TASK-RELATED BRAIN SYSTEMS REVEALED BY HUMAN IMAGING EXPERIMENTS

EDITED BY: Kuniyoshi L. Sakai and Yuji Naya

PUBLISHED IN: Frontiers in Behavioral Neuroscience and
Frontiers in Human Neuroscience

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88976-125-8
DOI 10.3389/978-2-88976-125-8

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

TASK-RELATED BRAIN SYSTEMS REVEALED BY HUMAN IMAGING EXPERIMENTS

Topic Editors:
Kuniyoshi L. Sakai, The University of Tokyo, Japan
Yuji Naya, Peking University, China

Citation: Sakai, K. L., Naya, Y., eds. (2022). Task-Related Brain Systems Revealed by Human Imaging Experiments. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88976-125-8

Table of Contents

frontiers | Frontiers in Behavioral Neuroscience

EDITORIAL
published: 19 April 2022
doi: 10.3389/fnbeh.2022.889486

Editorial: Task-Related Brain Systems Revealed by Human Imaging Experiments

Yuji Naya[1,2,3]* *and Kuniyoshi L. Sakai*[4]*

[1] School of Psychological and Cognitive Sciences, Peking University, Beijing, China, [2] IDG/McGovern Institute for Brain Research at Peking University, Beijing, China, [3] Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing, China, [4] Department of Basic Science, Graduate School of Arts and Sciences, The University of Tokyo, Tokyo, Japan

Keywords: neuroimaging (functional), human, non-human primates, task design, brain systems, language, cognition

Editorial on the Research Topic

Task-Related Brain Systems Revealed by Human Imaging Experiments

More than three decades have passed since the development of functional magnetic resonance imaging (fMRI), a non-invasive neuroimaging technique that allows us to look into neural activity of the human brain measured by local blood oxygenation level-dependent (BOLD) signals (Ogawa et al., 1990). Progress in neuroimaging studies has clarified a number of brain systems that are critical in higher cognitive functions, including learning and memory. Although "default mode networks" have been assessed without using any tasks, task-design development is still vitally important to reveal the specific brain networks responsible for individual cognitive functions. Therefore, in this Research Topic our goal was to address how task designs for cognitive neuroscience can be advanced, and which specific questions about cognitive functions can be addressed by neuroimaging approaches.

Considering the accumulation of tasks reported in previous neuroimaging and psychological studies, it would be a good start to utilize well-established tasks (e.g., a delayed matching-to-sample task) and combine those tasks with new stimuli and/or stimulus presentation conditions. For example, Zhou et al. used a "dual-feature delayed matching task" to examine the neural mechanisms underlying an attentional function. In their task, the participants attended to either the color or shape of stimuli, and the trans-magnetic stimulation to the right posterior superior temporal sulcus facilitated feature discrimination. In another study, Tsuruha and Tsukiura used a word-face association memory task, where face stimuli were categorized into two age-groups. They examined the effects of in-group (participants with ages close to those of the stimuli) and out-group members (participants with ages different from those of the stimuli) on the neural mechanisms underlying the recollection of association memory. Regarding stimulus presentation conditions, Chen and Naya took over a series of studies (Chen and Naya, 2020a,b) to examine a scene perception using a delayed matching task, in which the identity of an object and its location were encoded under two conditions: a foveal-view (F-V) and a peripheral-view (P-V). Under the F-V condition, the location information of an object was obtained as a gaze position, while under the P-V condition, that information was obtained as a peripheral retinotopic position. In an electrophysiological study of non-human primates, they found robust location signals in the ventral visual pathway, as well as an integration of object and location information in the medial temporal lobe only under the F-V condition.

