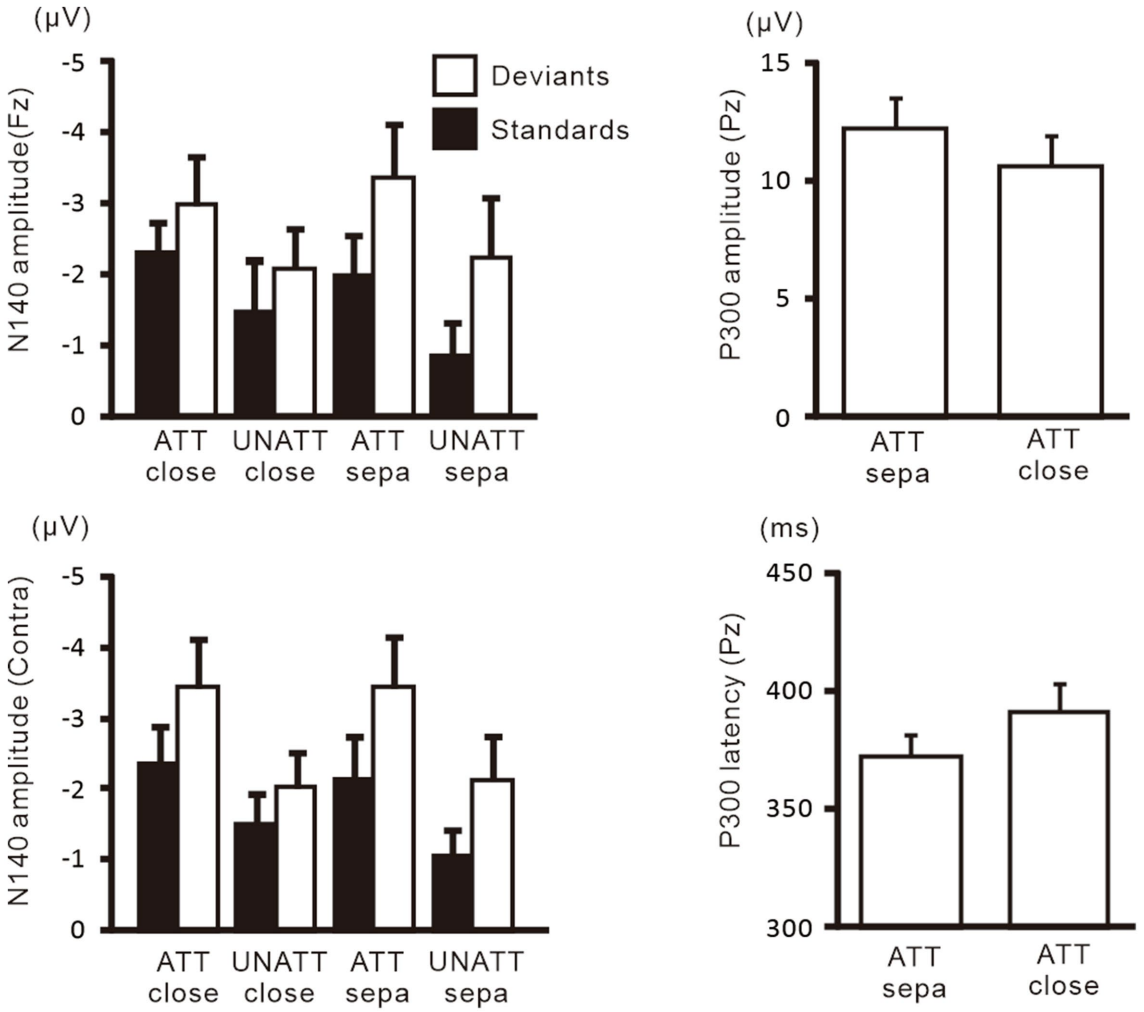


FIGURE 7
Mean values of ERP amplitudes and latencies across subjects in experiment 2. Data for attended and unattended stimuli in the hands-separated condition, which are identical to that for attended and unattended stimuli in Figure 5, are shown here for comparison with the hands closely-placed condition. Amplitudes and latencies averaged across left-and right-hand stimuli are shown. Close, hands closely-placed; sepa, hands-separated.

with RT was moderate in the unilaterally-focused attention/hands-separated condition (correlation coefficient $R = 0.52$) and low in the unilaterally-focused attention/hands closely-placed condition ($R = 0.38$) and divided attention condition ($R = 0.38$), all of which showed correlations (Table 4). The former showed a stronger correlation than the latter two.

In experiment 3, the two-way ANOVA of (2 forearm positions and 2 target sides) single-trial P300 latency showed a main effect of forearm posture ($F(1, 9) = 6.1$, $p < 0.05$; $\eta_p^2 = 0.41$), with latency being longer when the forearms were crossed than when they were uncrossed irrespective of the target stimulus hand (Table 5). Therefore, the same pattern of results was observed as averaged waveforms. The single-trial analysis of RT and P300 latency showed that RT preceded P300 latency in 42.5 and 50.7% of good trials in the uncrossed-and crossed-forearms conditions, respectively, for left-hand targets, and in 59.3 and 51.3%, respectively, for right-hand targets (Table 5). Therefore, right-hand targets had more trials with RT preceding P300 latency in both forearm positions, and crossing the forearms decreased the RT-P300 latency difference between left-and right-hand targets. The variabilities (SD) of single-trial P300 latency and RT were not significant in ANOVA with forearm position and the target stimulus hand. A correlation of single-trial P300 latency with RT was observed in the uncrossed-forearms condition ($R = 0.37$, significant), but not in the crossed-forearms condition ($R = 0.23$). Furthermore, a significant difference was observed in the correlation between the two conditions ($z = 1.98$) with a small effect ($q = 0.15$). When it was examined separately for each target side, the correlation was significantly higher in the uncrossed-forearms condition for right-hand targets ($R = 0.45$) than in the crossed-forearms condition ($R = 0.23$) ($z = 2.29$, $p < 0.05$) with a small effect ($q = 0.25$), whereas left-hand targets showed no significant difference between the crossed- and uncrossed-forearms conditions.

Another result on the correlation of single-trial P300 latency with RT was observed when left-and right-hand target stimuli were separately analyzed. The single-trial P300 latency-RT correlation in most of the conditions examined in experiments 2 and 3 was slightly stronger for right-hand target stimuli than for left-hand target stimuli (the only exception was the crossed-forearms condition), whereas no

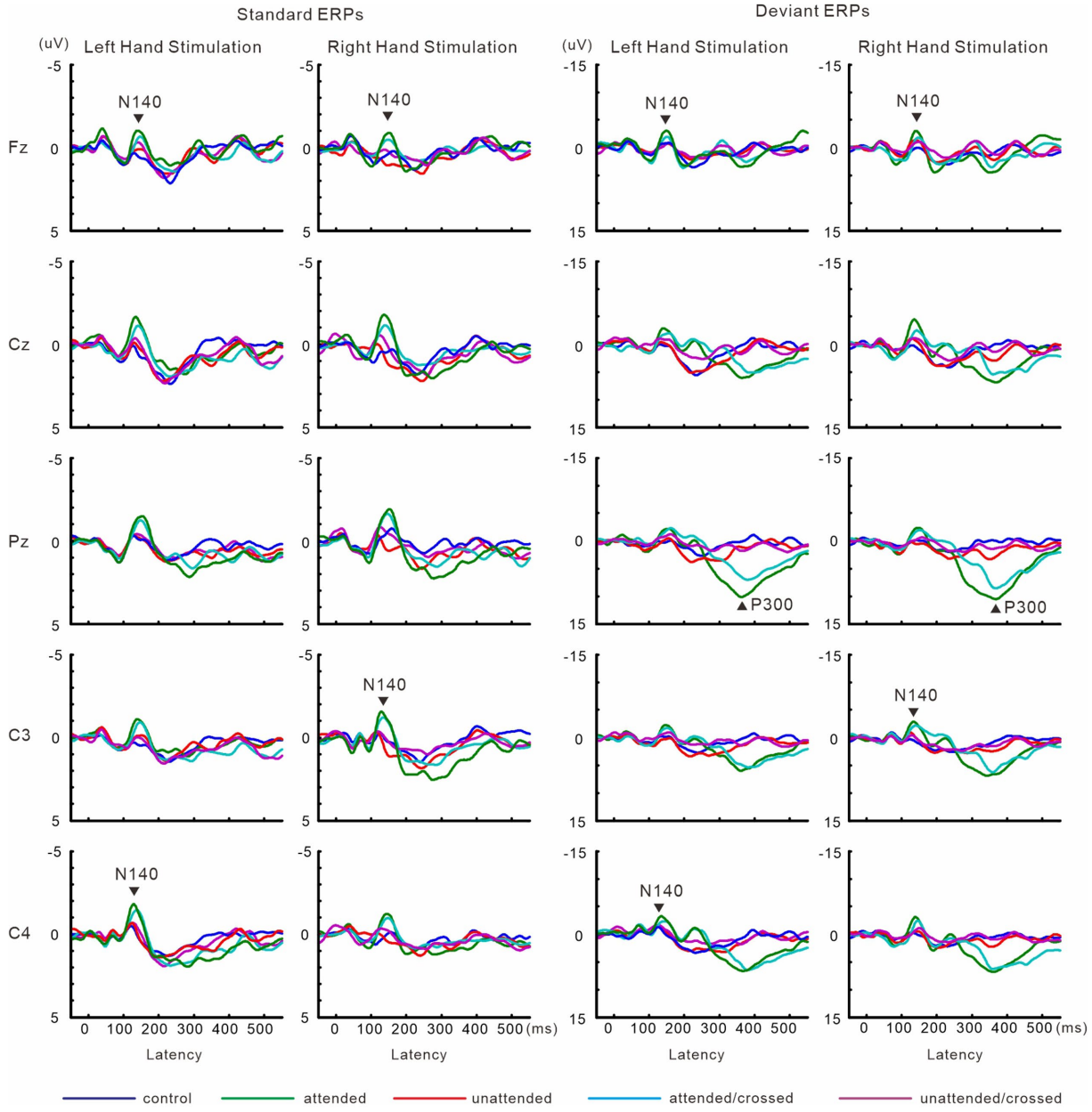


FIGURE 8
Grand-averaged waveforms of ERPs elicited by standard and target stimuli in experiment 3.

significant difference was observed when it was tested separately in each condition (Tables 4, 5). To increase the statistical power and examine correlation patterns general to all attention conditions, we concatenated latency and RT data from all attention conditions for each target side and then compared correlations between different target sides. The analysis revealed correlations for both target sides ($R = 0.31$ for left-hand target, $R = 0.41$ for right-hand target) and a significant difference in the correlation between left-and right-hand target stimuli ($z = 2.28$, $p < 0.05$) with a small effect ($q = 0.12$). Collectively, correlation patterns showed that right-hand attended targets (and responses) produced a stronger P300 latency-RT correlation with a small effect than left-hand targets, which was decreased by crossing the forearms.

Discussion

Replication of the attentional modulation of N140

The present study showed that the amplitude of N140 was modulated by directing attention to the unilateral hand, and was higher for attended stimuli than for unattended stimuli and identical stimuli in the control condition. Therefore, we successfully replicated previous findings on the effects of somatosensory attention on the amplitude of N140 (Desmedt and Robertson, 1977; Desmedt et al., 1984; Michie et al., 1987; Desmedt and Tomberg, 1989; Garcia-Larrea et al., 1995; Eimer and Forster, 2003a; Kida et al., 2004b, 2012a,b).

TABLE 4 Mean (±SE) and SD of single-trial P300 latency, the SD of RT, the ratio of trials with RT preceding single-trial P300 latency, and the correlation of single-trial P300 latency with RT in experiment 2.

	Focused (unilateral) attention				Divided attention	
	Hands separated		Hands closely placed			
	LH target	RH target	LH target	RH target	LH target	RH target
Mean (ms) of single-trial P300 latency	340.0 (3.3)	338.3 (4.1)	349.4 (4.6)**	357.9 (6.9)**	344.6 (6.2)	342.6 (7.4)
Mean of SD of single-trial P300 latency	59.2 (5.8)	61.9 (5.3)	59.6 (5.8)	65.8 (9.9)	63.7 (3.2)	63.9 (4.0)
Mean of SD of single-trial RT	58.9 (4.7)	58.7 (4.3)	55.7 (4.2)	59.4 (4.0)	62.1 (2.6)	61.8 (3.2)
Ratio (%) of trials with RT preceding P300 latency	46.0	54.6	36.9	43.1	36.6	39.2
Correlation of single-trial P300 latency with RT	0.40*	0.52*	0.31*	0.44*	0.34*	0.44*
Correlation of single-trial P300 latency with RT (based on data concatenated across LH and RH)	0.47*		0.38*		0.38*	

*Correlation, $p < 0.05$, vs. zero; **$p < 0.05$, vs. Focused attention/Hands separated.
Data shown here were calculated after excluding bad trials (using the ACF technique), trials with missed targets, and trials with artifacts.

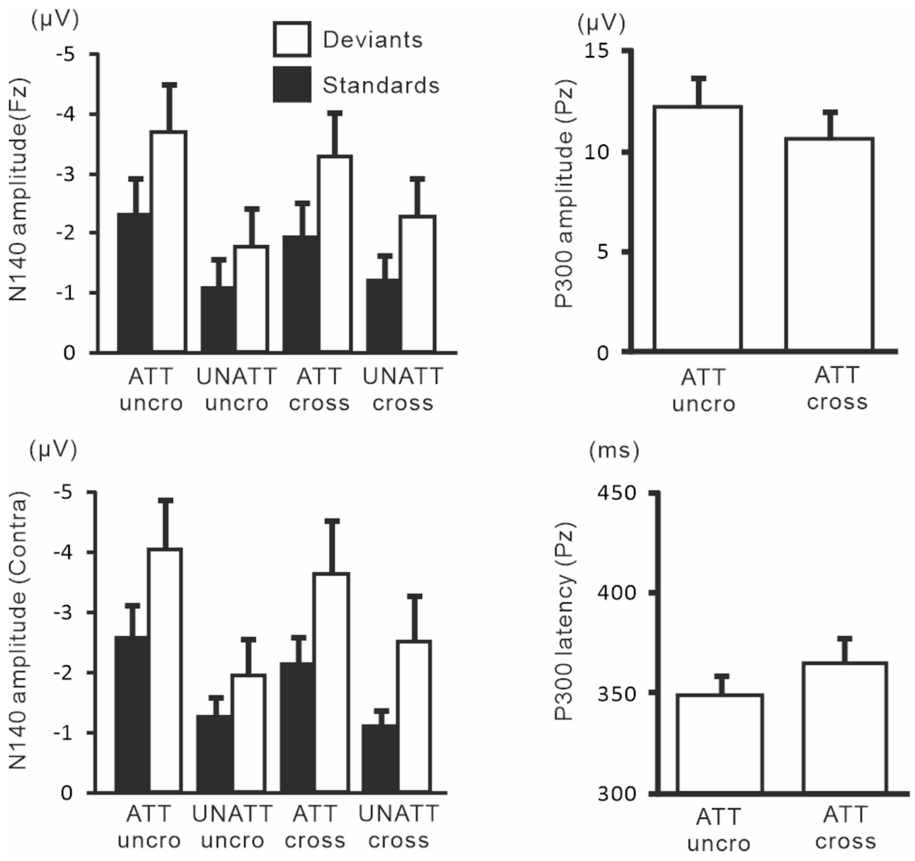


FIGURE 9
Mean values of ERP amplitudes and latencies across subjects in experiment 3. ERP modulations are shown to compare crossed- vs. uncrossed-forearms conditions. Amplitudes and latencies averaged between left-and right-hand stimuli are shown. Cross, crossed forearms; uncro, uncrossed forearms.

Effects of divided attention on ERPs

N140 amplitude in the divided attention condition was intermediate between those elicited by attended and unattended stimuli during the focused attention condition. Previous studies in audition showed that the amplitude of N1 in the divided attention condition was intermediate between those elicited by attended and unattended stimuli during the focused attention condition (Hink et al., 1977, 1978; Parasuraman, 1978), which was consistent with the present results. Therefore, dividing attention between the hands may be controlled similarly to auditory divided attention, which may be explained by a capacity model of attention (Hink et al., 1977). Some researchers have also suggested that the attentional modulation of early ERP components is associated with the perceptual resource

TABLE 5 Mean (±SE) and SD of single-trial P300 latency, the SD of RT, the ratio of trials with RT preceding single-trial P300 latency, and the correlation of single-trial P300 latency with RT in experiment 3.

	Forearms uncrossed		Forearms crossed	
	LH target	RH target	LH target	RH target
Mean (ms) of single-trial P300 latency	335.3 (4.9)	335.8 (5.6)	358.6 (7.8)**	363.8 (10.9)**
Mean of SD of single-trial P300 latency	57.8 (7.1)	61.3 (5.1)	61.8 (5.8)	63.4 (6.7)
Mean of SD of single-trial RT	56.7 (4.4)	55.8 (3.3)	59.9 (3.7)	57.2 (4.0)
Ratio (%) of trials with RT preceding P300 latency	52.5	55.3	50.7	51.3
Correlation of single-trial P300 latency with RT	0.32*	0.45*	0.24	0.23**
Correlation of single-trial P300 latency with RT (based on data concatenated across LH and RH)	0.37*		0.23**	

*Correlation, $p < 0.05$, vs. zero; **$p < 0.05$, vs. Forearms uncrossed.
Data shown here were calculated after excluding bad trials (using the ACF technique), trials with missed targets, and trials with artifacts.

(Kok, 1997; Kida et al., 2012a,b) in the framework of multiple resources, including perceptual, central, and response resources (Wickens, 1991; Wickens and McCarley, 2008).

Regarding the distribution of attention, previous studies demonstrated a spatial gradient of attention in different modalities using visual (Mangun and Hillyard, 1988; Wijers et al., 1989; Heinze et al., 1994), auditory (Teder-Salejarvi and Hillyard, 1998; Teder-Salejarvi et al., 1999), and somatosensory ERPs (Heed and Roder, 2010), suggesting modality-independent patterns of the distribution of attention. In addition, visual and auditory ERP studies reported that attention forms a unitary zone that may expand to multiple relevant locations, but also includes the area between them (Mangun and Hillyard, 1988; Wijers et al., 1989; Heinze et al., 1994; Teder-Salejarvi et al., 1999). In contrast, a previous study on touch reported that when attention was directed simultaneously to non-adjacent fingers within one hand, ERPs in response to stimuli delivered to spatially and anatomically intervening fingers showed no attentional modulations (Eimer and Forster, 2003b). This study concluded that, in contrast to vision, the focus of somatosensory attention may be split and directed simultaneously to non-adjacent areas, thereby excluding spatially and anatomically intermediate regions from attentional processing. An MEG study also reported that somatosensory attention has a gradient and may also be divided into non-adjacent areas (Kida et al., 2018), supporting ERP results. These electrophysiological findings indicate a somatosensory-specific pattern of attention. Based on the present and previous findings, we speculate that somatosensory attention may be split between non-adjacent fingers and also between the hands. The modality specificity of the distribution of attention needs to be examined in more detail in future studies.