It is worth noting that task conditions can change significantly according to the status or training of participants. Khaksari et al. used a motor task under either a self-action (actor) or an observation

Edited and reviewed by:
Denise Manahan-Vaughan, Ruhr University Bochum, Germany

***Correspondence:**
Yuji Naya
yujin@pku.edu.cn
Kuniyoshi L. Sakai
sakai@sakai-lab.jp

Specialty section:
This article was submitted to Learning and Memory, a section of the journal Frontiers in Behavioral Neuroscience

Received: *04 March 2022*
Accepted: *31 March 2022*
Published: *19 April 2022*

Citation:
Naya Y and Sakai KL (2022) Editorial: Task-Related Brain Systems Revealed by Human Imaging Experiments. Front. Behav. Neurosci. 16:889486. doi: 10.3389/fnbeh.2022.889486

(observer) condition, and observed brain lateralization when participants were actors. On the other hand, Ohbayashi conducted a series of studies to examine motor learning in non-human primates using two sequential reaching tasks (one visually-guided [random] and the other memory-guided [repeating]), while, respectively, inactivating the corresponding motor areas (Ohbayashi et al., 2016; Ohbayashi, 2020). In a subsequent article, Ohbayashi discussed the distinct roles of the motor areas, especially the dorsal premotor cortex and primary motor cortex, on the effects of training over 100 daily sessions. Although such repetitive and intensive training is usually difficult to study in human participants, Sakai et al. focused on second language acquisition in visitors to Japan over the course of several *months*, and succeeded in revealing functional changes in both modality-dependent networks and domain-special language areas. Moreover, these cortical regions were found to be selectively recruited for specified music processes (pitch, tempo, stress, and articulation) after several *years* of musical instrument training (Sakai et al., 2021).

To investigate neurological symptoms, some psychological tasks designed for patients have been tested in non-human primates as an animal model. Misonou and Jimura reversed this common procedure, and tested decision making processes (an immediate small reward vs. a delayed large reward) in human participants given a liquid supply like that used in monkey experiments. Another direction for a new task paradigm would be the use of multi-voxel pattern analyses (MVPA), with which residual bottom-up or top-down signals can be subtracted out from the original signals in each brain region. Pham et al. compared a visual perception task and a visual imagery task, which involved more salient bottom-up and top-down signals, respectively. Yuen et al. examined attentional effects on brain activity during a driving simulation task, suggesting the importance of oculomotor behavior.

It would be also interesting to conduct human neuroimaging studies in which our daily lives are represented or simulated using virtual-reality techniques or natural methods that avoid artificial rule learning. Umejima et al. examined the effects of the use of either paper notebooks or mobile devices on a subsequent memory recall, and found enhanced activations in the hippocampus, visual cortices, and language-related frontal regions for the group using paper notebooks. Moreover, during the natural acquisition of a new language, activations in the bilateral frontal/temporal regions were maintained at a higher level than the initial level during subsequent new grammar conditions for multilinguals (Umejima et al., 2021). These results suggest that individual brain networks become increasingly specialized and intricate to adapt to a constantly changing outer world.

Overall, the above-mentioned findings indicate that hypothesis-driven or top-down approaches are crucial in cognitive or systems neuroscience, together with insights into experimental design. This is why a sophisticated task is required for human neuroimaging studies, especially when studying functions such as cognition, thinking, and language. Such sophistication of task would also be crucial for electrophysiological/imaging studies of non-human primates, which contribute to our understanding of basic brain functions. The above-mentioned findings also open up new and attractive questions about human mind, which could be addressed in future research with much improved and sophisticated task designs.

AUTHOR CONTRIBUTIONS

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

REFERENCES

Chen, H., and Naya, Y. (2020a). Forward processing of object-location association from the ventral stream to medial temporal lobe in nonhuman primates. *Cereb. Cortex* 30, 1260–1271. doi: 10.1093/cercor/bhz164

Chen, H., and Naya, Y. (2020b). Automatic encoding of a view-centered background image in the macaque temporal lobe. *Cereb. Cortex*. 30, 6270–6283. doi: 10.1101/2020.03.01.971507

Ogawa, S., Lee, T. M., Kay, A. R., and Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc. Natl. Acad. Sci. U.S.A.* 87, 9868–9872. doi: 10.1073/pnas.87.24.9868

Ohbayashi, M. (2020). Inhibition of protein synthesis in M1 of monkeys disrupts performance of sequential movements guided by memory. *eLife* 9, e53038. doi: 10.7554/eLife.53038

Ohbayashi, M., Picard, N., and Strick, P. L. (2016). Inactivation of the dorsal premotor area disrupts internally generated, but not visually guided, sequential movements. *J. Neurosci.* 36, 1971–1976. doi: 10.1523/JNEUROSCI.2356-15

Sakai, K. L., Oshiba, Y., Horisawa, R., Miyamae, T., and Hayano, R. (2021). Music-experience-related and musical-error-dependent activations in the brain. *Cereb. Cortex.* doi: 10.1093/cercor/bhab478. [Epub ahead of print].

Umejima, K., Flynn, S., and Sakai, K. L. (2021). Enhanced activations in syntax-related regions for multilinguals while acquiring a new language. *Sci. Rep.* 11, 7296. doi: 10.1038/s41598-021-86710-4

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 in Behavioral Neuroscience

ORIGINAL RESEARCH
published: 19 March 2021
doi: 10.3389/fnbeh.2021.634158

Paper Notebooks vs. Mobile Devices: Brain Activation Differences During Memory Retrieval

*Keita Umejima[1], Takuya Ibaraki[2], Takahiro Yamazaki[2] and Kuniyoshi L. Sakai[1]**

[1]Department of Basic Science, Graduate School of Arts and Sciences, The University of Tokyo, Tokyo, Japan, [2]NTT Data Institute of Management Consulting, Inc., Tokyo, Japan

It remains to be determined how different inputs for memory-encoding, such as the use of paper notebooks or mobile devices, affect retrieval processes. We compared three groups of participants who read dialogues on personal schedules and wrote down the scheduled appointments on a calendar using a paper notebook (Note), an electronic tablet (Tablet), or a smartphone (Phone). After the retention period for an hour including an interference task, we tested recognition memory of those appointments with visually presented questions in a retrieval task, while scanned with functional magnetic resonance imaging. We obtained three major results. First, the duration of writing down schedules was significantly shorter for the Note group than the Tablet and Phone groups, and accuracy was much higher for the Note group in easier (i.e., more straightforward) questions. Because the input methods were equated as much as possible between the Note and Tablet groups, these results indicate that the cognitive processes for the Note group were deeper and more solid. Second, brain activations for all participants during the retrieval phase were localized in the bilateral hippocampus, precuneus, visual cortices, and language-related frontal regions, confirming the involvement of verbalized memory retrieval processes for appointments. Third, activations in these regions were significantly higher for the Note group than those for the Tablet and Phone groups. These enhanced activations for the Note group could not be explained by general cognitive loads or task difficulty, because overall task performances were similar among the groups. The significant superiority in both accuracy and activations for the Note group suggested that the use of a paper notebook promoted the acquisition of rich encoding information and/or spatial information of real papers and that this information could be utilized as effective retrieval clues, leading to higher activations in these specific regions.

Keywords: memory encoding, memory retrieval, hippocampus, language, fMRI

OPEN ACCESS

Edited by:
Malgorzata Maria Kossut, Nencki Institute of Experimental Biology (PAS), Poland

Reviewed by:
Emiliano Macaluso, Université Claude Bernard Lyon 1, France
Aneta Brzezicka, University of Social Sciences and Humanities, Poland

****Correspondence:***
Kuniyoshi L. Sakai
sakai@sakai-lab.jp

Specialty section:
This article was submitted to Learning and Memory, a section of the journal Frontiers in Behavioral Neuroscience

Received: *27 November 2020*
Accepted: *15 February 2021*
Published: *19 March 2021*