P300 was found for attended infrequent stimuli (targets) in both the focused and divided attention conditions, but not for identical stimuli in the control condition or unattended infrequent stimuli in the focused attention condition. This pattern is a common feature of the appearance of P300. In addition, P300 amplitude was lower when attention was directed simultaneously to both hands than when it was unilaterally focused on one hand. Previous studies reported that P300 amplitude reflects the amount of the modality-non-specific, perceptual-central resource allocated to a given task (Wickens et al., 1983; Kramer et al., 1985; Sirevaag et al., 1989; Kok, 1997, 2001; Kida et al., 2004a, 2012a,b). The gradual change observed in P300 amplitude in the present study (i.e., focused attention >divided attention >unattended or control condition) was consistent with the resource allocation view of P300 amplitude. Therefore, the decrease in P300 amplitude in the divided attention condition was assumed to be caused by the division of the modality-non-specific perceptual-central resource between the hands. The common pattern of P300 amplitude to mental counting and motor response tasks shows that the allocation of the modality-non-specific resource to both hands was independent of whether the response was covert (mental) or overt (motor).

In the divided attention condition, target probability was two-fold or target-to-target interval (TTI) was half that in the focused attention condition. P300 amplitude has been shown to decrease with high target probability and a short TTI (Kok, 2001; Polich, 2007). Therefore, it is unclear whether the decrease in P300 amplitude in the divided attention condition was due to resource division or changes in task difficulty following changes in target probability and TTI. However, the resource limitation explanation may account for the potential relationship of target probability and TTI with P300 amplitude. When target stimuli are presented more frequently (a higher target probability or shorter TTI), more resources are consumed in a given amount of time than with less frequently presented stimuli, and P300 amplitude is small. When stimuli are presented more infrequently (lower target probability or longer TTI), the structures involved in the generation of P300 may recover more fully and P300 amplitude is large (Gonsalvez and Polich, 2002; Gonsalvez et al., 2007). This resource limitation explanation accounts for the interaction between task difficulty and target stimulus probability (Kramer et al., 1986; Ruchkin et al., 1987; Polich et al., 1988). Therefore, even if the decrease in P300 amplitude by dividing attention to both hands is associated with higher target probability or shorter TTI, it may also be explained by the resource allocation view.

In contrast to the common pattern in P300 amplitude, P300 latency was longer in the divided attention condition than in the focused attention condition in the motor response task, but not in the mental counting task. This result suggests the functional dissociation of the amplitude and latency of P300. More specifically, the modality-non-specific resource may be divided between the hands without affecting the stimulus evaluation speed during the mental counting task, whereas dividing attention between the hands decreases the evaluation speed during the motor response task. It remains unclear whether motor response demands affect P300 latency, with some studies reporting no effects (Kutas et al., 1977; McCarthy and Donchin, 1981; Magliero et al., 1984; Doucet and Stelmack, 1999). In contrast, other studies noted the significant effects of motor response demands on P300 latency (Ragot and Renault, 1981; Ragot, 1984; Ragot and Renault, 1985; Pfefferbaum et al., 1986; Ragot and Lesevre, 1986; Christensen et al., 1996; Leuthold

and Sommer, 1998). However, the effects of motor response demand on P300 latency have not been investigated under divided versus focused attention conditions in the somatosensory modality. In the motor response task, attention needed to be divided between the hands for both stimulus and action processes. Therefore, attentional demands required for action may be sufficiently high to decrease the stimulus evaluation speed in the motor response task to lower than that in the mental counting task with no motor response demand.

Effects of hand and forearm postures on ERPs

Hand position, such as closely-placed hands, did not significantly affect the attentional modulation of N140 amplitude. This result shows that the attentional modulation of N140 largely depended on the anatomical space rather than the physical space. If an attentional effect on N140 amplitude is exclusively based on the physical (external) space, it is expected to disappear with closely-placed hands because spatial attention will operate equally on both to-be-attended and not-to-be-attended closely-placed hands. In contrast, if an attentional effect on N140 amplitude is based on the anatomical space, it is expected to appear in both separated-hands and closely-placed hands conditions. In addition, changes in hand position follows changes in arm posture. Therefore, a postural difference in the arms with closely-placed hands does not affect the attentional modulation of N140 amplitude.

In contrast to closely-placed hands, crossing the forearms reduced the attentional increase in N140 amplitude for standard stimuli. In a selective attention task, the ERP amplitude for standard stimuli generally reflects a pure selective attention effect, whereas that for target stimuli may contaminate a target-related effect or potential (Garcia-Larrea et al., 1995; Kida et al., 2004b, 2018). Therefore, the attentional increase in N140 amplitude in the present study represents the effects of somatosensory selective attention, which may have operated less efficiently at this stage when the forearms were crossed. Since positional and postural changes to the hands and forearms with closely-placed hands did not affect the attentional increase in N140 amplitude as discussed above, crossing the forearms may be a specific hand position and arm posture leading to changes in attentional somatosensory processing at this stage. A possible explanation for this result is based on the experimental condition that the hands and arms were placed on an unhabitual side with an unhabitual posture. This crossed-forearms condition will produce an incongruency between a representation of actual stimulus (and response) sides and a mental image of an internal space. This incongruency may result in the suppressive effect of crossing the forearms on the attentional increase in N140 amplitude. Previous studies demonstrated that attentional enhancements in early ERP amplitudes were smaller for crossed forearms than for uncrossed forearms (Eimer et al., 2001; Kennett et al., 2001; Eimer et al., 2003a; Gherri and Forster, 2012a,b), supporting the present results. Therefore, the result suggests that the attentional modulation of stimulus processing reflected by N140 amplitude is not only based on the anatomical space, but also the congruency between real and internal spaces depending on the hand position and arm posture. A psychophysical study reported that somatosensory attention was dependent on the physical space, but not on the anatomical space (Lakatos and Shepard, 1997). In contrast, ERP studies provided evidence to show that somatosensory attention was associated with an incongruency between different spatial coordinates (Eimer et al., 2001; Kennett et al., 2001). The present finding supports the latter ERP evidence. This pattern of attentional modulation associated with different spatial codes is also consistent with generators of N140, which originates from modality-specific and multisensory areas, including the second somatosensory, anterior cingulate, and prefrontal cortices (Allison et al., 1992; Tarkka et al., 1996; Waberski et al., 2002; Inui et al., 2003; Tanaka et al., 2008).

In contrast to the early N140 component, P300 amplitude was affected by both hand position and forearm posture. The behavioral measures, RT and RA, paralleled a decrease in P300 amplitude and increase in P300 latency in both the crossed-forearms and hands closely-placed conditions. P300 amplitude has been associated with equivocation or post-stimulus uncertainty (Sutton et al., 1965; Johnson, 1986; Kok, 2001) as well as resource allocation, whereas P300 latency was related to the stimulus evaluation time (Kutas et al., 1977). Therefore, crossing the forearms may decrease the resolution of post-stimulus uncertainty and increase the stimulus evaluation time through a congruency of real and learned spaces, whereas closely-placed hands exert the same effects by overlapping the attentional range at targets and non-targets.

Effects of crossed forearms on stimulus-response coupling

Crossing the forearms significantly decreased the correlation of single-trial P300 latency with RT. As discussed above, crossing the forearms may produce an incongruency between a representation of actual stimulus (and response) sides and a mental image of a learned physical space. A previous study suggested that following accuracy maximizing instructions, subjects hesitated before pressing the button because the task used reduced their confidence of a correct response, thereby decoupling P300 latency from RT (Pfefferbaum et al., 1983). Similarly, an incongruency between actual and learned spaces may be associated with this hesitation before responding, thereby resulting in the decoupling of stimulus- and response-related processing. Crossing the forearms also decreased P300 amplitude and increased P300 latency as already discussed. Therefore, we speculate that crossing the forearms caused the decoupling of stimulus- and response-related processing, decreased the resolution of post-stimulus uncertainty, and reduced the stimulus evaluation speed concomitantly through an incongruency between real and learned spaces.

The present and previous studies using the ACF technique showed the task-dependent nature of the correlation of single-trial P300 latency with RT. A historical study on P300 latency using ACF found that speed-maximizing instructions resulted in a weaker correlation of single-trial P300 latency with RT than accuracy-maximizing instructions, suggesting the loose coupling of stimulus- and response-related processing in the former instructions and motor command output before the stimulus has been fully evaluated (Kutas et al., 1977). In contrast, another study reported the reverse effect, i.e., a weaker correlation of P300 latency with RT under accuracy-maximizing instructions than speed-maximizing instructions (Pfefferbaum et al., 1983). The latter study suggested that the difference in tasks and strategic differences for task requirements explain the discrepancies in the findings obtained as discussed above. We also previously used the ACF technique to demonstrate stronger stimulus–response coupling by the performance of a dual task than a single task (Kida et al., 2012b). Two explanations for this result, a snap decision

strategy in the single-task and the lack of the resource allocated to the eliciting task during the dual-task performance, were suggested because stronger stimulus–response coupling by the dual-task performance was associated with a lower P300 amplitude and smaller number of trials with RT preceding P300 latency. However, the ratio of trials with RT preceding P300 latency was even lower in the present selective attention task (36.9–59.3%) than during single-task performance (76.8%) and was in the same range as that during the dual-task performance (53.7%). This is because the present study employed a type of selective attention task where somatosensory stimuli were presented to many fingers in bilateral hands and subjects had to discriminate one or two target stimuli, i.e., task demand was higher than the simple somatosensory oddball task used in the previous study where two types of somatosensory stimuli were presented to unilateral fingers. RT was earlier in the previous study (349.5 and 336.6 ms in experiments 1 and 2, respectively) than in the present study (ranging between 363.5 and 394.7 ms depending on the task condition, Tables 2, 5), supporting the selective attention task used herein being a more demanding task than the oddball task used in the previous study. Therefore, the snap decision is not the main cause for the changes observed in the coupling of stimulus- and response-related processing in the present study. Another suggestion, the lack of the perceptual-central resource, is also not straightforward to explain our previous findings and the present results because the two studies showed the reverse pattern for P300 amplitude and the correlation of single-trial P300 latency with RT. Therefore, the present and previous studies showed the importance of task features in interpreting the correlation of single-trial P300 latency with RT.

Hand preference of the P300 latency-RT correlation

In the present study, the correlation between single-trial P300 latency and RT was stronger for the right hand than for the left hand in right-handed subjects. Previous studies reported a hand preference (i.e., a difference between the dominant and non-dominant hands) in motor tasks and sensorimotor tasks (Beste et al., 2009), whereas others showed no hand preference in a tactile perceptual task (Finlayson and Reitan, 1976). Regarding neural activation, movements with the non-dominant hand were associated with stronger and more extended brain activation than those with the dominant hand (Leocani et al., 2001; Potgieser et al., 2015). In contrast, an ERP study found that theta power related to handwriting was higher with the dominant hand than with the non-dominant hand (Pei et al., 2021). In addition to these findings from psychophysics and neuroimaging, the present study provides evidence for a hand preference for stimulus–response coupling using the combined measure of behavioral and neural response times. This hand preference of stimulus–response coupling was eliminated by crossing the forearms, but not by placing the hands close together or dividing attention between the hands. As already discussed, crossing the forearms may produce an incongruency between real (external) and learned (internal) spaces, which may eliminate the hand preference in the stimulus–response translation process.

Somatosensory mismatch responses

In the present study, N140 amplitude was higher for deviant stimuli than for standard stimuli in all conditions, including attended and unattended stimuli. This amplitude increase was associated with somatosensory mismatch negativity (s-MMN), which has been reported to occur in the range of this latency in passive tasks (Kekoni et al., 1997; Kida et al., 2004c; Restuccia et al., 2007, 2009; Chen et al., 2014; Shen et al., 2018a; He et al., 2020). MMN is considered to reflect the automatic detection of stimulus changes in the sensory environment (Naatanen, 1992; Naatanen et al., 2007) and has recently been associated with predictive coding (Kimura, 2012; Stefanics et al., 2014). The first demonstration of MMN was in both the unattended and attended channels in an auditory selective attention task (Naatanen et al., 1978), suggesting the automatic nature of MMN. However, s-MMN was previously detected in passive tasks, such as reading (Kekoni et al., 1997; Restuccia et al., 2007, 2009) and video watching (Chen et al., 2014; Shen et al., 2018b), both of which are general MMN recording procedures. In contrast, the present study observed s-MMN in both attended and unattended deviant stimuli in a selective attention task, suggesting that the generation of the response is independent of attention. Similar attention-independent responses to stimulus onset, offset, and change have been demonstrated in both active and passive tasks (Yamashiro et al., 2008, 2009). Previous studies found s-MMN in frontal (Kekoni et al., 1997; Kida et al., 2004c) and parieto-occipital sites (Restuccia et al., 2007, 2009). We also detected s-MMN in frontal and central sites, supporting previous ERP findings. Placing the hands close together and crossing the forearms had no impact on the higher amplitude for deviant stimuli than for standard stimuli, thereby supporting the primarily automatic nature of MMN.

Limitations

Since we recorded ERPs from a limited number of electrodes, we were unable to perform a source-level analysis or current source density analysis. However, N140 and P300 have both been extensively examined and an abundant amount of information has been obtained on their generators (Allison et al., 1992; Tarkka et al., 1996; Valeriani et al., 2000, 2001; Tanaka et al., 2008). Therefore, it is possible to speculate about the brain regions involved in the modulation of ERP components. Furthermore, we did not perform the divided attention condition when the hands were closely located or when the forearms were crossed. These conditions may provide insights into the interaction between different spatial coordinates in somatosensory attention. Another limitation is the small sample size; therefore, we performed a post-hoc power analysis (Supplementary Material). The results obtained showed that most of the significant differences observed in ANOVAs, MANOVAs, and *t*-tests remained and, thus, our suggestions are effective. However, some analyzes showed low statistical power, suggesting higher type II error in some results, which was at least partly due to the small sample size. Therefore, further studies are warranted.

Conclusion

In conclusion, the present study demonstrated that the pattern of the somatosensory-attention effect on ERPs during focused and divided attention was similar to that in vision and audition. Therefore, somatosensory attention may be split between the hands, but follows some delay in modality-non-specific late processing by dividing resources between the hands depending on task demands (mental or

motor). The present results in the somatosensory modality support the classical resource allocation view of the P300 amplitude in both motor response and mental tasks and also provides additional evidence for s-MMN. The effect of somatosensory-spatial attention reflected by N140 amplitude may be affected by crossed forearms, but not by closely-placed hands, whereas modality-non-specific late processing interfered uniformly with both. A combined measure of neural (P300) and behavioral (RT) response times revealed specific changes in stimulus–response coupling and hand preference. Therefore, hand position and arm posture differently affected the attentional modulation of somatosensory processing at different stages as well as stimulus-response coupling.

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.

Ethics statement

The studies involving humans were approved by Ethics Committee, Graduate School of Comprehensive Human Sciences at the University of Tsukuba. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

TKi, TKa, and YN contributed to conception and design of the study. TKi and TKa performed the experiment and analysis. TKi wrote the first draft of the manuscript. All authors contributed to the article and approved the submitted version.

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.2023.1252686/full#supplementary-material

References

Adler, J., Giabbiconi, C. M., and Muller, M. M. (2009). Shift of attention to the body location of distracters is mediated by perceptual load in sustained somatosensory attention. *Biol. Psychol.* 81, 77–85. doi: 10.1016/j.biopsycho.2009.02.001

Akaiwa, M., Iwata, K., Saito, H., Shibata, E., Sasaki, T., and Sugawara, K. (2022). The effect of pedaling at different cadence on attentional resources. *Front. Hum. Neurosci.* 16:819232. doi: 10.3389/fnhum.2022.819232

Alho, K., Donauer, N., Paavilainen, P., Reinikainen, K., Sams, M., and Naatanen, R. (1987). Stimulus selection during auditory spatial attention as expressed by event-related potentials. *Biol. Psychol.* 24, 153–162. doi: 10.1016/0301-0511(87)90022-6

Allison, T., McCarthy, G., and Wood, C. C. (1992). The relationship between human long-latency somatosensory evoked potentials recorded from the cortical surface and from the scalp. *Electroencephalogr. Clin. Neurophysiol.* 84, 301–314. doi: 10.1016/0168-5597(92)90082-M

Beste, C., Konrad, C., Saft, C., Ukas, T., Andrich, J., Pfleiderer, B., et al. (2009). Alterations in voluntary movement execution in Huntington's disease are related to the dominant motor system: evidence from event-related potentials. *Exp. Neurol.* 216, 148–157. doi: 10.1016/j.expneurol.2008.11.018

Chen, J. C., Hammerer, D., D'Ostilio, K., Casula, E. P., Marshall, L., Tsai, C. H., et al. (2014). Bi-directional modulation of somatosensory mismatch negativity with transcranial direct current stimulation: an event related potential study. *J. Physiol.* 592, 745–757. doi: 10.1113/jphysiol.2013.260331

Christensen, C. A., Ford, J. M., and Pfefferbaum, A. (1996). The effect of stimulus-response incompatibility on P3 latency depends on the task but not on age. *Biol. Psychol.* 44, 121–141. doi: 10.1016/0301-0511(96)05203-9

Craig, J. C. (1985). Attending to two fingers: two hands are better than one. *Percept. Psychophys.* 38, 496–511. doi: 10.3758/BF03207059

Desmedt, J. E., Bourguet, M., Nguyen Tran, H., and Delacuvellerie, M. (1984). The P40 and P100 processing positivities that precede P300 closure in serial somatosensory decision tasks. *Ann. N. Y. Acad. Sci.* 425, 188–193. doi: 10.1111/j.1749-6632.1984.tb23531.x

Desmedt, J. E., and Robertson, D. (1977). Differential enhancement of early and late components of the cerebral somatosensory evoked potentials during forced-paced cognitive tasks in man. *J. Physiol.* 271, 761–782. doi: 10.1113/jphysiol.1977.sp012025

Desmedt, J. E., and Tomberg, C. (1989). Mapping early somatosensory evoked potentials in selective attention: critical evaluation of control conditions used for titrating by difference the cognitive P30, P40, P100 and N140. *Electroencephalogr. Clin. Neurophysiol.* 74, 321–346. doi: 10.1016/0168-5597(89)90001-4

Desmedt, J. E., and Tomberg, C. (1991). The search for 'neutral' conditions for recording control event-related potentials in order to assess cognitive components to both irrelevant and relevant stimuli: evidence for short-latency cognitive somatosensory effects. *Electroencephalogr. Clin. Neurophysiol.* 42, 210–221.

Doucet, C., and Stelmack, R. M. (1999). The effect of response execution on P3 latency, reaction time, and movement time. *Psychophysiology* 36, 351–363. doi: 10.1017/S0048577299980563

Downing, C. J., and Pinker, S. (1985). "The spatial structure of visual attention" in *Attention and performance XI*. eds. M. I. Posner and O. S. Marin (Hillsdale: Erlbaum), 171–187.