Citation:
Umejima K, Ibaraki T, Yamazaki T and Sakai KL (2021) Paper Notebooks vs. Mobile Devices: Brain Activation Differences During Memory Retrieval. Front. Behav. Neurosci. 15:634158. doi: 10.3389/fnbeh.2021.634158

INTRODUCTION

The properties of human memory have been investigated with several approaches, including clinical, psychological, and neuroimaging studies (Tulving, 2002; Schacter et al., 2007; Miyashita, 2019). It remains to be elucidated how brain activations during retrieval processes are modulated by different encoding procedures, because it has been

reported that retrieval performances on paired words became worse when the categorically similar target words were simultaneously encoded, suggesting the importance of the context-dependent encoding (Nairne, 2002; Goh and Lu, 2012). It is also possible that the manner with which specific information is encoded—e.g., whether by using a paper notebook, computer, or mobile device—may affect retrieval processes. A recent behavioral study showed that students who took longhand notes performed better on conceptual questions than those who took notes on laptop computers (Mueller and Oppenheimer, 2014). A reasonable explanation for this interesting finding would be that the use of a paper notebook enables users to summarize and reframe information in their own words for encoding, while the use of a laptop tends to encourage them to write down information more passively (i.e., more nearly verbatim). The former processes thus naturally ensure deeper and more solid encoding *via* the active process of making notes. Moreover, it has been reported that longhand note-taking enhanced the performance of students on recognition of memorized words, even though typing on a computer keyboard allowed greater speed (Aragón-Mendizábal et al., 2016).

Another possible explanation for the superiority of longhand note-taking for conceptual understanding is related to the use of paper for writing/reading since a behavioral study reported the superiority of paper to computer screens in terms of reading comprehension (Wästlund et al., 2005; Mangen et al., 2013). These studies indicated the importance of visual and tactile cues for perceiving *constant* physical sizes and spatial locations, because "the material substrate of paper provides physical, tactile, spatiotemporally fixed cues to the length of the text" (Mangen et al., 2013). We hypothesized that the use of a paper notebook, together with longhand note-taking, would enhance both memory encoding and later retrieval processes that could then be investigated at the brain level. More specifically, the utilization of the paper likely enhances the processes of associating episodic (*what*) and spatial (*where*) information, especially in the hippocampus, given its well-established role in the integration of what/where/when information (Broadbent et al., 2004; Eichenbaum, 2004; Chadwick et al., 2010).

To address this issue, we compared three groups of participants who used a paper notebook (Note), electronic tablet (Tablet), or smartphone (Phone) during the encoding phase. Participants in the Tablet group used a stylus pen, thereby controlling for the effects of longhand writing with a pen in the Note group. It should be noted that physical sizes and spatial locations of a document remain constant for a paper notebook, whereas they become variable on the display of a tablet or smartphone. Moreover, not only the physical interaction of the hand with the pen/paper during note-taking but the actual writing of notes relative to each page of the real paper provides more concrete encoding information, because that information can be easily erased and updated by new information on the physically same screen of a tablet or smartphone.

We asked participants to write down scheduled appointments, and then, after one hour during which they performed an interference task, we conducted a retrieval task in which we tested participants' recognition memory of those appointments (**Figure 1**). We further hypothesized that the interaction with physical paper, rather than the mental editing/preparation of the notes or the physical act of handwriting, provides episodic and spatial information of notes relative to each page of real paper, together with visual/tactile information from the paper. These properties and cues of papers could help to retrieve specific information, and thus lead to increased activations in specified brain regions for the Note group, compared with the other groups using mobile devices lacking such processes.

It has been proposed that the hippocampus and the prefrontal cortex support complementary functions in episodic memory and that the bidirectional information flow between these regions may play a crucial role in integrating and consolidating individual information (Moscovitch et al., 2016; Eichenbaum, 2017). A previous functional magnetic resonance imaging (fMRI) study reported that episodic memory of a word or picture is related to a functional network that includes the left posterior precuneus and the left lateral prefrontal cortex (Lundstrom et al., 2003). On the other hand, language function is critically involved in human episodic memory, and some language-related regions would be recruited during both memory encoding and retrieval. The left lateral premotor cortex (LPMC) and left opercular/triangular parts of the inferior frontal gyrus (F3op/F3t) are suggested to have central roles in syntactic processing, whereas the left angular/supramarginal gyri (AG/SMG) make a major contribution to lexical processing (Sakai, 2005). Moreover, the right frontal cortex was identified as a supportive region for syntactic processing (Kinno et al., 2014). Activations in these regions would be observed during memory retrieval because fMRI studies showed that the hippocampus and language-related regions involved in the encoding phase were also activated during the retrieval phase (Rugg et al., 2008). The retrieval task we used critically involved episodic memory of scheduled appointments, and thus activations in these regions would be increased more for the Note group than the other groups.

MATERIALS AND METHODS

Participants

University student volunteers (48 native Japanese speakers, 18 females) aged 18–29 years were openly recruited from multiple sources, including the University of Tokyo and Sophia University, as well as the participant pool of the NTT Data Institute of Management Consulting. The laterality quotient (LQ) was measured according to the Edinburgh inventory (Oldfield, 1971); all participants but one were right-handed, and the exception was both-handed (LQ: −14). As stated above, the participants were divided into three groups: Note, Tablet, and Phone groups (**Table 1**). These three groups were age- and LQ-matched (Kruskal–Wallis test, $p > 0.1$), as well as gender-matched (Fisher's exact test for count data, $p = 0.17$). Each participant first answered a questionnaire