Eimer, M., Cockburn, D., Smedley, B., and Driver, J. (2001). Cross-modal links in endogenous spatial attention are mediated by common external locations: evidence from event-related brain potentials. *Exp. Brain Res.* 139, 398–411. doi: 10.1007/s002210100773

Eimer, M., and Driver, J. (2000). An event-related brain potential study of cross-modal links in spatial attention between vision and touch. *Psychophysiology* 37, 697–705. doi: 10.1111/1469-8986.3750697

Eimer, M., and Forster, B. (2003a). Modulations of early somatosensory ERP components by transient and sustained spatial attention. *Exp. Brain Res.* 151, 24–31. doi: 10.1007/s00221-003-1437-1

Eimer, M., and Forster, B. (2003b). The spatial distribution of attentional selectivity in touch: evidence from somatosensory ERP components. *Clin. Neurophysiol.* 114, 1298–1306. doi: 10.1016/S1388-2457(03)00107-X

Eimer, M., Forster, B., Fieger, A., and Harbich, S. (2004). Effects of hand posture on preparatory control processes and sensory modulations in tactile-spatial attention. *Clin. Neurophysiol.* 115, 596–608. doi: 10.1016/j.clinph.2003.10.015

Eimer, M., Forster, B., and Van Velzen, J. (2003a). Anterior and posterior attentional control systems use different spatial reference frames: ERP evidence from covert tactile-spatial orienting. *Psychophysiology* 40, 924–933. doi: 10.1111/1469-8986.00110

Eimer, M., van Velzen, J., and Driver, J. (2002). Cross-modal interactions between audition, touch, and vision in endogenous spatial attention: ERP evidence on preparatory states and sensory modulations. *J. Cogn. Neurosci.* 14, 254–271. doi: 10.1162/089892902317236885

Eimer, M., van Velzen, J., Forster, B., and Driver, J. (2003b). Shifts of attention in light and in darkness: an ERP study of supramodal attentional control and crossmodal links in spatial attention. *Brain Res. Cogn. Brain Res.* 15, 308–323. doi: 10.1016/S0926-6410(02)00203-3

Evans, P. M., Craig, J. C., and Rinker, M. A. (1992). Perceptual processing of adjacent and nonadjacent tactile nontargets. *Percept. Psychophys.* 52, 571–581. doi: 10.3758/BF03206719

Finlayson, M. A., and Reitan, R. M. (1976). Handedness in relation to measures of motor and tactile-perceptual functions in normal children. *Percept. Mot. Skills* 42, 475–481. doi: 10.2466/pms.1976.43.2.475

Forster, B., and Eimer, M. (2004). The attentional selection of spatial and non-spatial attributes in touch: ERP evidence for parallel and independent processes. *Biol. Psychol.* 66, 1–20. doi: 10.1016/j.biopsycho.2003.08.001

Forster, B., and Eimer, M. (2005). Covert attention in touch: behavioral and ERP evidence for costs and benefits. *Psychophysiology* 42, 171–179. doi: 10.1111/j.1469-8986.2005.00268.x

Garcia-Larrea, L., Lukaszewicz, A. C., and Mauguiere, F. (1995). Somatosensory responses during selective spatial attention: the N120-to-N140 transition. *Psychophysiology* 32, 526–537. doi: 10.1111/j.1469-8986.1995.tb01229.x

Gherri, E., and Eimer, M. (2008). Links between eye movement preparation and the attentional processing of tactile events: an event-related brain potential study. *Clin. Neurophysiol.* 119, 2587–2597. doi: 10.1016/j.clinph.2008.07.214

Gherri, E., and Eimer, M. (2010). Manual response preparation disrupts spatial attention: an electrophysiological investigation of links between action and attention. *Neuropsychologia* 48, 961–969. doi: 10.1016/j.neuropsychologia.2009.11.017

Gherri, E., and Forster, B. (2012a). Crossing the hands disrupts tactile spatial attention but not motor attention: evidence from event-related potentials. *Neuropsychologia* 50, 2303–2316. doi: 10.1016/j.neuropsychologia.2012.05.034

Gherri, E., and Forster, B. (2012b). The orienting of attention during eye and hand movements: ERP evidence for similar frame of reference but different spatially specific modulations of tactile processing. *Biol. Psychol.* 91, 172–184. doi: 10.1016/j.biopsycho.2012.06.007

Gherri, E., White, F., and Venables, E. (2023). On the spread of spatial attention in touch: evidence from event-related brain potentials. *Biol. Psychol.* 178:108544. doi: 10.1016/j.biopsycho.2023.108544

Gillmeister, H., and Forster, B. (2012). Hands behind your back: effects of arm posture on tactile attention in the space behind the body. *Exp. Brain Res.* 216, 489–497. doi: 10.1007/s00221-011-2953-z

Gonsalvez, C. J., Barry, R. J., Rushby, J. A., and Polich, J. (2007). Target-to-target interval, intensity, and P300 from an auditory single-stimulus task. *Psychophysiology* 44, 245–250. doi: 10.1111/j.1469-8986.2007.00495.x

Gonsalvez, C. L., and Polich, J. (2002). P300 amplitude is determined by target-to-target interval. *Psychophysiology* 39, 388–396. doi: 10.1017/S0048577201393137

He, X., Zhang, J., Zhang, Z., Go, R., Wu, J., Li, C., et al. (2020). Effects of visual attentional load on the tactile sensory memory indexed by somatosensory mismatch negativity. *Front. Neuroinform.* 14:575078. doi: 10.3389/fninf.2020.575078

Heed, T., and Roder, B. (2010). Common anatomical and external coding for hands and feet in tactile attention: evidence from event-related potentials. *J. Cogn. Neurosci.* 22, 184–202. doi: 10.1162/jocn.2008.21168

Heinze, H. J., Luck, S. J., Munte, T. F., Gos, A., Mangun, G. R., and Hillyard, S. A. (1994). Attention to adjacent and separate positions in space: an electrophysiological analysis. *Percept. Psychophys.* 56, 42–52. doi: 10.3758/BF03211689

Hillyard, S. A., and Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proc. Natl. Acad. Sci. U. S. A.* 95, 781–787. doi: 10.1073/pnas.95.3.781

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

Hillyard, S. A., Vogel, E. K., and Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 353, 1257–1270. doi: 10.1098/rstb.1998.0281

Hink, R. F., Fenton, W. H. Jr., Pfefferbaum, A., Tinklenberg, J. R., and Kopell, B. S. (1978). The distribution of attention across auditory input channels: an assessment using the human evoked potential. *Psychophysiology* 15, 466–473. doi: 10.1111/j.1469-8986.1978.tb01417.x

Hink, R. F., Van Voorhis, S. T., Hillyard, S. A., and Smith, T. S. (1977). The division of attention and the human auditory evoked potential. *Neuropsychologia* 15, 597–605. doi: 10.1016/0028-3932(77)90065-3

Inui, K., Tran, T. D., Qiu, Y., Wang, X., Hoshiyama, M., and Kakigi, R. (2003). A comparative magnetoencephalographic study of cortical activations evoked by noxious and innocuous somatosensory stimulations. *Neuroscience* 120, 235–248. doi: 10.1016/S0306-4522(03)00261-6

Johannes, S., Munte, T. F., Heinze, H. J., and Mangun, G. R. (1995). Luminance and spatial attention effects on early visual processing. *Brain Res. Cogn. Brain Res.* 2, 189–205. doi: 10.1016/0926-6410(95)90008-X

Johnson, R. (1986). A triarchic model of P300 amplitude. *Psychophysiology* 23, 367–384. doi: 10.1111/j.1469-8986.1986.tb00649.x

Josiassen, R. C., Shagass, C., Roemer, R. A., Ercegovac, D. V., and Straumanis, J. J. (1982). Somatosensory evoked potential changes with a selective attention task. *Psychophysiology* 19, 146–159. doi: 10.1111/j.1469-8986.1982.tb02536.x

Keil, J., Pomper, U., Feuerbach, N., and Senkowski, D. (2017). Temporal orienting precedes intersensory attention and has opposing effects on early evoked brain activity. *NeuroImage* 148, 230–239. doi: 10.1016/j.neuroimage.2017.01.039

Kekoni, J., Hamalainen, H., Saarinen, M., Grohn, J., Reinikainen, K., Lehtokoski, A., et al. (1997). Rate effect and mismatch responses in the somatosensory system: ERP-recordings in humans. *Biol. Psychol.* 46, 125–142. doi: 10.1016/S0301-0511(97)05249-6

Kennett, S., Eimer, M., Spence, C., and Driver, J. (2001). Tactile-visual links in exogenous spatial attention under different postures: convergent evidence from psychophysics and ERPs. *J. Cogn. Neurosci.* 13, 462–478. doi: 10.1162/08989290152001899

Kida, T., Kaneda, T., and Nishihira, Y. (2012a). Dual-task repetition alters event-related brain potentials and task performance. *Clin. Neurophysiol.* 123, 1123–1130. doi: 10.1016/j.clinph.2011.10.001

Kida, T., Kaneda, T., and Nishihira, Y. (2012b). Modulation of somatosensory processing in dual tasks: an event-related brain potential study. *Exp. Brain Res.* 216, 575–584. doi: 10.1007/s00221-011-2961-z

Kida, T., Nishihira, Y., Hatta, A., and Wasaka, T. (2003a). Somatosensory N250 and P300 during discrimination tasks. *Int. J. Psychophysiol.* 48, 275–283. doi: 10.1016/S0167-8760(03)00021-7

Kida, T., Nishihira, Y., Hatta, A., Wasaka, T., Nakata, H., and Sakamoto, M. (2003b). Stimulus context affects P300 and reaction time during a somatosensory discrimination task. *Adv. Exerc. Sports Physiol.* 9, 105–110.

Kida, T., Nishihira, Y., Hatta, A., Wasaka, T., Nakata, H., Sakamoto, M., et al. (2003c). Changes in the somatosensory N250 and P300 by the variation of reaction time. *Eur. J. Appl. Physiol.* 89, 326–330. doi: 10.1007/s00421-003-0801-y

Kida, T., Nishihira, Y., Hatta, A., Wasaka, T., Tazoe, T., Sakajiri, Y., et al. (2004a). Resource allocation and somatosensory P300 amplitude during dual task: effects of tracking speed and predictability of tracking direction. *Clin. Neurophysiol.* 115, 2616–2628. doi: 10.1016/j.clinph.2004.06.013

Kida, T., Nishihira, Y., Wasaka, T., Nakata, H., and Sakamoto, M. (2004b). Differential modulation of temporal and frontal components of the somatosensory N140 and the effect of interstimulus interval in a selective attention task. *Brain Res. Cogn. Brain Res.* 19, 33–39. doi: 10.1016/j.cogbrainres.2003.10.016

Kida, T., Nishihira, Y., Wasaka, T., Nakata, H., and Sakamoto, M. (2004c). Passive enhancement of the somatosensory P100 and N140 in an active attention task using deviant alone condition. *Clin. Neurophysiol.* 115, 871–879. doi: 10.1016/j.clinph.2003.11.037

Kida, T., Tanaka, E., and Kakigi, R. (2018). Adaptive flexibility of the within-hand attentional gradient in touch: an MEG study. *NeuroImage* 179, 373–384. doi: 10.1016/j.neuroimage.2018.06.063

Kida, T., Wasaka, T., Nakata, H., Akatsuka, K., and Kakigi, R. (2006). Active attention modulates passive attention-related neural responses to sudden somatosensory input against a silent background. *Exp. Brain Res.* 175, 609–617. doi: 10.1007/s00221-006-0578-4

Kimura, M. (2012). Visual mismatch negativity and unintentional temporal-context-based prediction in vision. *Int. J. Psychophysiol.* 83, 144–155. doi: 10.1016/j.ijpsycho.2011.11.010

Kok, A. (1997). Event-related-potential (ERP) reflections of mental resources: a review and synthesis. *Biol. Psychol.* 45, 19–56. doi: 10.1016/S0301-0511(96)05221-0

Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology* 38, 557–577. doi: 10.1017/S0048577201990559

Kramer, A., Schneider, W., Fisk, A., and Donchin, E. (1986). The effects of practice and task structure on components of the event-related brain potential. *Psychophysiology* 23, 33–47. doi: 10.1111/j.1469-8986.1986.tb00590.x

Kramer, A. F., Wickens, C. D., and Donchin, E. (1985). Processing of stimulus properties: evidence for dual-task integrality. *J. Exp. Psychol. Hum. Percept. Perform.* 11, 393–408. doi: 10.1037/0096-1523.11.4.393

Kutas, M., McCarthy, G., and Donchin, E. (1977). Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. *Science* 197, 792–795. doi: 10.1126/science.887923

LaBerge, D. (1983). Spatial extent of attention to letters and words. *J. Exp. Psychol. Hum. Percept. Perform.* 9, 371–379. doi: 10.1037/0096-1523.9.3.371

Lakatos, S., and Shepard, R. N. (1997). Time-distance relations in shifting attention between locations on one's body. *Percept. Psychophys.* 59, 557–566. doi: 10.3758/BF03211864

Leocani, L., Toro, C., Zhuang, P., Gerloff, C., and Hallett, M. (2001). Event-related desynchronization in reaction time paradigms: a comparison with event-related potentials and corticospinal excitability. *Clin. Neurophysiol.* 112, 923–930. doi: 10.1016/S1388-2457(01)00530-2

Leuthold, H., and Sommer, W. (1998). Postperceptual effects and P300 latency. *Psychophysiology* 35, 34–46. doi: 10.1111/1469-8986.3510034

Magliero, A., Bashore, T. R., Coles, M. G., and Donchin, E. (1984). On the dependence of P300 latency on stimulus evaluation processes. *Psychophysiology* 21, 171–186. doi: 10.1111/j.1469-8986.1984.tb00201.x

Mangun, G. R., and Hillyard, S. A. (1988). Spatial gradients of visual attention: behavioral and electrophysiological evidence. *Electroencephalogr. Clin. Neurophysiol.* 70, 417–428. doi: 10.1016/0013-4694(88)90019-3

McCarthy, G., and Donchin, E. (1981). A metric for thought: A comparison of P300 latency and reaction time. *Science* 211, 77–80. doi: 10.1126/science.7444452

Michie, P. T., Bearpark, H. M., Crawford, J. M., and Glue, L. C. (1987). The effects of spatial selective attention on the somatosensory event-related potential. *Psychophysiology* 24, 449–463. doi: 10.1111/j.1469-8986.1987.tb00316.x

Mondor, T. A., and Zatorre, R. J. (1995). Shifting and focusing auditory spatial attention. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 387–409. doi: 10.1037/0096-1523.21.2.387

Munoz, F., Reales, J. M., Sebastian, M. A., and Ballesteros, S. (2014). An electrophysiological study of haptic roughness: effects of levels of texture and stimulus uncertainty in the P300. *Brain Res.* 1562, 59–68. doi: 10.1016/j.brainres.2014.03.013

Naatanen, R. (1992) *Attention and brain function*. Erlbaum, Hillsdale, NJ.

Naatanen, R. (2000). SPR award, 1999. For distinguished contributions to psychophysiology: Steven a Hillyard. *Psychophysiology* 37, 269–274. doi: 10.1111/1469-8986.3730269

Naatanen, R., Gaillard, A. W., and Mantysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychol.* 42, 313–329. doi: 10.1016/0001-6918(78)90006-9

Naatanen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* 118, 2544–2590. doi: 10.1016/j.clinph.2007.04.026

Novicic, M., and Savic, A. M. (2023). Somatosensory event-related potential as an electrophysiological correlate of endogenous spatial tactile attention: prospects for Electrotactile brain-computer Interface for sensory training. *Brain Sci.* 13:766. doi: 10.3390/brainsci13050766

Parasuraman, R. (1978). Auditory evoked potentials and divided attention. *Psychophysiology* 15, 460–465. doi: 10.1111/j.1469-8986.1978.tb01416.x

Pei, L., Longcamp, M., Leung, F. K., and Ouyang, G. (2021). Temporally resolved neural dynamics underlying handwriting. *NeuroImage* 244:118578. doi: 10.1016/j.neuroimage.2021.118578

Pfefferbaum, A., Christensen, C., Ford, J. M., and Kopell, B. S. (1986). Apparent response incompatibility effects on P3 latency depend on the task. *Electroencephalogr. Clin. Neurophysiol.* 64, 424–437. doi: 10.1016/0013-4694(86)90076-3

Pfefferbaum, A., Ford, J., Johnson, R. Jr., Wenegrat, B., and Kopell, B. S. (1983). Manipulation of P3 latency: speed vs. accuracy instructions. *Electroencephalogr. Clin. Neurophysiol.* 55, 188–197. doi: 10.1016/0013-4694(83)90187-6

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148. doi: 10.1016/j.clinph.2007.04.019

Polich, J. (2020). 50+ years of P300: where are we now? *Psychophysiology* 57:e13616. doi: 10.1111/psyp.13616

Polich, J., Aung, M., and Dalessio, D. J. (1988). Long latency auditory evoked potentials: intensity, inter-stimulus interval, and habituation. *Pavlov. J. Biol. Sci.* 23, 35–40. doi: 10.1007/BF02910543

Potgieser, A. R., van der Hoorn, A., and de Jong, B. M. (2015). Cerebral activations related to writing and drawing with each hand. *PLoS One* 10:e0126723. doi: 10.1371/journal.pone.0126723

Press, C., Heyes, C., Haggard, P., and Eimer, M. (2008). Visuotactile learning and body representation: an ERP study with rubber hands and rubber objects. *J. Cogn. Neurosci.* 20, 312–323. doi: 10.1162/jocn.2008.20022

Ragot, R. (1984). Perceptual and motor space representation: an event-related potential study. *Psychophysiology* 21, 159–170. doi: 10.1111/j.1469-8986.1984.tb00199.x

Ragot, R., and Lesevre, N. (1986). Electrophysiological study of intrahemispheric S-R compatibility effects elicited by visual directional cues. *Psychophysiology* 23, 19–27. doi: 10.1111/j.1469-8986.1986.tb00586.x

Ragot, R., and Renault, B. (1981). P300, as a function of S-R compatibility and motor programming. *Biol. Psychol.* 13, 289–294. doi: 10.1016/0301-0511(81)90044-2

Ragot, R., and Renault, B. (1985). P300 and S-R compatibility: a reply to Magliero et al. *Psychophysiology* 22, 349–352. doi: 10.1111/j.1469-8986.1985.tb01614.x

Restuccia, D., Della Marca, G., Valeriani, M., Leggio, M. G., and Molinari, M. (2007). Cerebellar damage impairs detection of somatosensory input changes. A somatosensory mismatch-negativity study. *Brain* 130, 276–287. doi: 10.1093/brain/awl236

Restuccia, D., Zanini, S., Cazzagon, M., Del Piero, I., Martucci, L., and Della Marca, G. (2009). Somatosensory mismatch negativity in healthy children. *Dev. Med. Child Neurol.* 51, 991–998. doi: 10.1111/j.1469-8749.2009.03367.x

Reuter, E. M., Voelcker-Rehage, C., Vieluf, S., and Godde, B. (2014). Effects of age and expertise on tactile learning in humans. *Eur. J. Neurosci.* 40, 2589–2599. doi: 10.1111/ejn.12629

Reuter, E. M., Voelcker-Rehage, C., Vieluf, S., Winneke, A. H., and Godde, B. (2013). A parietal-to-frontal shift in the P300 is associated with compensation of tactile discrimination deficits in late middle-aged adults. *Psychophysiology* 50, 583–593. doi: 10.1111/psyp.12037

Rinker, M. A., and Craig, J. C. (1994). The effect of spatial orientation on the perception of moving tactile stimuli. *Percept. Psychophys.* 56, 356–362. doi: 10.3758/BF03209769

Rorden, C., and Driver, J. (2001). Spatial deployment of attention within and across hemifields in an auditory task. *Exp. Brain Res.* 137, 487–496. doi: 10.1007/s002210100679

Ruchkin, D. S., Sutton, S., and Mahaffey, D. (1987). Functional differences between members of the P300 complex: P3e and P3b. *Psychophysiology* 24, 87–103. doi: 10.1111/j.1469-8986.1987.tb01867.x

Savic, A. M., Novicic, M., Ethordevic, O., Konstantinovic, L., and Miler-Jerkovic, V. (2023). Novel electrotactile brain-computer interface with somatosensory event-related potential based control. *Front. Hum. Neurosci.* 17:1096814. doi: 10.3389/fnhum.2023.1096814

Shen, G., Smyk, N. J., Meltzoff, A. N., and Marshall, P. J. (2018a). Using somatosensory mismatch responses as a window into somatotopic processing of tactile stimulation. *Psychophysiology* 55:e13030. doi: 10.1111/psyp.13030

Shen, G., Weiss, S. M., Meltzoff, A. N., and Marshall, P. J. (2018b). The somatosensory mismatch negativity as a window into body representations in infancy. *Int. J. Psychophysiol.* 134, 144–150. doi: 10.1016/j.ijpsycho.2018.10.013