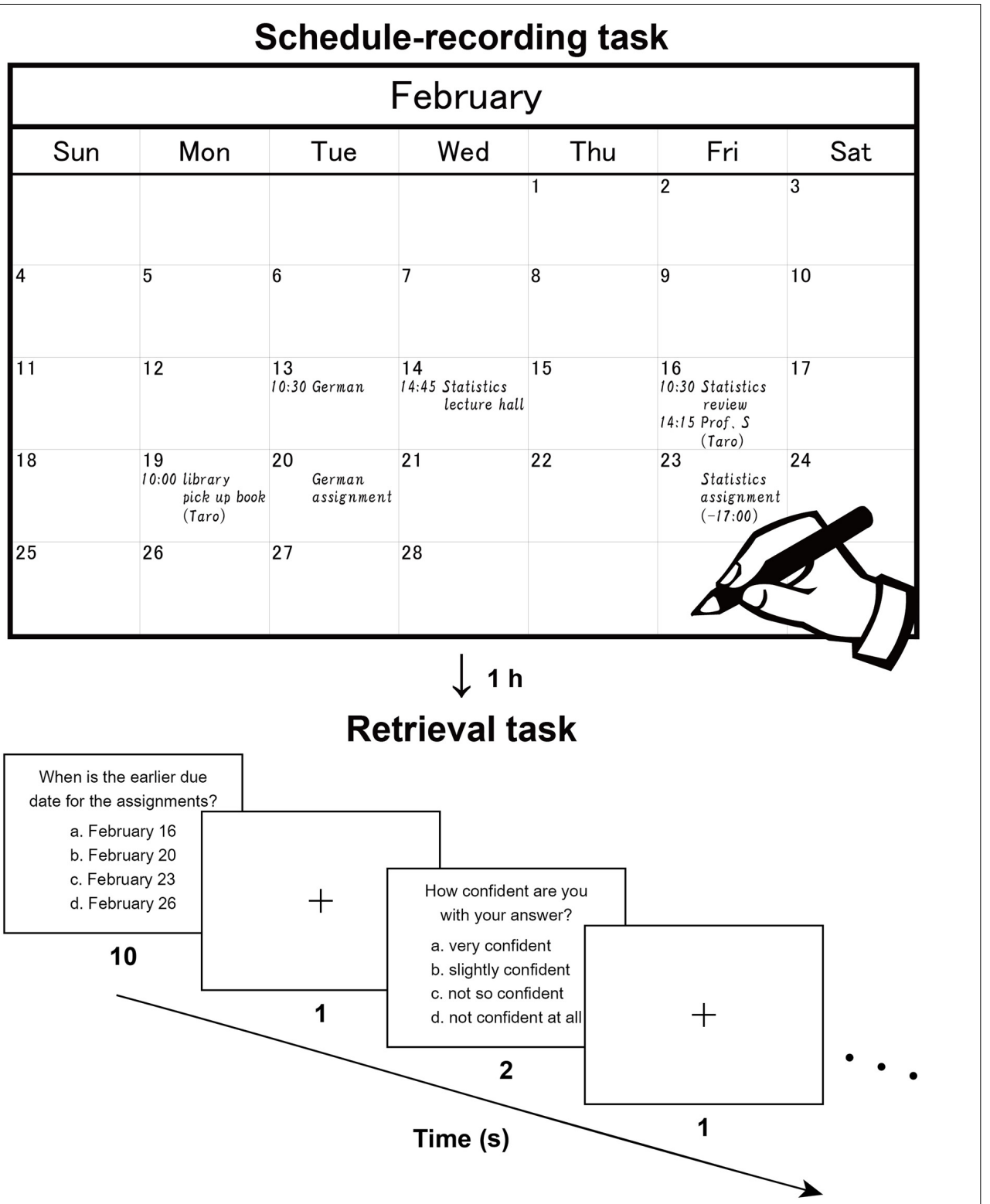

FIGURE 1 | Recording and retrieval of schedule information. Participants first read dialogues (in Japanese), then extracted scheduled appointments contained in the dialogues, and wrote them down with a paper notebook (*Note* group), electronic tablet (*Tablet* group), or smartphone (*Phone* group). This procedure reproduces the daily making of to-do lists and naturally involves encoding processes. The upper panel shows a typical example (English translation) written by a participant. After an hour including an interference task, the participants were asked to answer questions about the appointments and reported their level of confidence in their answer to each question. The lower panel shows a typical trial in this retrieval task.

TABLE 1 | Basic data on participants.

Experimental groups	Number of participants	Age (year)	LQ	The maximum length of memorized sequences
Paper notebook (Note)	16 (8)	20.8 ± 1.6 (18.7–24.1)	90 ± 10 (71–100)	6.5 ± 1.2 (4–8)
Electronic tablet (Tablet)	16 (3)	20.1 ± 1.0 (18.7–22.4)	88 ± 12 (65–100)	6.4 ± 1.0 (5–8)
Smartphone (Phone)	16 (7)	21.8 ± 2.6 (19.0–27.7)	83 ± 28 (−14–100)	6.9 ± 1.1 (4–8)

For the number of participants, numbers of females are shown in parentheses. For age, laterality quotient (LQ), and the maximum length of memorized sequences, averaged data (mean ± standard deviation) and their ranges (in parentheses) are shown.

on their daily use of paper notebooks, electronic tablets, and smartphones for scheduling in an academic or personal context (seven-point scale for each). Based on this result, electronic tablet users were assigned to the Tablet group, and smartphone users (those on the highest scale for smartphone use) were assigned to either the Tablet or Phone group. To estimate short-term memory ability, we used the number-letter sequencing in the Wechsler Adult Intelligence Scale—Fourth Edition (Drozdick et al., 2012), and the maximum length of memorized sequences was not significantly different among the three groups (p = 0.4). All participants in the Note group used paper notebooks for daily schedule management, whereas eight and seven participants in the Tablet and Phone groups, respectively, also used paper notebooks for that purpose. To control the experience and accustomedness of using paper notebooks for daily schedule management, these 15 participants (with eight females) were separately designated the *Device* group, which was used in behavioral and activation analyses.