Shulman, G. L., Sheehy, J. B., and Wilson, J. (1986). Gradients of spatial attention. *Acta Psychol.* 61, 167–181. doi: 10.1016/0001-6918(86)90029-6

Shulman, G. L., Wilson, J., and Sheehy, J. B. (1985). Spatial determinants of the distribution of attention. *Percept. Psychophys.* 37, 59–65. doi: 10.3758/BF03207139

Sirevaag, E. J., Kramer, A. F., Coles, M. G., and Donchin, E. (1989). Resource reciprocity: an event-related brain potentials analysis. *Acta Psychol.* 70, 77–97. doi: 10.1016/0001-6918(89)90061-9

Stefanics, G., Kremlacek, J., and Czigler, I. (2014). Visual mismatch negativity: a predictive coding view. *Front. Hum. Neurosci.* 8:666. doi: 10.3389/fnhum.2014.00666

Sutton, S., Braren, M., Zubin, J., and John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science* 150, 1187–1188. doi: 10.1126/science.150.3700.1187

Tanaka, E., Inui, K., Kida, T., Miyazaki, T., Takeshima, Y., and Kakigi, R. (2008). A transition from unimodal to multimodal activations in four sensory modalities in humans: an electrophysiological study. *BMC Neurosci.* 9:116. doi: 10.1186/1471-2202-9-116

Tarkka, I. M., Micheloyannis, S., and Stokic, D. S. (1996). Generators for human P300 elicited by somatosensory stimuli using multiple dipole source analysis. *Neuroscience* 75, 275–287. doi: 10.1016/0306-4522(96)00287-4

Teder-Salejarvi, W. A., and Hillyard, S. A. (1998). The gradient of spatial auditory attention in free field: an event-related potential study. *Percept. Psychophys.* 60, 1228–1242. doi: 10.3758/BF03206172

Teder-Salejarvi, W. A., Hillyard, S. A., Roder, B., and Neville, H. J. (1999). Spatial attention to central and peripheral auditory stimuli as indexed by event-related potentials. *Brain Res. Cogn. Brain Res.* 8, 213–227. doi: 10.1016/S0926-6410(99)00023-3

Valeriani, M., Le Pera, D., Niddam, D., Arendt-Nielsen, L., and Chen, A. C. (2000). Dipolar source modeling of somatosensory evoked potentials to painful and nonpainful median nerve stimulation. *Muscle Nerve* 23, 1194–1203. doi: 10.1002/1097-4598(200008)23:8<1194::AID-MUS6>3.0.CO;2-E

Valeriani, M., Le Pera, D., and Tonali, P. (2001). Characterizing somatosensory evoked potential sources with dipole models: advantages and limitations. *Muscle Nerve* 24, 325–339. doi: 10.1002/1097-4598(200103)24:3<325::AID-MUS1002>3.0.CO;2-0

Valeriani, M., Ranghi, F., and Giaquinto, S. (2003). The effects of aging on selective attention to touch: a reduced inhibitory control in elderly subjects? *Int. J. Psychophysiol.* 49, 75–87. doi: 10.1016/S0167-8760(03)00094-1

van Velzen, J., Forster, B., and Eimer, M. (2002). Temporal dynamics of lateralized ERP components elicited during endogenous attentional shifts to relevant tactile events. *Psychophysiology* 39, 874–878. doi: 10.1111/1469-8986.3960874

Waberski, T. D., Gobbele, R., Darvas, F., Schmitz, S., and Buchner, H. (2002). Spatiotemporal imaging of electrical activity related to attention to somatosensory stimulation. *NeuroImage* 17, 1347–1357. doi: 10.1006/nimg.2002.1222

Wickens, C. D. (1991). "Processing resources and attention" in *Multi-task performance*. ed. D. L. Damos (Milton Park Abingdon: Taylor and Francis), 3–34.

Wickens, C., Kramer, A., Vanasse, L., and Donchin, E. (1983). Performance of concurrent tasks: a psychophysiological analysis of the reciprocity of information-processing resources. *Science* 221, 1080–1082. doi: 10.1126/science.6879207

Wickens, C. D., and McCarley, J. S. (2008) *Applied attention theory*. CRC Press, Boca Raton.

Wijers, A. A., Lamain, W., Slopsema, J. S., Mulder, G., and Mulder, L. J. (1989). An electrophysiological investigation of the spatial distribution of attention to colored stimuli in focused and divided attention conditions. *Biol. Psychol.* 29, 213–245. doi: 10.1016/0301-0511(89)90021-5

Woody, C. D. (1967). Characterization of an adaptive filter for the analysis of variable latency neuroelectric signals. *Med. Biol. Eng. Comput.* 5, 539–554. doi: 10.1007/BF02474247

Yamashiro, K., Inui, K., Otsuru, N., Kida, T., Akatsuka, K., and Kakigi, R. (2008). Somatosensory off-response in humans: an ERP study. *Exp. Brain Res.* 190, 207–213. doi: 10.1007/s00221-008-1468-8

Yamashiro, K., Inui, K., Otsuru, N., Kida, T., and Kakigi, R. (2009). Somatosensory off-response in humans: an MEG study. *NeuroImage* 44, 1363–1368. doi: 10.1016/j.neuroimage.2008.11.003

frontiers | Frontiers in Human Neuroscience

TYPE Original Research
PUBLISHED 29 December 2023
DOI 10.3389/fnhum.2023.1160392

OPEN ACCESS

EDITED BY
Tetsuo Kida,
Aichi Developmental Disability Center,
Institute for Developmental Research, Japan

REVIEWED BY
Swethasri Padmanjani Dravida,
Massachusetts Institute of Technology,
United States
Noman Naseer,
Air University, Pakistan
Erol Yildirim,
Istanbul Medipol University, Türkiye

*CORRESPONDENCE
Ippeita Dan
✉ dan@brain-lab.jp

†These authors have contributed equally to this work

RECEIVED 07 February 2023
ACCEPTED 28 September 2023
PUBLISHED 29 December 2023

CITATION
Ohshima S, Koeda M, Kawai W, Saito H, Niioka K, Okuno K, Naganawa S, Hama T, Kyutoku Y and Dan I (2023) Cerebral response to emotional working memory based on vocal cues: an fNIRS study.
Front. Hum. Neurosci. 17:1160392.
doi: 10.3389/fnhum.2023.1160392

Cerebral response to emotional working memory based on vocal cues: an fNIRS study

Saori Ohshima[1†], Michihiko Koeda[2,3†], Wakana Kawai[1], Hikaru Saito[1], Kiyomitsu Niioka[1], Koki Okuno[1], Sho Naganawa[1], Tomoko Hama[4,5], Yasushi Kyutoku[1] and Ippeita Dan[1*]

[1]Applied Cognitive Neuroscience Laboratory, Faculty of Science and Engineering, Chuo University, Bunkyo, Japan, [2]Department of Neuropsychiatry, Graduate School of Medicine, Nippon Medical School, Bunkyo, Japan, [3]Department of Mental Health, Nippon Medical School Tama Nagayama Hospital, Tama, Japan, [4]Department of Medical Technology, Ehime Prefectural University of Health Sciences, Iyo-gun, Japan, [5]Department of Clinical Laboratory Medicine, Faculty of Health Science Technology, Bunkyo Gakuin University, Tokyo, Japan

Introduction: Humans mainly utilize visual and auditory information as a cue to infer others' emotions. Previous neuroimaging studies have shown the neural basis of memory processing based on facial expression, but few studies have examined it based on vocal cues. Thus, we aimed to investigate brain regions associated with emotional judgment based on vocal cues using an N-back task paradigm.

Methods: Thirty participants performed N-back tasks requiring them to judge emotion or gender from voices that contained both emotion and gender information. During these tasks, cerebral hemodynamic response was measured using functional near-infrared spectroscopy (fNIRS).

Results: The results revealed that during the Emotion 2-back task there was significant activation in the frontal area, including the right precentral and inferior frontal gyri, possibly reflecting the function of an attentional network with auditory top-down processing. In addition, there was significant activation in the ventrolateral prefrontal cortex, which is known to be a major part of the working memory center.

Discussion: These results suggest that, compared to judging the gender of voice stimuli, when judging emotional information, attention is directed more deeply and demands for higher-order cognition, including working memory, are greater. We have revealed for the first time the specific neural basis for emotional judgments based on vocal cues compared to that for gender judgments based on vocal cues.

KEYWORDS

emotional judgment, voice, n-back, working memory, functional near-infrared spectroscopy, VLPFC, dorsal attention network

1 Introduction

When we communicate with others, we mainly use visual and auditory information as a cue to infer the emotions of others. Visual information includes facial expressions, gestures, eye contact, and distance from others, while auditory information includes the voices of others (Gregersen, 2005; Esposito et al., 2009). In general, humans are more highly evolved to detect and process visual information from faces than are other species. It has been shown that we prefer facial information or face-like patterns to other visual information or nonface-like patterns and look at such patterns for longer periods of time (Morton and Johnson, 1991; Valenza et al., 1996). Thus, facial expressions are considered to be major cues for inferring the emotions of others (Neta et al., 2011; Neta and Whalen, 2011).

On the other hand, it has been pointed out that the effectiveness of voice for inferring emotions cannot be ignored because it accurately reflects people's intended emotions (Johnstone and Scherer, 2000). For example, we can interpret that the speaker is happy when hearing a bright, high-pitched voice, or that the speaker is afraid when hearing a screeching, high-pitched voice. There are several studies focusing on the influence of voice on the inference of emotions (De Gelder and Vroomen, 2000; Collignon et al., 2008). De Gelder and Vroomen (2000) showed interaction between visual and auditory information while inferring emotions, using voices and pictures expressing happy or sad emotions. They confirmed that inferences of emotions depicted in pictures were more accurate when the emotions in the pictures matched those of the vocal stimuli than when they did not match. Collignon et al. (2008) prepared voice and facial expression stimuli depicting two emotions, fear and disgust. Then, they presented them alone or in combination, and asked participants to infer the emotions depicted. Participants inferred emotions more quickly and more accurately when voice and facial expressions containing congruent emotions were presented than when either one of them was presented alone. Thus, the voice also works as a complementary cue for inferring other people's emotions, and it is believed that accurate and faster decisions can be made when auditory information is added to visual information.

Recent studies on the physiological mechanisms behind inferring emotions by focusing on vocal stimuli have been conducted using functional magnetic resonance imaging (fMRI). For example, Koeda et al. (2013a) used fMRI to examine the brain regions associated with emotion inference. They compared the brain activation during the inference of emotions in healthy subjects with that of schizophrenic patients. In general, schizophrenic patients have more difficulty inferring emotions than do healthy subjects, which leads to difficulty in communicating (Bucci et al., 2008; Galderisi et al., 2013). The voice stimuli in their study were presented in the form of greetings such as "hello" and "good morning" uttered by males or females with positive, negative, or no emotion. Participants were asked to infer the emotion or gender depicted in each voice stimulus. Greater activation of the left superior temporal gyrus in healthy subjects was observed compared to that of schizophrenic patients. Thus, they concluded that the left superior temporal gyrus is relevant when inferring emotions from voices. Koeda et al. (2013a) study focused mainly on the aspects of semantic processing of emotions. In other words, they focused on the discrimination of specific emotions.

Moreover, other research focusing on the aspect of cognitive mechanisms related to the inference of emotions has also been conducted using facial expressions as visual stimuli (Neta and Whalen, 2011). Models for face perception assume a difference between the neural circuits that support the perception of changeable facial features (e.g., emotional expression, gaze) and of invariant facial features (e.g., facial structure, identity) (Haxby et al., 2000; Calder and Young, 2005). For example, within the core system of the cognitive processing of faces, it has been shown that changeable features elicited greater activation in the superior temporal sulcus, and invariant features elicited greater activation in the lateral spindle gyrus (Haxby et al., 2000). Furthermore, in conjunction with this core system, it has been shown that activation in other regions that extract relevant meanings from faces (e.g., amygdala/insula for emotional information, intraparietal sulcus for spatial attention, and auditory cortex for speech) was also observed (Haxby et al., 2000).

To understand the processing in inference of emotions, one study asked subjects to retain information about faces in working memory (Neta and Whalen, 2011). Working memory is an essential component of many cognitive operations, from complex decision making to selective attention (Baddeley and Hitch, 1974; Baddeley, 1998). Working memory is commonly examined using an N-back task, in which participants are asked to judge whether the current stimulus matches the stimulus presented N-stimuli before. Importantly, the N-back task has been used in some studies on face processing (Hoffman and Haxby, 2000; Braver et al., 2001; Leibenluft et al., 2004; Gobbini and Haxby, 2007; Weiner and Grill-Spector, 2010).

The N-back task has also been used to identify neuroanatomical networks, respectively, related to changeable facial features (e.g., emotional expression) and invariant facial features (e.g., identity) (e.g., Haxby et al., 2000; Calder and Young, 2005). LoPresti et al. (2008) confirmed that the sustained activation of the left orbitofrontal cortex was greater during an emotion N-back task than during an identity N-back task using faces as stimuli. In addition, transient activations of the temporal and occipital cortices, including the right inferior occipital cortex, were greater during the identity N-back task. On the other hand, those of the right superior temporal sulcus and posterior parahippocampal cortex were greater during the emotion N-back task.

Based on these findings, some studies have focused on the neural processing of the inference of emotions from facial expressions, which are changeable features of faces. Neta and Whalen (2011) used fMRI to evaluate brain activation during two types of 2-back tasks, using the standard design. During the tasks, participants were required to judge whether the emotion or gender depicted by the current stimulus was the same as that of the stimulus presented 2 stimuli back. The results showed that the activations in the right posterior superior temporal sulcus and the bilateral inferior frontal gyrus during the Emotion 2-back task were significantly greater than those during the Identity 2-back task. In contrast, the rostral/ventral anterior cingulate cortex, bilateral precuneus, and right temporoparietal junction were significantly more active during the Identity 2-back task than during the Emotion 2-back task. Moreover, participants who exhibited greater activation in the DLPFC during both tasks also exhibited greater activation in the amygdala during the Emotion 2-back task and in the lateral spindle gyrus during the Identity 2-back task than those who had less activation in the DLPFC. Interestingly, the activation levels of the DLPFC and amygdala/spindle gyrus were significantly correlated with behavioral performance factors, such as accuracy (ACC) and reaction time (RT), for each task. Based on these findings, Neta and Whalen (2011) postulated that the DLPFC acts as a part of the core system for working memory tasks in general.

On the other hand, it has been suggested that not only the DLPFC but also the VLPFC may relate to WM (Braver et al., 2001; Kostopoulos and Petrides, 2004). In particular, Braver et al. (2001) revealed that the right VLPFC worked selectively for processing non-verbal items such as unfamiliar faces. It should be noted that the bilateral VLPFC was recruited in the Emotion 2-back task as mentioned above (Neta and Whalen, 2011).

Although the neural bases of the cognitive process related to inference of emotions based on facial expression have become clearer, those for vocal cues have not to date been clarified despite their importance. In fact, often we are forced to infer the emotions of others correctly based only on auditory information, not visual information (e.g., during a telephone call). In the processing of inferred emotions

based on vocal cues, the DLPFC and VLPFC should be involved as parts of the core system of working memory, as with visual cues (Braver et al., 2001; Neta and Whalen, 2011). However, the findings the Braver et al. (2001) and Neta and Whalen (2011) studies are not directly generalizable due to differences in the sensory modalities. In other words, the inference of emotions based on vocal cues might involve cognitive components specific to auditory processing.

Therefore, the purpose of our study was to clarify the neural basis of processing related to inferring emotions based on vocal cues using the N-back task paradigm. We named one version of the N-back task, in which participants were asked to judge the emotion depicted by the stimuli, the Emotion task, and, the other version of the task, in which they were asked to judge the gender depicted, the Identity task. In addition, as did Neta and Whalen (2011), we examined whether greater activation in specific brain regions was associated with behavioral performance advantages while inferring emotions based on vocal cues.

Moreover, in order to examine differences in judgment strategies between the Emotion and Identity tasks, we also examined the correlations between the behavioral performance for each. For example, if it were shown that the faster the RT, the higher the percentage of correct answers (i.e., negative correlation), then it would be possible that responding quickly leads to better performance. In such a case, there would be no need for cognitive control to be applied to the automatized process. On the other hand, if it were shown that the slower the RT, the higher the percentage of correct responses (i.e., positive correlation), then careful responses may be associated with better performance. In such a case, the inhibitory function would be required, and the cognitive demand associated with recognition should increase.

We utilized functional near-infrared spectroscopy (fNIRS) to measure brain functions. Although fNIRS cannot measure the deep regions of the brain, it has advantages that fMRI does not. fNIRS can measure cerebral hemodynamic responses caused by brain activation relatively easily by simply attaching a probe to the subject's head. Notably, fNIRS measurement equipment is less constraining and is relatively quiet. It does not require a special measurement environment. fMRI, however, unavoidably creates a noisy environment, which is not conducive to processing emotion judgments based on vocal cues. Thus, fNIRS was judged to be the best method to achieve the purpose of this study.

2 Methods

2.1 Participants and ethics

Thirty right-handed, healthy volunteers (13 males and 17 females, mean age 21.83 ± 0.97 years, range 20–24) participated in this study. All participants were native Japanese speakers with no history of neurological, psychiatric, or cardiac disorders. They had normal or corrected-to-normal vision and normal color vision. Handedness was assessed by means of the Edinburgh Inventory (Oldfield, 1971). This study was approved by the institutional ethics committee of Chuo University, and the protocol was in accordance with the Declaration of Helsinki guidelines. Two participants were removed from the sample due to unavailability for complete experimental data. As a result, the final sample contained 28 participants.

2.2 Experimental procedure and stimulus

Participants performed an Emotion task to judge emotional information and an Identity task to determine gender information as in a previous study on facial expression (Neta and Whalen, 2011). We used two audio sets for the emotion condition. The first expressed specific emotional vocalizations by Canadian actors: "Anger," "Disgust," "Fear," "Pain," "Sadness," "Surprise," "Happiness," and "Pleasure" (Montreal Affective Voices: MAV) (Belin et al., 2008). The second was created using Japanese actors with the same eight emotions expressed. We named it Tokyo Affective Voices (TAV). Simple "ah" sounds were used as a control for the influence of lexical-semantic processing (Figure 1).

According to Koeda et al. (2013b), when Japanese participants listened to the MAVs, they recognized the positive emotion "Happy" and the opposite negative emotion "Angry" with high ACC. Therefore, we used these for the vocal stimuli; we selected two emotions, "Happy" and "Angry," expressed by both "Male" and "Female" voices from both MAV and TAV, resulting in eight vocal stimuli.

The task load consisted of 0-back, 1-back, and 2-back. In accordance with Nakao et al. (2011) and Li et al. (2014), we performed 0-back, 1-back, and 2-back in sequence, in order to avoid the effect of tension caused by a sudden increase in load, for example jumping from 0-back to 2-back. Each condition was repeated twice for a total of 12 blocks. First, participants performed an N-back task in which they were asked to retain either the emotion (emotional information) or the identity (gender information) in their working memory. Second, in the 0-back condition, participants were asked to respond to the target emotion or gender by pressing a particular key on a keyboard, and a different key for non-targets. Subsequently, the participants were asked to judge whether the current emotion or gender being presented was the same or not the same as the emotion or gender presented immediately before (1-back) or two trials (stimulus presentations) before (2-back) by pressing a key as in the 0-back condition.

We presented each emotion ("Happy" and "Angry") and each gender ("Male" and "Female") in each type of voice (MAV and TAV) the same number of times. Each block condition included 10 trials (4 target trials and 6 non-target trials) in random order, each being presented once per block for 2 s, with an inter-stimulus interval of 1.5 s, and a 15-s blank between conditions. Before the start of a new block, the word "0-back," "1-back," or "2-back" appeared in the center of the screen to indicate to the participant which task they should prepare to perform. The participants answered by pressing "F" or "J" on a keyboard with their left and right index fingers, respectively, to indicate whether the presented voices were target or non-target trials. We used E-prime 2.0 (Psychology Software Tools) to create these tasks. ACC and RT were obtained as behavioral measures for each trial.

Conventionally, when the distance from the speaker to the participant's ear is 1 meter, the sound pressure is 65 dB (Shiraishi and Kanda, 2010). As each stimulus had a different sound pressure, we set the highest sound pressure at 65 dB. We used NL-27 (RION Corporation, Kokubunji, Japan) as the sound pressure meter. It was mounted at a height of 138.5 cm above the ground and at a distance of 118 cm from the speaker. Stimuli were presented using a JBL Pebbles speaker (HARMAN International Corporation, Northridge, CA, United States).

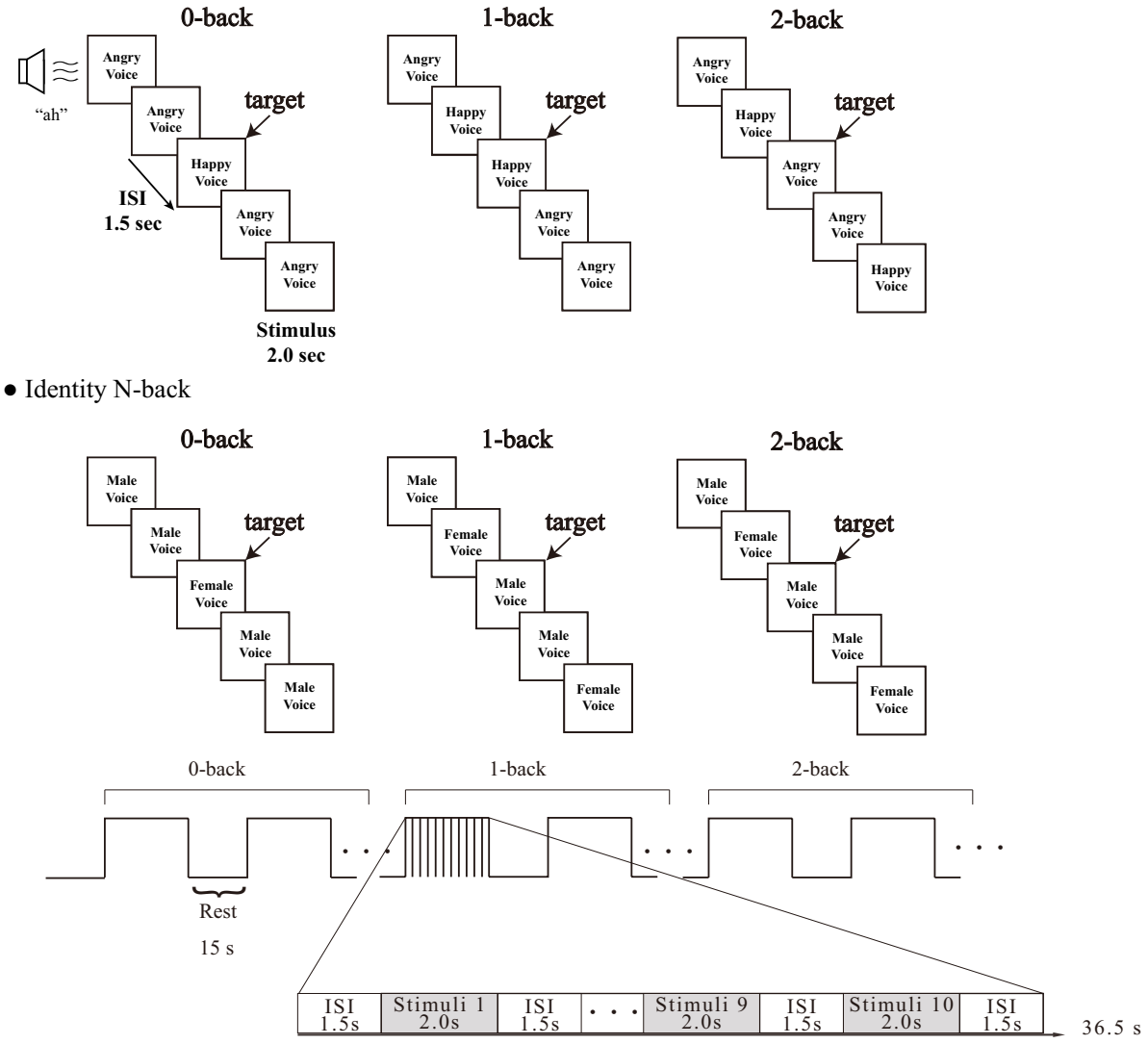


FIGURE 1
Experimental design of each task.

2.3 fNIRS measurement

We used a multichannel fNIRS system, ETG-4000 (Hitachi Corporation, Tokyo, Japan), which uses two wavelengths of near-infrared light (695 and 830 nm). We analyzed the optical data based on the modified Beer–Lambert Law (Cope et al., 1988) as previously described (Maki et al., 1995). This method enabled us to calculate signals reflecting the oxygenated hemoglobin (oxy-Hb), deoxygenated hemoglobin (deoxy-Hb), and total hemoglobin (total-Hb) signal changes, calculated in units of millimolar × millimeter (mM × mm). The sampling rate was set at 10 Hz. We used the oxy-Hb for analysis because it is the most sensitive indicator of regional cerebral hemodynamic response (Hoshi et al., 2001; Strangman et al., 2002; Hoshi, 2003).

2.4 fNIRS probe placement

We used a 3 × 11 multichannel probe holder that consisted of 17 illuminating and 16 detecting probes arranged alternately at an inter-probe distance of 3 cm. The probe was fixed using one 9 × 34 cm rubber shell and bandages mainly over the frontal areas. We defined channel positions in compliance with the international 10–20 system for EEG (Klem et al., 1999; Jurcak et al., 2007). The lowest probes were positioned along the Fpz, T3, and T4 line (horizontal reference curve).

Probabilistic spatial registration was employed to register the positions of each channel to Montreal Neurological Institute (MNI) standard brain space (Tsuzuki et al., 2007; Tsuzuki and Dan, 2014). Specifically, the positions for channels and reference points, which included the Nz (nasion), Cz (midline central), and left and right

preauricular points, were measured using a three-dimensional digitizer in real-world (RW) space. We affine-transformed each RW reference point to the corresponding MRI-database reference point and then replaced them to MNI space (Okamoto and Dan, 2005; Singh et al., 2005). Adopting the same transformation parameters enabled us to obtain the MNI coordinate values for the position of each channel in order to obtain the most likely estimate of the location of given channels for the group of participants and the spatial variability associated with the estimation. Finally, the estimated locations were anatomically labeled using a MATLAB® function that reads anatomical labeling information coded in a microanatomical brain atlas, LBPA40 (Shattuck et al., 2008) and Brodmann's atlas (Rorden and Brett, 2000; Figures 2, 3).

2.5 Behavioral performance data

To validate the results of brain activation analyses, we measured ACC and RT for each stimulus during the N-back tasks. Then, ACC and RT means were calculated for each participant. The calculated data were subjected to three-way repeated measures ANOVA, 3 (load: 0-back, 1-back, 2-back) × 2 (stimulus: Emotion, Identity) × 2 (trial: target, non-target), using IBM SPSS Statistics 26. For main effects, Bonferroni correction was performed. The significance level was set at 0.05.

2.6 Preprocessing of fNIRS data

We used Platform for Optical Topography Analysis Tools (POTATo) (Sutoko et al., 2016) for data preprocessing. Individual participants' timeline data for the oxy-Hb signal of each channel were preprocessed with a first-degree polynomial fitting and high-pass filter using cut-off frequencies of 0.0107-Hz to remove baseline drift, and a 0.1-Hz low-pass filter to remove heartbeat and pulse. We fit the baseline for task blocks with averages of 10 s before the introduction period.

2.7 fNIRS analysis

From the preprocessed time series data, we computed channel-wise and subject-wise contrasts by calculating the inter-trial mean of differences between the oxy-Hb signals for target periods (13–36.5 s after the start of the N-back task) and baseline periods (10 s before the start of the N-back task). We performed a one-sample *t*-test against zero (two tails) on preprocessed data.

A previous study (Zhou et al., 2021) demonstrated that brain activation is greater when the task load is efficient than when it is excessive or insufficient. Therefore, we explored conditions with the most sufficient cognitive loads to detect brain activations related to the N-back task in our study.

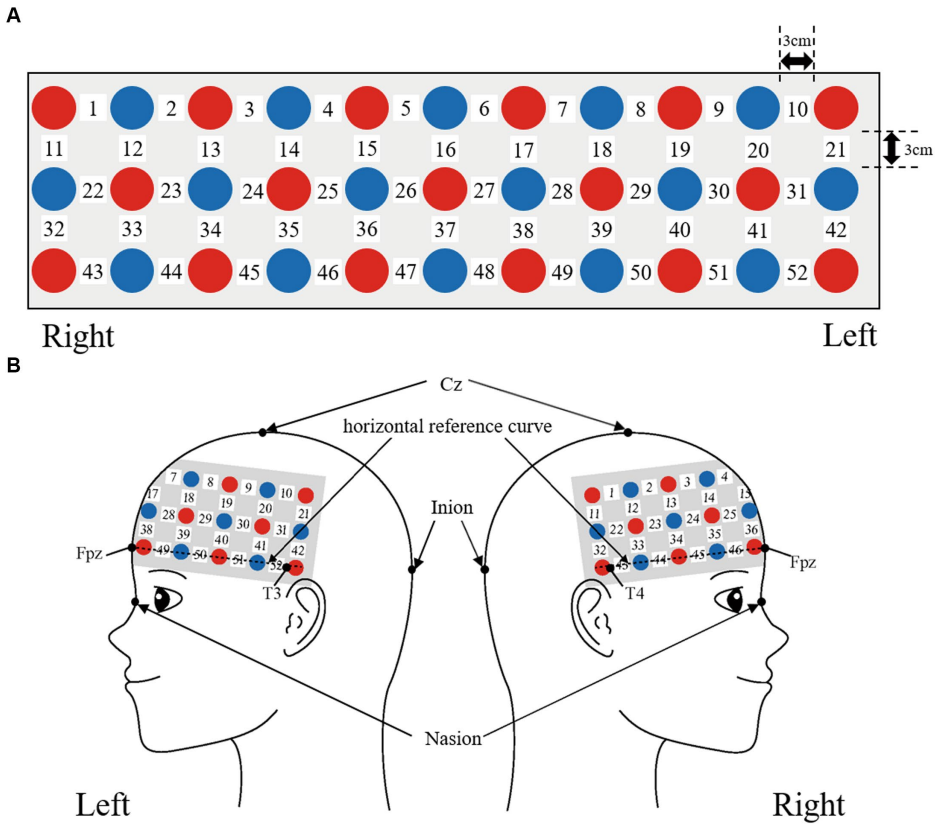


FIGURE 2
Spatial profiles of fNIRS channels. Detectors are indicated with blue circles, illuminators with red circles, and channels with white squares. **(A)** Position of probes and channels. **(B)** Left and right side views of the probe arrangement are exhibited with fNIRS channel orientation.

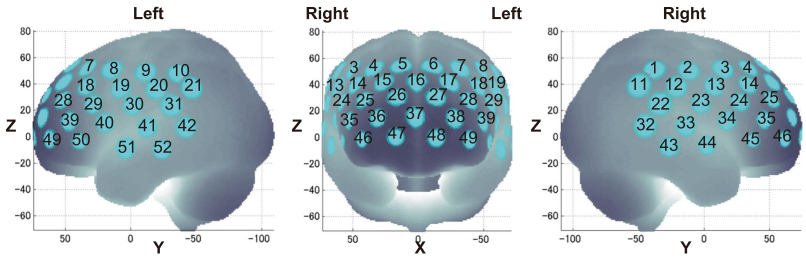


FIGURE 3
Channel locations on the brain are exhibited in left, frontal, and right side views. Probabilistically estimated fNIRS channel locations (centers of blue circles) and their spatial variability of 8 standard deviation (radii of the blue circles) associated with the estimation are depicted in Montreal Neurological Institute (MNI) space.

Activated channels were estimated from the effect size. We used G*Power (release 3.1.9.7) to consider a reasonable effect size for the 28 participants (Faul et al., 2007). Power analysis was conducted under the conditions of sample size = 28, one-sample *t*-test with $\alpha = 0.01$ and power = 0.80. A reasonable effect size of 0.57 was obtained to maintain the balance between Type I and Type II errors, based on the study by Cohen (1992). Therefore, we defined a reasonable effect size of 0.57 or more as activation of brain function in this analysis.

2.8 Between participants correlations

We conducted Spearman's correlation analyses for behavioral performance and brain activation to explore whether there was a connection between brain activation and behavioral performance. In addition, we conducted Spearman's correlation analyses between ACC and RT to examine whether judgment strategy differed between judging emotion and gender. We used IBM SPSS Statistics version 26. The significance level was set at 0.05.

3 Results

3.1 Behavior results

We calculated behavioral performance separately for each load, stimulus, and trial. For both ACC and RT, we conducted a three-way repeated-measures ANOVA, 3 (load: 0-back, 1-back, 2-back) × 2 (stimulus: Emotion, Identity) × 2 (trial: target, non-target). Figure 4 shows the results of the three-way ANOVA.

3.1.1 Accuracy

We revealed a significant main effect of trials [$F(1, 27) = 10.092$, $p < 0.01$, $\eta_p^2 = 0.272$]. Corrected pairwise comparisons (Bonferroni correction) revealed that participants were more accurate for non-target trials than target trials ($p < 0.01$).

3.1.2 Reaction time

We revealed a significant main effect of all factors [loads: $F(2, 54) = 28.523$, $p < 0.001$, $\eta_p^2 = 0.514$; stimuli: $F(1, 27) = 32.326$, $p < 0.001$, $\eta_p^2 = 0.545$; trials: $F(1, 27) = 18.382$, $p < 0.001$, $\eta_p^2 = 0.405$]. There was also a three-way interaction [$F(2, 54) = 4.364$, $p < 0.05$, $\eta_p^2 = 0.139$]. The *post-hoc* test showed that RTs were significantly longer between loads in the following order (longest to shortest): 2-back, 1-back, 0-back. The RTs were also significantly longer for the Emotion task than for the Identity task, and longer for non-target trials than for target trials. These results show that the 2-back task had the lowest ACC and the longest RT of the three loads. Therefore, we used performance data from during the 2-back task, which produced the best possible cognitive loads from among the three conditions, for further analysis (Tables 1, 2).

3.1.3 fNIRS results

Figure 5 shows the results of a one-sample *t*-test (vs. 0) on the calculated interval means. Our results showed a significant oxy-Hb signal increase during the Emotion 2-back task in seven channels and during the Identity 2-back task in seven channels. Tables 3, 4 show the locations of channels with significant activation. As task difficulty increased, activated channels increased progressively from the 0-back tasks with no activations, to one channel during Emotion and two channels during Identity 1-back tasks, and, finally, to all the channels indicated in Figure 5 during the respective 2-back tasks.

3.2 Between-participant correlations

First, we conducted Spearman's correlation analysis between ACC and brain activation. There were no significant correlations in the Identity 2-back task (Table 5).

Second, we conducted Spearman's correlation analysis between RT and brain activation. There were no significant correlations in the Emotion 2-back task. In the Identity 2-back task, we found significant positive correlations in five channels: channel 20, $\rho(28) = 0.43$, $p < 0.05$; channel 41, $\rho(28) = 0.47$, $p < 0.05$; channel 42, $\rho(28) = 0.44$, $p < 0.05$; channel 51, $\rho(28) = 0.48$, $p < 0.05$; and channel 52, $\rho(28) = 0.41$, $p < 0.05$. There was also a significant negative correlation in one channel: channel 24, $\rho(28) = -0.43$, $p < 0.05$. Table 6 shows the locations of channels with significant correlations.

Finally, we conducted Spearman's correlation analysis between ACC and RT. We found a significant positive correlation for the Emotion 2-back task [$\rho(28) = 0.375$, $p < 0.05$]. There were no significant correlations for the Identity 2-back task.

4 Discussion

In our present study, to clarify the cognitive mechanisms for the inference of emotions based on vocal speech, we examined brain

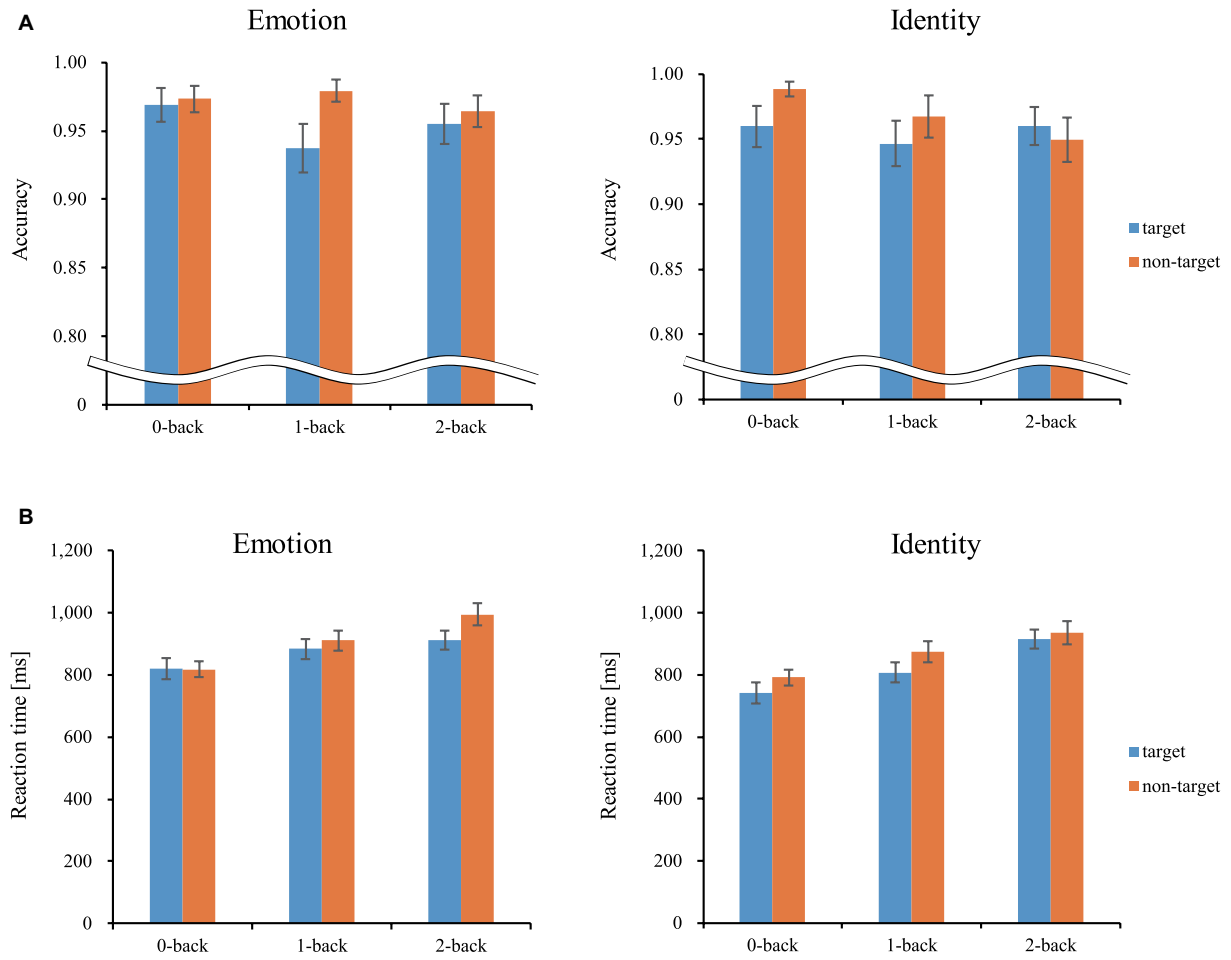


FIGURE 4
Mean accuracy and reaction time. Error bars represent standard error. ms, millisecond. **(A)** Accuracy. **(B)** Reaction time.