Before they participated in the study, the nature and possible consequences of the studies were explained to each participant and written informed consent was obtained afterward. None of the participants had a history of neurological or psychological disorders. Approval for the experiments was obtained from the institutional review board of the University of Tokyo, Komaba.

Stimuli and Tasks

Two sets of written dialogues between two or three persons (a set of dialogues on academic matters and a set on personal matters) were presented to the participants, who were asked to imagine that they were participating in those dialogues. There were seven daily scheduled appointments for the academic context and seven for the personal context (in February and March, respectively). While silently reading the dialogues, participants were asked to enter each of these appointments into a monthly calendar (**Figure 1**, upper panel). The participants used either a paper notebook [Noritsu NOLTY Notebook (2017), size 20.6 × 17.6 cm^2 when opened], an electronic tablet [iPad Pro 10.5 inch (2017), screen size 21.4 × 16.1 cm^2 in landscape orientation], or a smartphone (Google Nexus 5 LG-D821, screen size 6.2 × 10.9 cm^2 in portrait orientation), where the paper notebook and electronic tablet were similar in physical layout (size and orientation). All three types of calendars had a day, week, and month view, but we used only the month view. In the case of the paper notebook and electronic tablet, appointments could only be viewed and edited individually in the relevant month (i.e., discrete views). In the smartphone, individual weeks could be viewed and edited by swiping continuously (i.e., continuous views). This difference was notable, in that schedule information would be encoded relative to the *spatial configuration* of one month (see **Figure 1**) for the paper notebook and electronic tablet.

Regarding input methods, a four-color pen was used to write in the paper notebook [the use of color(s) was up to each participant], and a stylus pen was used to write on the electronic tablet with a free choice of multiple colors (without using a virtual keyboard). In the case of the smartphone, the text was written by either flick input with the finger(s) or by using a virtual keyboard. In Japanese, there are three types of characters (*hiragana*, *katakana*, and *kanji*; kanji basically consists of Chinese characters), and kana-kanji transformation is usually used for inputs in mobile devices and computers (kana-kanji transformation converts a limited number of hiragana to vast numbers of kanji by requiring users to select appropriate kanji from multiple candidates). The flick input utilizes a telephone keypad with a three by four layout, and one hiragana character can be selected by either tapping a keypad or flicking from a keypad to one of four directions (up, down, left, or right) to enter one of five hiragana characters sharing the same initial consonant.

We measured the time required by participants to write down the appointments, but we set no time limit. When the participants finished writing down, they were instructed to review the calendar for 30 s. Then, after the retention period for an hour including an interference task, participants were asked to recall those appointments in a retrieval task; the experimental purpose of writing down the appointments was not disclosed to them. The interference task involved listening comprehension; participants were informed that they would hear a story, and then be asked about its contents while lying in an MRI scanner. We used the first 6 min of a narrated version of a Japanese classic short story called "*Ma-jutsu* (*Magic*)" (written by Ryūnosuke Akutagawa, narrated by Takeshi Sasaki, and published by Pan Rolling, Japan). This story was unfamiliar to all participants. The auditory stimuli were presented through a headphone and participants were not permitted to take notes while listening. Sixteen questions about the detailed contents of the story were displayed inside the scanner (two questions per run), and the participants pressed one of four buttons to select the right answer.

After a short break outside the scanner to adjust the time between the encoding and retrieval phases to 1 h, participants performed a retrieval task inside the scanner (**Figure 1**, lower panel), in which 16 questions about detailed contents of the

appointments were displayed (two questions per run). Out of the 16 questions, seven required recalling of the relationships between multiple appointments, one required the conversion from the date to the day of the week (using the spatial information of the calendar), and three required recalling from similar or confusing appointments. The remaining five questions were more straightforward and thus considered as the *easier* questions. In each trial, a question was presented with four choices, and the participants pressed a button to select the right answer within 10 s. After an interval of 1 s, participants reported their level of confidence (1–4 scale, 4 = very confident) for that answer by pressing one of four buttons within 2 s. These responses were used to assess the correctness of each participant's self-evaluation, where the true positive rate vs. the false positive rate was plotted for each of the four levels of confidence. By connecting these plots, we obtained a receiver operating characteristics curve (Fawcett, 2006), and we used the area under the curve (AUC) for this assessment (0 = perfectly wrong; 0.5 = no distinction; 1 = perfectly correct).