TABLE 1 Summary of results of activation analysis of Emotion 2-back task.

CH	*M*	*SD*	*t*	Sig	*d*
11	0.037	0.05	4.19	<0.001	0.79
13	0.054	0.09	3.20	0.004	0.60
21	0.035	0.05	3.65	0.001	0.69
24	0.044	0.07	3.26	0.003	0.62
39	0.035	0.05	3.99	<0.001	0.75
42	0.050	0.08	3.33	0.003	0.63
46	0.039	0.06	3.46	0.002	0.65

SD, standard deviation; *t*, *t*-value.

activation patterns during Emotion and Identity N-back tasks in combination with fNIRS measurements. Our results reveal different patterns not only in behavioral performance but also in brain activation between the tasks. The non-target trials elicited higher ACC than the target trials in both tasks. The Emotion task required longer RTs than the Identity task. Also, RTs were longer for non-target trials than for target trials, decreasing in order of 2-back, 1-back, and 0-back. Moreover, we observed different cortical activation patterns, possibly reflecting the differences in behavioral variables. During the Emotion task, there was activation in the right inferior frontal gyrus. During the Identity task, we observed activations in the left supramarginal, left postcentral, right middle temporal, right inferior frontal, and left superior temporal gyri.

The behavioral results showed that ACC did not differ across the two tasks or between cognitive loads, whereas the RT increased with increasing cognitive loads. In general, the larger the value of N in the N-back task was, the higher the subjective difficulty and the greater the cognitive load were (Yen et al., 2012). This suggests that the cognitive load of the 2-back task was highest, which led to a greater increase in RT compared to the other N-back tasks. Similar to facial expression experiments by Neta and Whalen (2011) and Xin and Lei (2015), RTs were also significantly longer for the Emotion task than for the Identity task. Therefore, it can be supposed that the inference of emotions generally demands a greater cognitive load than the inference of genders across different modalities.

As for the brain activation results, we focused on the 2-back conditions for both tasks since it required a sufficient cognitive load to effectively elicit brain activation reflecting cognitive processing. The number of activated channels increased progressively from no activations in the 0-back condition to a few in the 1-back condition and to several in the 2-back condition. This pattern serves as a good

indication that activations during the emotion task were due not to emotion-related physiological responses but to increased cognitive loads for processing emotion-related information since emotional loads are thought to be similar among corresponding n-back conditions.

Our results revealed distinct cortical regions specifically recruited for either the Emotion or Identity task. The right prefrontal regions in channel 24 exhibited specific focused activation during the Emotion task. The Emotion task should require attention to information about a specific emotion (happy or angry), which requires top-down processing based on voice information such as tone and pitch.

Many theoretical accounts of cognitive control have assumed that there would be a single system that mediates such endogenous attention (Posner and Petersen, 1990; Spence and Driver, 1997; Corbetta et al., 2008). Typically, such a system is referred to as the "dorsal attention network" (DAN) (Corbetta et al., 2008), which is a frontal–parietal network including the superior parietal lobule (SPL), the frontal eye field (FEF) in the superior frontal gyrus, and the middle frontal gyrus (MFG) (Corbetta et al., 2008). It has been shown that a top-down attention process is involved (Kincade et al., 2005; Vossel et al., 2006). DAN has been shown to be elicited not only in attention to visual stimuli (Büttner-Ennever and Horn, 1997; Behrmann et al., 2004; Corbetta et al., 2008), but also to auditory stimuli (Driver and Spence, 1998; Bushara et al., 1999; Linden et al., 1999; Downar et al., 2000; Maeder et al., 2001; Macaluso et al., 2003; Mayer et al., 2006; Shomstein and Yantis, 2006; Sridharan et al., 2007; Wu et al., 2007; Langner et al., 2012).

However, recent studies have confirmed the modality specificity of DAN. For visual attention, it has been speculated that the frontoparietal network, including the SPL, FEF, and MFG, plays an important role (Driver and Spence, 1998; Davis et al., 2000; Macaluso et al., 2003; Langner et al., 2011). However, for auditory attention, a network including the MFG and posterior middle temporal gyrus has been shown to be mainly at work (Braga et al., 2013). The right prefrontal regions activated in the current study's Emotion task were roughly consistent with the regions shown to make up the attentional network for auditory top-down processing (Braga et al., 2013), although channel 13 was located slightly posterior and channels 24 and 46 slightly inferior, respectively.

TABLE 2 Summary results of activation analysis for Identity task.

CH	*M*	*SD*	*t*	*p*	*d*
20	0.048	0.080	3.16	0.004	0.60
21	0.035	0.056	3.34	0.002	0.63
30	0.074	0.104	3.77	0.001	0.71
31	0.070	0.110	3.36	0.002	0.63
32	0.054	0.076	3.81	0.001	0.72
42	0.075	0.114	3.48	0.002	0.66
46	0.043	0.059	3.92	0.001	0.74

SD, standard deviation; *t*, t-value.

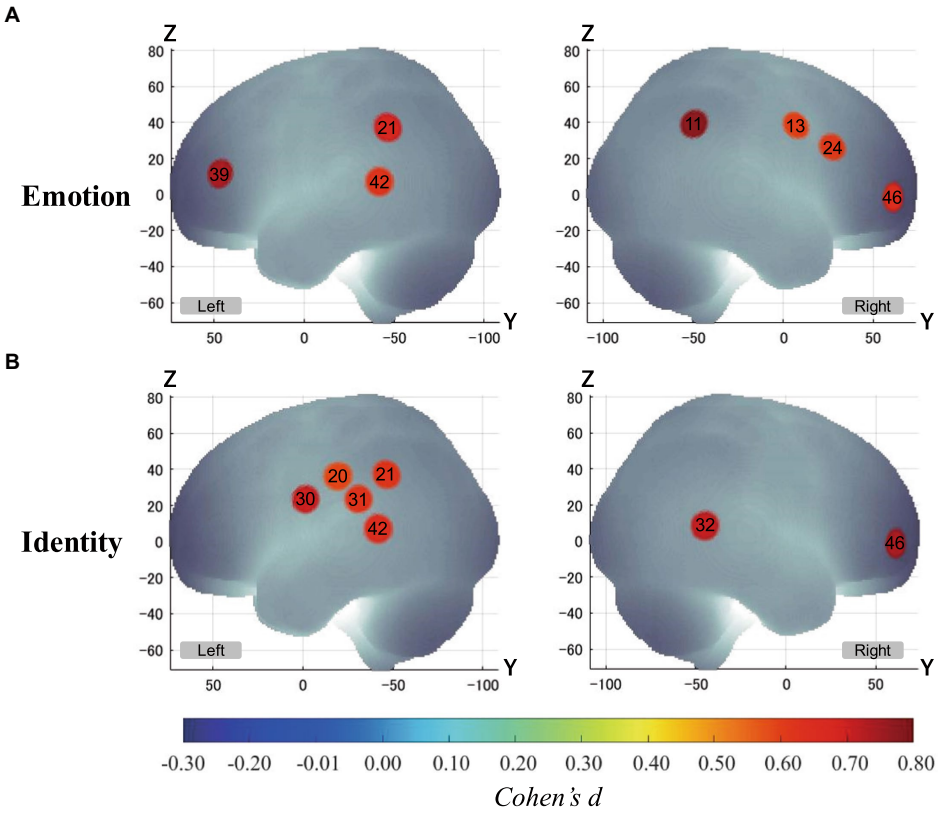


FIGURE 5
Activation t-maps of oxy-Hb signal increase during 2-back task. Only channels defined as activated (α <0.01 and Cohen's d >0.57) are exhibited in MNI space. Among them, Ch39 was found activated during the 1-back Emotion task and Chs21 and 32 were activated during the 1-back Identity task. For both 2-back task conditions, all channels indicated above were activated, and no channels were found activated in any of the 0-back tasks. **(A)** Emotion task. **(B)** Identity task.

TABLE 3 Estimated most likely locations of activated channels from spatial registration for Emotion 2-back task.

CH	MNI-coordinates				LPBA40	Prob. (%)	Brodmann area	Prob. (%)
	x	*y*	*z*	*SD*				
11	64.7	−48.3	39.7	10.0	R AG	78	SMG part of Wernicke's area (BA40)	90
13	61.0	7.3	38.3	8.5	R PRG	93	Pre-SMA(BA6)	81
21	−65.0	−46.3	37.3	9.5	L SMG	67	SMG part of Wernicke's area (BA40)	87
24	57.3	27.7	26.3	7.8	R IFG	87	Pars triangularis Broca's area (BA45)	85
39	−49.0	46.7	11.7	8.3	L IFG	71	Pars triangularis Broca's area (BA45)	52
42	−70.0	−41.7	7.3	8.8	L STG	51	STG (BA22)	79
46	42.7	61.3	−1.3	6.8	R IFG	89	FPA (BA10)	74

Prob, probability; SD, standard deviation. SD is expressed in millimeters (mm). AG, angular gyrus; IFG, inferior frontal gyrus; FPA, frontopolar cortex; Pre-SMA, pre-motor and supplementary motor cortex; PRG, precentral gyrus; SMG, supramarginal gyrus; STG, superior temporal gyrus.

TABLE 4 Estimated most likely locations of activated channels from spatial registration for Identity 2-back task.

CH	MNI-coordinates				LPBA40	Prob. (%)	Brodmann area	Prob. (%)
	x	*y*	*z*	*SD*				
20	−66.0	−19.3	36.7	8.9	L SMG	72	S1 (BA2)	49
21	−65.0	−46.3	37.3	9.5	L SMG	67	SMG part of Wernicke's area (BA40)	87
30	−65.7	−1.3	24.3	8.1	L PoCG	57	Subcentral area (BA43)	87
31	−69.0	−30.7	24.3	9.1	L SMG	61	S1 (BA2)	42
32	71.0	−44.3	9.3	9.3	R MTG	76	STG (BA22)	77
42	−70.0	−41.7	7.3	8.8	L STG	51	STG (BA22)	79
46	42.7	61.3	−1.3	6.8	R IFG	89	FPA (BA10)	74

Prob, probability; SD, standard deviation. SD is expressed in millimeters (mm). FPA, frontopolar cortex; IFG, inferior frontal gyrus; MTG, middle temporal gyrus; PoCG, post central gyrus; S1, primary somatosensory cortex; SMG, supramarginal gyrus; STG, superior temporal gyrus.

The slight differences between our results and those described by Braga et al. (2013) may reflect differences of task demands. Braga et al. (2013) requested participants to detect physical features (i.e., changes of pitch) from stereo, naturalistic background sounds. On the other hand, we asked participants to infer the emotions of others from human voices in the Emotion task. Our task would have involved not only physical feature detection but also other cognitive processes (e.g., considering the mood of others), which would, in turn, involve more complex task demands.

From a different perspective, channels 24 and 39 were located over the ventrolateral prefrontal cortex (VLPFC), which is known to play important roles in working memory (Owen et al., 2005; Barbey et al., 2013). Wagner (1999) indicated that VLPFC activity had a privileged role in domain-specific WM processes of phonological and visuospatial rehearsal. However, Jolles et al. (2010) demonstrated that greater activation in the VLPFC was observed in response to the increasing task demands of a WM task. Indeed, our results suggest that the Emotion task would demand more complex cognitive processes since the RTs for the Emotion task were longer than those for the Identity task. Such differences in task demands may be reflected in the cortical activation in the VLPFC observed only during the Emotion task.

On the other hand, there were several regions specifically activated during the Identity task, in particular the left temporal region, including the left supramarginal gyrus, at channels 20, 30, and 31. The supramarginal gyrus has been shown to be involved in short-term auditory memory (Gaab et al., 2003, 2006; Vines et al., 2006). Moreover, hemispheric differences have been reported: During a short-term auditory memory task, there was greater activation in the left supramarginal gyrus than in the right (Gaab et al., 2006). The results of these previous studies suggest that the left supramarginal gyrus works as a center for auditory short-term memory and also influences the allocation of processing in the auditory domain as part of a top-down system.

The right middle temporal gyrus was also activated only during the Identity task. The middle temporal gyrus is known to play an important role in semantic memory processing (Onitsuka et al., 2004; Visser et al., 2012). Clinically, activation in the right middle temporal gyrus is known to be reduced in schizophrenic patients, and its reduction is associated with auditory verbal hallucinations and difficulty in semantic memory processing (Woodruff et al., 1995; Lennox et al., 1999, 2000). These findings suggest that the right temporal gyrus plays an important role in auditory semantic memory processing.

By interpreting the functions of these brain regions during both tasks, we can begin to integratively explain the physiological mechanisms involved in the inference of emotions based on auditory information. Specifically, when judging emotional information compared to judging the gender of voice stimuli, auditory attention could be directed more deeply. Therefore, the regions responsible for

TABLE 5 Correlation between brain activation and ACC during Emotion 2-back task.

CH	MNI-coordinates				LPBA40	Prob. (%)	Brodmann area	Prob. (%)
	x	*y*	*z*	*SD*				
4	32.7	33.7	52.0	9.4	R MFG	95	DLPFC (BA9)	66
12	68.0	−20.3	38.7	8.6	R SMG	100	S1 (BA1)	54
13	61.0	7.3	38.3	8.5	R PRG	93	Pre-SMA (BA6)	81
22	71.0	−32.7	26.3	9.5	R SMG	64	S1 (BA2)	31
24	57.3	27.7	26.3	7.8	R IFG	87	Pars triangularis Broca's area (BA45)	85

Prob, probability; SD, standard deviation. SD is expressed in millimeters (mm). DLPFC, dorsolateral prefrontal cortex; IFG, inferior frontal gyrus; MFG, middle frontal gyrus; Pre-SMA, pre-motor and supplementary motor cortex; PRG, precentral gyrus; S1, primary somatosensory cortex; SMG, supramarginal gyrus.

TABLE 6 Correlation between brain activation signals and RT during Identity 2-back task.

CH	MNI-coordinates				LPBA40	Prob. (%)	Brodmann area	Prob. (%)
	x	*y*	*z*	*SD*				
20	−66.0	−19.3	36.7	8.9	L SMG	72	S1 (BA2)	49
24	57.3	27.7	26.3	7.8	R IFG	87	Pars triangularis Broca's area (BA45)	85
31	−69.0	−30.7	24.3	9.1	L SMG	61	S1 (BA2)	42
41	−67.7	−12.7	7.3	8.7	L STG	66	STG (BA22)	94
42	−70.0	−41.7	7.3	8.8	L STG	51	STG (BA22)	79
51	−62.7	3.7	−8.3	9.0	L STG	84	MTG (BA21)	62
52	−71.0	−23.3	−9.3	7.8	L MTG	100	MTG (BA21)	100

Prob, probability; SD, standard deviation. SD is expressed in millimeters (mm). IFG, inferior frontal gyrus; MTG, middle temporal gyrus; S1, primary somatosensory cortex; SMG, supramarginal gyrus; STG, superior temporal gyrus.

complex cognition, including working memory, would be working together. On the other hand, when judging gender rather than emotional information, the demand of processing physical information from a vocal stimulus would, rather, be greater. It is possible that we are making judgments based on the characteristics of the sound itself, including pitch information. This interpretation is rational because the activation in the supramarginal gyrus was localized to the left hemisphere during the Identity task. Previous studies support the predominance of working memory system requirements for emotional judgments as well as the predominance of speech information processing requirements for gender judgments (e.g., Whitney et al., 2011a,b). The left temporal region at channels 20 and 31 which activated during the Identity task and the left inferior frontal gyrus at channel 39 which activated during the Emotion task have been found to play different roles in semantic cognition. The left temporal region is involved in accessing information in the semantic store (Hodges et al., 1992; Vandenberghe et al., 1996; Hickok and Poeppel, 2004; Indefrey and Levelt, 2004; Rogers et al., 2004; Vigneau et al., 2006; Patterson et al., 2007; Binder et al., 2009; Binney et al., 2010). On the other hand, the left inferior frontal gyrus is involved in executive mechanisms that direct the activation of semantics appropriate to the current task and context (Thompson-Schill et al., 1997; Wagner et al., 2001; Bookheimer, 2002; Badre et al., 2005; Ye and Zhou, 2009). The present Identity 2-back task would have required these cognitive processes.

We also examined the correlation between behavioral performance and brain activation to determine how brain activation associated with voice-based inference of emotions is related to behavioral performance. In the Emotion task, the greater the activation in the right frontal region to the supramarginal gyrus was, the higher the ACC was. This suggests that these regions contribute to correctly inferring emotions based on speech information. Regarding the relationship between RT and ACC, in the Emotion task alone, the slower the RT was, the higher the ACC was. This indicates that there was a trade-off between RT and ACC when inferring emotions. Thus, the Emotion task and Identity task elicited different brain activation patterns under different cognitive strategies derived from different cognitive demands.

For the Emotion task, considering the activation of regions related to the auditory attention network, we submit that the cognitive strategy was not an immediate response to auditory cues, but rather an analytical processing of the features of the auditory information with auditory attention for accurate inferences. On the other hand, for the Identity task, no correlation was found between RT and ACC. This indicates that the ACC was not determined by RT, and accurate inferences could be made without sacrificing speed. We can interpret these results with reference to the findings of a previous study examining facial-expression-based inference of emotions (Neta and Whalen, 2011). That is, changeable features (e.g., emotions) involve more complex judgments requiring relatively more auditory attention, similarly to the visual attention mentioned in Neta and Whalen (2011), than do invariant features (e.g., gender, identity).

Though N-back tasks are known to require working memory processes (Owen et al., 2005; Kane et al., 2007), the VLPFC in the right inferior frontal regions, which is the center of working memory, elicited activation only during the Emotion task, and not during the Identity task. However, there was a negative correlation between brain activation and RT in the region located on the right VLPFC (channel 24) during the Identity task (see Figure 6). This suggests that the VLPFC served an important function in the Identity task as well as in

the Emotion task. Considering the difference in cognitive demands between the Emotion and Identity tasks, it is likely that the demands of speech information processes were more dominant than working memory processes during the Identity task. Thus, no significant activation was observed in the area located on the VLPFC (channel 24) during the Identity task (see Figure 6).

From the results of our study, we can infer that different cognitive demands underlie each task (Emotion and Identity). To determine gender based on voice information, it is sufficient to access the semantic store from the voice features and check whether the feature information is consistent with the knowledge in the semantic store. On the other hand, when judging emotions in speech information, it is necessary not only to check whether the characteristics of the voice are consistent with the emotional information in the semantic store, but also to judge the emotional information in conjunction with the context behind it. In other words, it is necessary to infer the mental state of the speaker by considering not only physically produced qualities such as pitch, but also the context within which the speaker is vocalizing. Such differences in cognitive demand should produce different patterns of brain activation during the Emotion and the Identity N-back tasks. Indeed, during the Emotion task, the reaction time was longer due to the additional processing needed to infer the mental state of the speaker in addition to judging the morphological characteristics of the voice.

In the present study, there are several limitations that should be considered. First, the activation of the amygdala, which is widely known to be related to emotional functioning (e.g., Gallagher and Chiba, 1996; Phelps, 2006; Pessoa, 2010), was not considered; because fNIRS can only measure the activation of the cortical surface, it cannot measure the activation of the amygdala. Since voice-based emotional judgments occur with social interactions, their neural basis is expected to be complex, involving a range of cortical and subcortical areas and connectivity pathways. In the present study, we used speech voices as stimuli, which provide only auditory information. However, as with Neta and Whalen (2011) use of facial expressions as visual information, the amygdala may be involved in emotional judgments regardless of the sensory modality. Though future studies using fMRI may clarify the role of the pyramid in the inference of emotions based on auditory information, there would inevitably be technical problems related to the noisy environment.

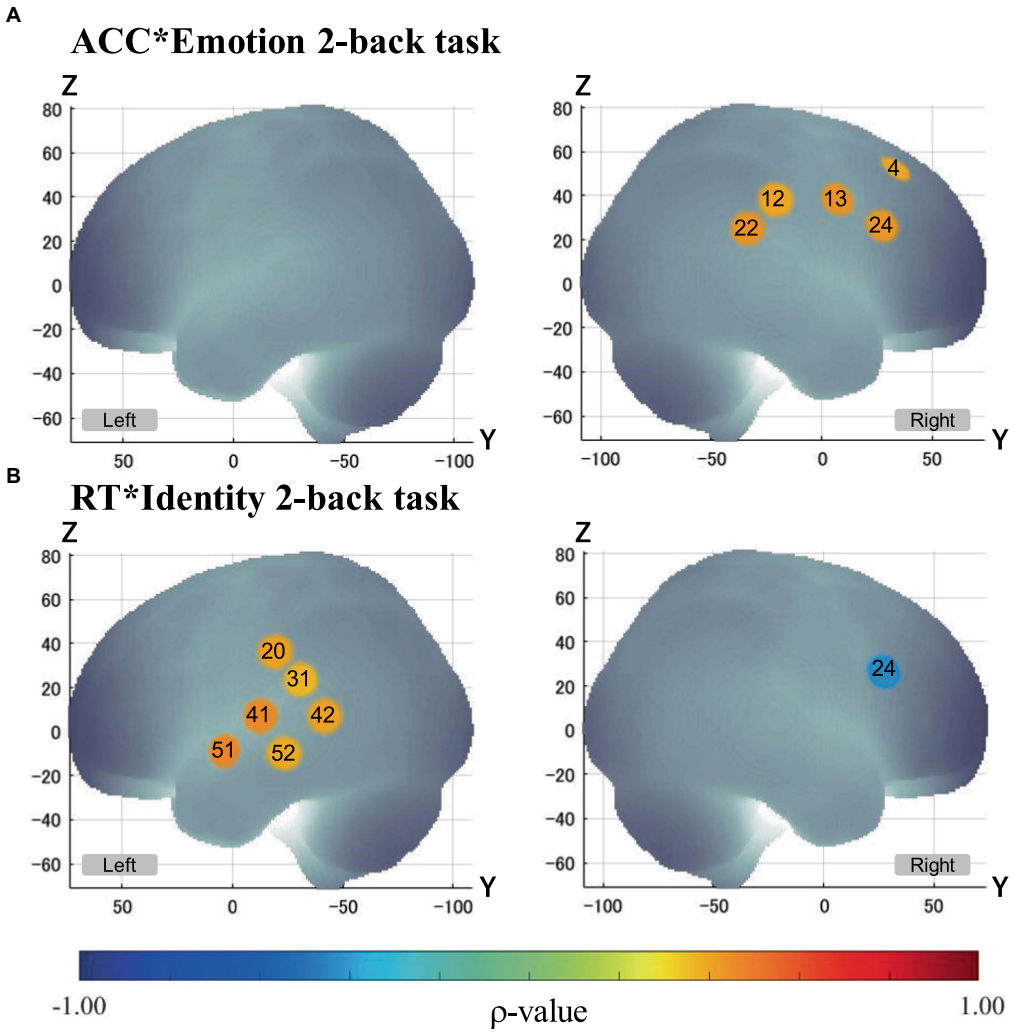


FIGURE 6
Correlation between brain activation and behavioral performance. Only channels with significant Spearman's correlation (α <0.05) are exhibited in MNI space. *p*-values are as indicated in the color bar. **(A)** Correlation between ACC and brain activation in the Emotion 2-back task (ACC*Emotion 2-back task). **(B)** Correlation between RT and brain activation in the Identity 2-back task (ACC*Identity 2-back task).

Second, the current experimental design did not allow us to discuss the physiological mechanisms that depend on the type of emotion. We aimed to clarify the underlying mechanisms of emotional judgments using two clearly distinguishable emotions (Happy/Angry) measured in the framework of a block design, and we were able to achieve this goal. However, the differences in the relevant domains are unclear due to the differences in the emotional valence. In general, it is known that basic emotions are expressed by two axes, valence and arousal (Yang and Chen, 2012), of which Happy and Angry have equal arousal but significantly different valence (Fehr and Russell, 1984; Yang and Chen, 2012). As a result, Happy is considered a positive emotion and Angry is considered a negative emotion.

In our present study, we used voices with happy and angry emotions as the stimuli; however, there are, of course, other basic emotions, such as sadness, fear, disgust, contempt, and surprise (Ekman and Cordaro, 2011). Different emotions have varied emotional valence and arousal, and such differences could result in different cognitive processing mechanisms. One possible future study could use an event-related design with voices expressing other basic emotions, which would allow us to clarify the mechanisms for each individual basic emotion.

Third, although the current study provides an important step toward the clinical applicability of emotional working memory tasks, we have yet to optimize the analytical methods applied here. Instead of using a regression analysis of the observed time-series data to modeled hemodynamic responses in a time series, a procedure typically referred to as a general linear model (GLM) analysis (Uga et al., 2014), we used a method that averages hemodynamic responses over a fixed time window. This was to ensure robust analyses of observed data with both temporal and spatial heterogeneity. However, to achieve a better understanding of the cognitive processes underlying the emotional and identity n-back tasks, we should optimize either hemodynamic response models, task structures or both.

Fourth, in the current study, we used continuous-wave fNIRS with an equidistant probe setting. However, this entails the possibility of physiological noises such as skin blood flow and systemic signals being included in the observed hemodynamic responses. While such physiological noises tend to be distributed globally, the cortical activation patterns observed in the current study mostly exhibited asymmetric distributions. Thus, it is expected that activations detected in the current study may well represent cognitive components associated with tasks and task contrasts. For further validation, we must extract genuine cognitive components for identity and emotional working memory by reducing the effects of physiological noise to a minimum.

Fifth, in a related vein, data analysis parameters used for the current study were optimized for oxy-Hb data. Had we applied those optimized for deoxy-Hb, it might have led to different results. Although we focused on oxy-Hb results in the current study and reported deoxy-Hb data using the same parameter sets in Supplementary material, future studies might be able to build upon our results and elucidate the potential use of deoxy-Hb signals as they relate to emotion and identity N-back tasks.

Despite the limitations mentioned above, here we have used fNIRS to provide new insights into the neurological mechanisms underlying the inference of emotions based on audio information for the first time. The primary importance of the findings of this study lie in the great potential for clinical applications. Many researchers have shown the association between difficulties inferring emotions and several types of disorders: neurological (Bora et al., 2016a,b), psychiatric (Gray and Tickle-Degnen, 2010; Dalili et al., 2015), and developmental (Bora and Pantelis, 2016). In particular, social cognitive dysfunction, including difficulties inferring emotions based on facial expression cues, has been identified as a potential marker for the developmental disorder autism and the psychopathology schizophrenia (Derntl and Habel, 2011).

Neurodegenerative diseases such as behavioral variant frontopolar dementia (bvFTD) and amyotrophic lateral sclerosis (ALS) have also been found to manifest social cognitive impairment, including difficulties inferring emotions (Bora et al., 2016c; Bora, 2017). Thus, the neurological measurement of such difficulties could serve as a marker for nerve deterioration, disease progression, and even treatment response (Cotter et al., 2016). Cotter et al. (2016) conducted a meta-analysis of related studies and provided a comprehensive review. They concluded that difficulties with the inference of emotions cued by facial expressions could be a core cognitive phenotype of many developmental, neurological, and psychiatric disorders. In addition, though most studies on the relationship between the inference of emotions and disease have used facial expressions denoting emotions, some experimental studies have focused on the inference of emotions using speech voices for patients with Parkinson's disease (PD) and with eating disorders (ED) (Gray and Tickle-Degnen, 2010; Caglar-Nazali et al., 2014). Since these studies also involved discriminating emotions based on speech voices, it is clearly possible to examine the brain regions of patients using our experimental tasks (especially the voice-based Emotion N-back task). Our findings provide essential information contributing to the identification of biomarkers associated with difficulties inferring emotions based on vocal cues. In other words, there is a possibility that the emotion-specific brain activation observed in healthy participants may not be the same as that observed in the clinical group. In the future, it will be necessary to clarify the brain activation associated with the inference of emotions based on vocal cues in clinical groups.

For future clinical application, fNIRS offers unique merits in assessing auditory Emotion and Identification tasks. By using fNIRS, we were able to detect regions related to the inference of emotions with speech voices, controlling the effects of sounds other than the stimuli. While fMRI has the great advantage of being able to reveal and measure relevant regions, including deep cortical regions, it is difficult to completely eliminate the effects of non-stimulus noise during measurement. On the other hand, fNIRS can be used in a relatively quiet environment to control the influence of ambient sound on brain activation. Therefore, we believe that fNIRS is the most reliable method of examining the neural basis for the inference of emotions based on vocal cues. In this sense, our present study has clarified the cognitive processing associated with the inference of emotions based on vocal cues using fNIRS, which is of great clinical and academic significance.

5 Conclusion

We investigated the neural basis of working memory processing during auditory emotional judgment in healthy participants and revealed that different brain activation patterns can be observed between the two tasks (Emotion, Identity). During the Emotion task, there was significant activation of brain regions known to be part of an attentional network in auditory top-down processing. In addition, there was significant positive correlations between the improvement of behavioral performance and activated cortical regions including the right frontal region and the supramarginal gyrus. Therefore, the current

study suggests that the cognitive demands of analyzing and processing the features of auditory information with auditory attention in order to make emotional judgments based solely on auditory information differ from those based on visual information such as facial expressions. Thus, we have provided an inclusive view of the cognitive mechanisms behind inferring the emotions of others based on analyses of fNIRS data and behavior data. In the future, it will be possible to apply our findings to the clinical diagnoses of some diseases associated with deficits of emotional judgment to contribute to early intervention and to providing criteria for selecting appropriate treatments.

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 humans were approved by the Institutional Ethics Committee of Chuo University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

MK, HS, KO, and TH contributed to the conception and design of the study. HS, KO, and KN performed the experiments. SO, WK, KN, KO, and SN analyzed the data and performed the statistical analyses. YK supervised the statistical analysis plan. SO and KN wrote the manuscript. WK, MK, YK, and ID reviewed and edited the manuscript. WK and ID prepared the Supplementary material. ID supervised the study. All authors have reviewed the manuscript and approved the final version for publication.

Funding

This study was in part supported by a Grant-in-Aid for Scientific Research (22K18653 to ID, 22H00681 to ID, 19H00632 to ID, and 19K08028 to MK), and [KAKENHI (Multi-year Fund)] -fund for the Promotion of Joint International Research (Fostering Joint International Research) (16KK0212 to MK).

Acknowledgments

We would like to thank Yuko Arakaki and Melissa Noguchi for their help with the language editing of this paper. Also, we would like to thank Hiroko Ishida for administrative assistance.

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

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnhum.2023.1160392/full#supplementary-material

References

Baddeley, A. (1998). Recent developments in working memory. *Curr. Opin. Neurobiol.* 8, 234–238. doi: 10.1016/S0959-4388(98)80145-1

Baddeley, A. D., and Hitch, G. (1974). "Working memory" in *Psychology of learning and motivation*. ed. G. A. Bower (New York, NY: Academic press), 47–89.

Badre, D., Poldrack, R. A., Paré-Blagoev, E. J., Insler, R. Z., and Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron* 47, 907–918. doi: 10.1016/j.neuron.2005.07.023

Barbey, A. K., Koenigs, M., and Grafman, J. (2013). Dorsolateral prefrontal contributions to human working memory. *Cortex* 49, 1195–1205. doi: 10.1016/j.cortex.2012.05.022

Behrmann, M., Geng, J. J., and Shomstein, S. (2004). Parietal cortex and attention. *Curr. Opin. Neurobiol.* 14, 212–217. doi: 10.1016/j.conb.2004.03.012

Belin, P., Fillion-Bilodeau, S., and Gosselin, F. (2008). The Montreal affective voices: a validated set of nonverbal affect bursts for research on auditory affective processing. *Behav. Res. Methods* 40, 531–539. doi: 10.3758/BRM.40.2.531

Binder, J. R., Desai, R. H., Graves, W. W., and Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebellum* 19, 2767–2796. doi: 10.1093/cercor/bhp055

Binney, R. J., Embleton, K. V., Jefferies, E., Parker, G. J., and Lambon Ralph, M. A. (2010). The ventral and inferolateral aspects of the anterior temporal lobe are crucial in semantic memory: evidence from a novel direct comparison of distortion-corrected fMRI, rTMS, and semantic dementia. *Cerebellum* 20, 2728–2738. doi: 10.1093/cercor/bhq019

Bookheimer, S. (2002). Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annu. Rev. Neurosci.* 25, 151–188. doi: 10.1146/annurev.neuro.25.112701.142946

Bora, E. (2017). Meta-analysis of social cognition in amyotrophic lateral sclerosis. *Cortex* 88, 1–7. doi: 10.1016/j.cortex.2016.11.012

Bora, E., Özakbaş, S., Velakoulis, D., and Walterfang, M. (2016a). Social cognition in multiple sclerosis: a meta-analysis. *Neuropsychol. Rev.* 26, 160–172. doi: 10.1007/s11065-016-9320-6

Bora, E., and Pantelis, C. (2016). Meta-analysis of social cognition in attention-deficit/hyperactivity disorder (ADHD): comparison with healthy controls and autistic spectrum disorder. *Psychol. Med.* 46, 699–716. doi: 10.1017/S0033291715002573

Bora, E., Velakoulis, D., and Walterfang, M. (2016b). Meta-analysis of facial emotion recognition in behavioral variant frontotemporal dementia: comparison with Alzheimer disease and healthy controls. *J. Geriatr. Psychiatry Neurol.* 29, 205–211. doi: 10.1177/0891988716640375

Bora, E., Velakoulis, D., and Walterfang, M. (2016c). Social cognition in Huntington's disease: a meta-analysis. *Behav. Brain Res.* 297, 131–140. doi: 10.1016/j.bbr.2015.10.001

Braga, R. M., Wilson, L. R., Sharp, D. J., Wise, R. J., and Leech, R. (2013). Separable networks for top-down attention to auditory non-spatial and visuospatial modalities. *Neuroimage* 74, 77–86. doi: 10.1016/j.neuroimage.2013.02.023

Braver, T. S., Barch, D. M., Kelley, W. M., Buckner, R. L., Cohen, N. J., Miezin, F. M., et al. (2001). Direct comparison of prefrontal cortex regions engaged by working and long-term memory tasks. *Neuroimage* 14, 48–59. doi: 10.1006/nimg.2001.0791

Bucci, S., Startup, M., Wynn, P., Baker, A., and Lewin, T. J. (2008). Referential delusions of communication and interpretations of gestures. *Psychiatry Res.* 158, 27–34. doi: 10.1016/j.psychres.2007.07.004

Bushara, K. O., Weeks, R. A., Ishii, K., Catalan, M. J., Tian, B., Rauschecker, J. P., et al. (1999). Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. *Nat. Neurosci.* 2, 759–766. doi: 10.1038/11239

Büttner-Ennever, J. A., and Horn, A. K. (1997). Anatomical substrates of oculomotor control. *Curr. Opin. Neurobiol.* 7, 872–879. doi: 10.1016/S0959-4388(97)80149-3

Caglar-Nazali, H. P., Corfield, F., Cardi, V., Ambwani, S., Leppanen, J., Olabintan, O., et al. (2014). A systematic review and meta-analysis of 'systems for social processes' in eating disorders. *Neurosci. Biobehav.* 42, 55–92. doi: 10.1016/j.neubiorev.2013.12.002

Calder, A. J., and Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nat. Rev. Neurosci.* 6, 641–651. doi: 10.1038/nrn1724

Cohen, J. (1992). A power primer. *Psychol. Bull.* 112, 155–159. doi: 10.1037//0033-2909.112.1.155

Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., et al. (2008). Audio-visual integration of emotion expression. *Brain Res.* 1242, 126–135. doi: 10.1016/j.brainres.2008.04.023

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*. eds. M. Mochizuki, C. R. Honig, T. Koyama, T. K. Goldstick and D. F. Bruley (Springer: New York, NY), 183–189.

Corbetta, M., Patel, G., and Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron* 58, 306–324. doi: 10.1016/j.neuron.2008.04.017

Cotter, J., Firth, J., Enzinger, C., Kontopantelis, E., Yung, A. R., Elliott, R., et al. (2016). Social cognition in multiple sclerosis: a systematic review and meta-analysis. *Neurology* 87, 1727–1736. doi: 10.1212/WNL.0000000000003236

Dalili, M. N., Penton-Voak, I. S., Harmer, C. J., and Munafò, M. R. (2015). Meta-analysis of emotion recognition deficits in major depressive disorder. *Psychol. Med.* 45, 1135–1144. doi: 10.1017/S0033291714002591

Davis, K. D., Hutchison, W. D., Lozano, A. M., Tasker, R. R., and Dostrovsky, J. O. (2000). Human anterior cingulate cortex neurons modulated by attention-demanding tasks. *J. Neurophysiol.* 83, 3575–3577. doi: 10.1152/jn.2000.83.6.3575

De Gelder, B., and Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cogn Emot* 14, 289–311. doi: 10.1080/026999300378824

Derntl, B., and Habel, U. (2011). Deficits in social cognition: a marker for psychiatric disorders? *Eur. Arch. Psychiatry Clin. Neurosci.* 261, 145–149. doi: 10.1007/s00406-011-0244-0

Downar, J., Crawley, A. P., Mikulis, D. J., and Davis, K. D. (2000). A multimodal cortical network for the detection of changes in the sensory environment. *Nat. Neurosci.* 3, 277–283. doi: 10.1038/72991

Driver, J., and Spence, C. (1998). Cross–modal links in spatial attention. *Philos. Trans. R. Soc.* 353, 1319–1331. doi: 10.1098/rstb.1998.0286

Ekman, P., and Cordaro, D. (2011). What is meant by calling emotions basic. *Emot. Rev.* 3, 364–370. doi: 10.1177/1754073911410740

Esposito, A., Riviello, M. T., and Bourbakis, N. (2009). "Cultural specific effects on the recognition of basic emotions: a study on Italian subjects" in *HCI and usability for e-inclusion*. eds. A. Holzinger and K. Miesenberger (Heidelberg, Berlin: Springer), 135–148.

Faul, F., Erdfelder, E., Lang, A. G., and Buchner, A. (2007). G* power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/BF03193146

Fehr, B., and Russell, J. A. (1984). Concept of emotion viewed from a prototype perspective. *J. Exp. Psychol. Gen.* 113, 464–486. doi: 10.1037/0096-3445.113.3.464

Gaab, N., Gaser, C., and Schlaug, G. (2006). Improvement-related functional plasticity following pitch memory training. *Neuroimage* 31, 255–263. doi: 10.1016/j.neuroimage.2005.11.046

Gaab, N., Gaser, C., Zaehle, T., Jancke, L., and Schlaug, G. (2003). Functional anatomy of pitch memory—an fMRI study with sparse temporal sampling. *Neuroimage* 19, 1417–1426. doi: 10.1016/S1053-8119(03)00224-6

Galderisi, S., Mucci, A., Bitter, I., Libiger, J., Bucci, P., Fleischhacker, W. W., et al. (2013). Persistent negative symptoms in first episode patients with schizophrenia: results from the European first episode schizophrenia trial. *Eur. Neuropsychopharmacol.* 23, 196–204. doi: 10.1016/j.euroneuro.2012.04.019

Gallagher, M., and Chiba, A. A. (1996). The amygdala and emotion. *Curr. Opin. Neurobiol.* 6, 221–227. doi: 10.1016/S0959-4388(96)80076-6

Gobbini, M. I., and Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia* 45, 32–41. doi: 10.1016/j.neuropsychologia.2006.04.015

Gray, H. M., and Tickle-Degnen, L. (2010). A meta-analysis of performance on emotion recognition tasks in Parkinson's disease. *Neuropsychology* 24, 176–191. doi: 10.1037/a0018104

Gregersen, T. S. (2005). Nonverbal cues: clues to the detection of foreign language anxiety. *Foreign Lang. Ann.* 38, 388–400. doi: 10.1111/j.1944-9720.2005.tb02225.x

Haxby, J. V., Hoffman, E. A., and Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends Cogn. Sci.* 4, 223–233. doi: 10.1016/S1364-6613(00)01482-0

Hickok, G., and Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition* 92, 67–99. doi: 10.1016/j.cognition.2003.10.011

Hodges, J. R., Patterson, K., Oxbury, S., and Funnell, E. (1992). Semantic dementia: progressive fluent aphasia with temporal lobe atrophy. *Brain* 115, 1783–1806. doi: 10.1093/brain/115.6.1783

Hoffman, E. A., and Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nat. Neurosci.* 3, 80–84. doi: 10.1038/71152

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

Hoshi, Y., Kobayashi, N., and Tamura, M. (2001). Interpretation of near-infrared spectroscopy signals: a study with a newly developed perfused rat brain model. *J. Appl. Physiol.* 90, 1657–1662. doi: 10.1152/jappl.2001.90.5.1657

Indefrey, P., and Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition* 92, 101–144. doi: 10.1016/j.cognition.2002.06.001

Johnstone, T., and Scherer, K. (2000). "Vocal communication of emotion" in *Handbook of emotions*. eds. M. Lewis and J. Haviland (New York, NY: Guilford), 220–235.

Jolles, D. D., Grol, M. J., Van Buchem, M. A., Rombouts, S. A., and Crone, E. A. (2010). Practice effects in the brain: changes in cerebral activation after working memory practice depend on task demands. *Neuroimage* 52, 658–668. doi: 10.1016/j.neuroimage.2010.04.028

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

Kane, M. J., Conway, A. R., Miura, T. K., and Colflesh, G. J. (2007). Working memory, attention control, and the N-back task: a question of construct validity. *J. Exp. Psychol. Learn. Mem. Cogn.* 33:615. doi: 10.1037/0278-7393.33.3.615

Kincade, J. M., Abrams, R. A., Astafiev, S. V., Shulman, G. L., and Corbetta, M. (2005). An event-related functional magnetic resonance imaging study of voluntary and stimulus-driven orienting of attention. *J. Neurosci.* 25, 4593–4604. doi: 10.1523/JNEUROSCI.0236-05.2005

Klem, G. H., Lüders, H. O., Jasper, H. H., and Elger, C. (1999). The ten-twenty electrode system of the international federation. International Federation of Clinical Neurophysiology. *Electroencephalogr. Clin. Neurophysiol.* 52, 3–6.

Koeda, M., Belin, P., Hama, T., Masuda, T., Matsuura, M., and Okubo, Y. (2013b). Cross-cultural differences in the processing of non-verbal affective vocalizations by Japanese and Canadian listeners. *Front. Psychol.* 4:105. doi: 10.3389/fpsyg.2013.00105

Koeda, M., Takahashi, H., Matsuura, M., Asai, K., and Okubo, Y. (2013a). Cerebral responses to vocal attractiveness and auditory hallucinations in schizophrenia: a functional MRI study. *Front. Hum. Neurosci.* 7:221. doi: 10.3389/fnhum.2013.00221

Kostopoulos, P., and Petrides, M. (2004). The role of the prefrontal cortex in the retrieval of mnemonic information. *Hell. J. Psychol.* 1, 247–267.

Langner, R., Kellermann, T., Boers, F., Sturm, W., Willmes, K., and Eickhoff, S. B. (2011). Modality-specific perceptual expectations selectively modulate baseline activity in auditory, somatosensory, and visual cortices. *Cerebellum* 21, 2850–2862. doi: 10.1093/cercor/bhr083

Langner, R., Kellermann, T., Eickhoff, S. B., Boers, F., Chatterjee, A., Willmes, K., et al. (2012). Staying responsive to the world: modality-specific and-nonspecific contributions to speeded auditory, tactile, and visual stimulus detection. *Hum. Brain Mapp.* 33, 398–418. doi: 10.1002/hbm.21220

Leibenluft, E., Gobbini, M. I., Harrison, T., and Haxby, J. V. (2004). Mothers' neural activation in response to pictures of their children and other children. *Biol. Psychiatry* 56, 225–232. doi: 10.1016/j.biopsych.2004.05.017

Lennox, B. R., Bert, S., Park, G., Jones, P. B., and Morris, P. G. (1999). Spatial and temporal mapping of neural activity associated with auditory hallucinations. *Lancet* 353:644. doi: 10.1016/S0140-6736(98)05923-6

Lennox, B. R., Park, S. B. G., Medley, I., Morris, P. G., and Jones, P. B. (2000). The functional anatomy of auditory hallucinations in schizophrenia. *Psychiatry Res.* 100, 13–20. doi: 10.1016/S0925-4927(00)00068-8

Li, L., Men, W. W., Chang, Y. K., Fan, M. X., Ji, L., and Wei, G. X. (2014). Acute aerobic exercise increases cortical activity during working memory: a functional MRI study in female college students. *PLoS One* 9:e99222. doi: 10.1371/journal.pone.0099222

Linden, D. E., Prvulovic, D., Formisano, E., Völlinger, M., Zanella, F. E., Goebel, R., et al. (1999). The functional neuroanatomy of target detection: an fMRI study of visual and auditory oddball tasks. *Cerebellum* 9, 815–823. doi: 10.1093/cercor/9.8.815

LoPresti, M. L., Schon, K., Tricarico, M. D., Swisher, J. D., Celone, K. A., and Stern, C. E. (2008). Working memory for social cues recruits orbitofrontal cortex and amygdala: a functional magnetic resonance imaging study of delayed matching to sample for emotional expressions. *J. Neurosci.* 28, 3718–3728. doi: 10.1523/JNEUROSCI.0464-08.2008

Macaluso, E., Eimer, M., Frith, C. D., and Driver, J. (2003). Preparatory states in crossmodal spatial attention: spatial specificity and possible control mechanisms. *Exp. Brain Res.* 149, 62–74. doi: 10.1007/s00221-002-1335-y

Maeder, P. P., Meuli, R. A., Adriani, M., Bellmann, A., Fornari, E., Thiran, J. P., et al. (2001). Distinct pathways involved in sound recognition and localization: a human fMRI study. *Neuroimage* 14, 802–816. doi: 10.1006/nimg.2001.0888

Maki, A., Yamashita, Y., Ito, Y., Watanabe, E., Mayanagi, Y., and Koizumi, H. (1995). Spatial and temporal analysis of human motor activity using noninvasive NIR topography. *Med. Phys.* 22, 1997–2005. doi: 10.1118/1.597496

Mayer, A. R., Harrington, D., Adair, J. C., and Lee, R. (2006). The neural networks underlying endogenous auditory covert orienting and reorienting. *Neuroimage* 30, 938–949. doi: 10.1016/j.neuroimage.2005.10.050

Morton, J., and Johnson, M. H. (1991). CONSPEC and CONLERN: a two-process theory of infant face recognition. *Psychol. Rev.* 98, 164–181. doi: 10.1037/0033-295X.98.2.164

Nakao, Y., Tamura, T., Kodabashi, A., Fujimoto, T., and Yarita, M. (2011). Working memory load related modulations of the oscillatory brain activity. N-back ERD/ERS study. *Seitai Ikougaku* 49, 850–857.

Neta, M., Davis, F. C., and Whalen, P. J. (2011). Valence resolution of ambiguous facial expressions using an emotional oddball task. *Emotion* 11, 1425–1433. doi: 10.1037/a0022993

Neta, M., and Whalen, P. J. (2011). Individual differences in neural activity during a facial expression vs. identity working memory task. *Neuroimage* 56, 1685–1692. doi: 10.1016/j.neuroimage.2011.02.051

Okamoto, M., and Dan, I. (2005). Automated cortical projection of head-surface locations for transcranial functional brain mapping. *Neuroimage* 26, 18–28. doi: 10.1016/j.neuroimage.2005.01.018

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4

Onitsuka, T., Shenton, M. E., Salisbury, D. F., Dickey, C. C., Kasai, K., Toner, S. K., et al. (2004). Middle and inferior temporal gyrus gray matter volume abnormalities in chronic schizophrenia: an MRI study. *Am. J. Psychiatry* 161, 1603–1611. doi: 10.1176/appi.ajp.161.9.1603

Owen, A. M., McMillan, K. M., Laird, A. R., and Bullmore, E. (2005). N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. *Hum. Brain Mapp.* 25, 46–59. doi: 10.1002/hbm.20131

Patterson, K., Nestor, P. J., and Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat. Rev. Neurosci.* 8, 976–987. doi: 10.1038/nrn2277

Pessoa, L. (2010). Emotion and cognition and the amygdala: from "what is it?" to "what's to be done?". *Neuropsychologia* 48, 3416–3429. doi: 10.1016/j.neuropsychologia.2010.06.038

Phelps, E. A. (2006). Emotion and cognition: insights from studies of the human amygdala. *Annu. Rev. Psychol.* 57, 27–53. doi: 10.1146/annurev.psych.56.091103.070234

Posner, M. I., and Petersen, S. E. (1990). The attention system of the human brain. *Annu. Rev. Neurosci.* 13, 25–42. doi: 10.1146/annurev.ne.13.030190.000325

Rogers, M. A., Kasai, K., Koji, M., Fukuda, R., Iwanami, A., Nakagome, K., et al. (2004). Executive and prefrontal dysfunction in unipolar depression: a review of neuropsychological and imaging evidence. *Neurosci. Res.* 50, 1–11. doi: 10.1016/j.neures.2004.05.003

Rorden, C., and Brett, M. (2000). Stereotaxic display of brain lesions. *Behav. Neurol.* 12, 191–200. doi: 10.1155/2000/421719

Shattuck, D. W., Mirza, M., Adisetiyo, V., Hojatkashani, C., Salamon, G., Narr, K. L., et al. (2008). Construction of a 3D probabilistic atlas of human cortical structures. *Neuroimage* 39, 1064–1080. doi: 10.1016/j.neuroimage.2007.09.031

Shiraishi, K., and Kanda, Y. (2010). Measurements of the equivalent continuous sound pressure level and equivalent A-weighted continuous sound pressure level during conversational Japanese speech. *Audiol. Japan* 53, 199–207. doi: 10.4295/audiology.53.199

Shomstein, S., and Yantis, S. (2006). Parietal cortex mediates voluntary control of spatial and nonspatial auditory attention. *J. Neurosci.* 26, 435–439. doi: 10.1523/JNEUROSCI.4408-05.2006

Singh, A. K., Okamoto, M., Dan, H., Jurcak, V., and Dan, I. (2005). Spatial registration of multichannel multi-subject fNIRS data to MNI space without MRI. *Neuroimage* 27, 842–851. doi: 10.1016/j.neuroimage.2005.05.019

Spence, C., and Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Percept. Psychophys.* 59, 1–22. doi: 10.3758/BF03206843

Sridharan, D., Levitin, D. J., Chafe, C. H., Berger, J., and Menon, V. (2007). Neural dynamics of event segmentation in music: converging evidence for dissociable ventral and dorsal networks. *Neuron* 55, 521–532. doi: 10.1016/j.neuron.2007.07.003

Strangman, G., Culver, J. P., Thompson, J. H., and Boas, D. A. (2002). A quantitative comparison of simultaneous BOLD fMRI and NIRS recordings during functional brain activation. *Neuroimage* 17, 719–731. doi: 10.1006/nimg.2002.1227

Sutoko, S., Sato, H., Maki, A., Kiguchi, M., Hirabayashi, Y., Atsumori, H., et al. (2016). Tutorial on platform for optical topography analysis tools. *Neurophotonics* 3:010801. doi: 10.1117/1.NPh.3.1.010801

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., and Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proc. Natl. Acad. Sci. U. S. A.* 94, 14792–14797. doi: 10.1073/pnas.94.26.14792

Tsuzuki, D., and Dan, I. (2014). Spatial registration for functional near-infrared spectroscopy: from channel position on the scalp to cortical location in individual and group analyses. *Neuroimage* 85, 92–103. doi: 10.1016/j.neuroimage.2013.07.025

Tsuzuki, D., Jurcak, V., Singh, A. K., Okamoto, M., Watanabe, E., and Dan, I. (2007). Virtual spatial registration of stand-alone fNIRS data to MNI space. *Neuroimage* 34, 1506–1518. doi: 10.1016/j.neuroimage.2006.10.043

Uga, M., Dan, I., Sano, T., Dan, H., and Watanabe, E. (2014). Optimizing the general linear model for functional near-infrared spectroscopy: an adaptive hemodynamic response function approach. *Neurophotonics* 1:015004. doi: 10.1117/1.NPh.1.1.015004

Valenza, E., Simion, F., Cassia, V. M., and Umiltà, C. (1996). Face preference at birth. *J. Exp. Psychol. Hum. Percept.* 22:892. doi: 10.1037/0096-1523.22.4.892

Vandenberghe, R., Price, C., Wise, R., Josephs, O., and Frackowiak, R. S. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature* 383, 254–256. doi: 10.1038/383254a0

Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houde, O., et al. (2006). Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *Neuroimage* 30, 1414–1432. doi: 10.1016/j.neuroimage.2005.11.002

Vines, B. W., Schnider, N. M., and Schlaug, G. (2006). Testing for causality with transcranial direct current stimulation: pitch memory and the left supramarginal gyrus. *Neuroreport* 17, 1047–1050. doi: 10.1097/01.wnr.0000223396.05070.a2

Visser, M., Jefferies, E., Embleton, K. V., and Lambon Ralph, M. A. (2012). Both the middle temporal gyrus and the ventral anterior temporal area are crucial for multimodal semantic processing: distortion-corrected fMRI evidence for a double gradient of information convergence in the temporal lobes. *J. Cogn. Neurosci.* 24, 1766–1778. doi: 10.1162/jocn_a_00244

Vossel, S., Thiel, C. M., and Fink, G. R. (2006). Cue validity modulates the neural correlates of covert endogenous orienting of attention in parietal and frontal cortex. *Neuroimage* 32, 1257–1264. doi: 10.1016/j.neuroimage.2006.05.019

Wagner, A. D. (1999). Working memory contributions to human learning and remembering. *Neuron* 22, 19–22. doi: 10.1016/S0896-6273(00)80674-1

Wagner, A. D., Paré-Blagoev, E. J., Clark, J., and Poldrack, R. A. (2001). Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. *Neuron* 31, 329–338. doi: 10.1016/S0896-6273(01)00359-2

Weiner, K. S., and Grill-Spector, K. (2010). Sparsely-distributed organization of face and limb activations in human ventral temporal cortex. *Neuroimage* 52, 1559–1573. doi: 10.1016/j.neuroimage.2010.04.262

Whitney, C., Jefferies, E., and Kircher, T. (2011a). Heterogeneity of the left temporal lobe in semantic representation and control: priming multiple versus single meanings of ambiguous words. *Cerebellum* 21, 831–844. doi: 10.1093/cercor/bhq148

Whitney, C., Kirk, M., O'Sullivan, J., Lambon Ralph, M. A., and Jefferies, E. (2011b). The neural organization of semantic control: TMS evidence for a distributed network in left inferior frontal and posterior middle temporal gyrus. *Cerebellum* 21, 1066–1075. doi: 10.1093/cercor/bhq180

Woodruff, P. W., McManus, I. C., and David, A. S. (1995). Meta-analysis of corpus callosum size in schizophrenia. *J. Neurol. Neurosurg. Psychiatry* 58, 457–461. doi: 10.1136/jnnp.58.4.457

Wu, C. T., Weissman, D. H., Roberts, K. C., and Woldorff, M. G. (2007). The neural circuitry underlying the executive control of auditory spatial attention. *Brain Res.* 1134, 187–198. doi: 10.1016/j.brainres.2006.11.088

Xin, F., and Lei, X. (2015). Competition between frontoparietal control and default networks supports social working memory and empathy. *Soc. Cogn. Affect. Neurosci.* 10, 1144–1152. doi: 10.1093/scan/nsu160

Yang, Y. H., and Chen, H. H. (2012). Machine recognition of music emotion: a review. *ACM Trans. Intell. Syst. Technol.* 3, 1–30. doi: 10.1145/2168752.2168754

Ye, Z., and Zhou, X. (2009). Conflict control during sentence comprehension: fMRI evidence. *Neuroimage* 48, 280–290. doi: 10.1016/j.neuroimage.2009.06.032

Yen, J. Y., Chang, S. J., Long, C. Y., Tang, T. C., Chen, C. C., and Yen, C. F. (2012). Working memory deficit in premenstrual dysphoric disorder and its associations with difficulty in concentrating and irritability. *Compr. Psychiatry* 53, 540–545. doi: 10.1016/j.comppsych.2011.05.016

Zhou, C., Song, X., and Ye, C. (2021). The inhibition of high load task on individual cognition: a functional near-infrared spectroscope study. *Res. Sq.* doi: 10.21203/rs.3.rs-494113/v1

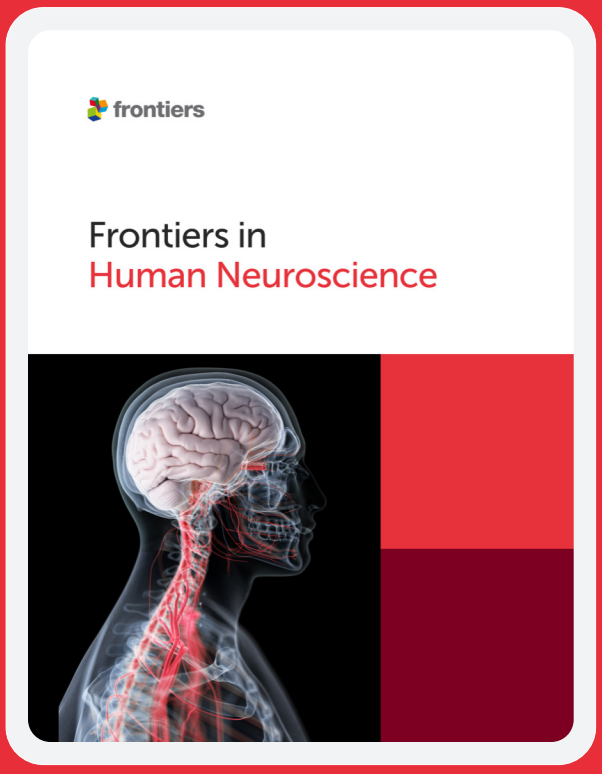
Frontiers in
Human Neuroscience

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