As a control condition, we added a 2-back task into the run with the retrieval task. In each trial of the 2-back task, two different non-words, each with three Japanese characters, were sequentially displayed (each for 2 s). These characters were randomly selected from those used in the retrieval tasks, where the same type of characters (either hiragana, katakana, or kanji) was presented in a block of trials. Then four choices were shown for 5 s with a new non-word to be remembered. The correct answer was the non-word that appeared 2-back before but in a different order of three characters. There were two to four continuous trials with button pressings in each block.

Each run consisted of three 2-back blocks and two retrieval task trials, in which a 2-back block always started first, and the 2-back blocks and retrieval task trials were alternated. As fMRI events, we estimated the 6-s memory retrieval phase [determined by response times (RTs)] and the subsequent 4-s *post hoc* period from each 10-s period of the retrieval task, as well as a 5-s event for the 2-back task. With regards to contrasts between events, we always applied an exclusive mask of negative activations for the control conditions (one-sample *t*-test, uncorrected $p < 0.05$). During the scans, the participants wore earplugs and an eyeglass-like MRI-compatible display (resolution = 800 × 600 pixels, framerate = 60 fps; VisuaStim Digital, Resonance Technology Inc., Northridge, CA, USA). The stimuli were all presented in yellow letters on a black background. For fixation, a small red cross was shown at the center of the screen when a stimulus was not shown. The stimulus presentation and collection of behavioral data (accuracy and RTs) were controlled using the Presentation software package (Neurobehavioral Systems, Albany, CA, USA).

MRI Data Acquisition

The MRI scans were conducted in a 3.0 T scanner (Signa HDxt; GE Healthcare, Milwaukee, WI, USA) with a bird-cage head coil. Each participant was in a supine position, and his or her head was immobilized inside the coil. As regards the structural images, high-resolution T1-weighted images of the whole brain (136 axial slices, 1 × 1 × 1 mm^3) were acquired with a three-dimensional fast spoiled gradient-echo (3D FSPGR) acquisition [repetition time (TR) = 8.6 ms, echo time (TE) = 2.6 ms, flip angle (FA) = 25°, field of view (FOV) = 256 × 256 mm^2]. With respect to the time-series data of fMRI, we used a gradient-echo echo-planar imaging (EPI) sequence (TR = 2 s, TE = 30 ms, FA = 78°, FOV = 192 × 192 mm^2, resolution = 3 × 3 mm^2). We scanned a set of 30 axial slices that were 3-mm thick with a 0.5-mm gap, covering the range of −38.5 to 66 mm from the line of the anterior commissure to the posterior commissure (AC-PC). In a single scanning run, we obtained 45 volumes and dropped the initial four volumes from analyses due to MR signal increases.

fMRI Data Analyses

The fMRI data were analyzed in a standard manner using SPM12 statistical parametric mapping software (Wellcome Trust Center for Neuroimaging[1]; Friston et al., 1994) implemented on MATLAB (Math Works, Natick, MA, USA). The acquisition timing of each slice was corrected using the middle slice (the 15th slice chronologically) as a reference for the functional images. We spatially realigned each volume to the first volume of consecutive runs, and a mean volume was obtained. We set the threshold of head movement during a single run as follows: within a displacement of 2 mm in any of the three directions, and a rotation of 1.4° around any of the three axes. These thresholds were empirically determined in our previous studies (Kinno et al., 2008). If a run included one or several images over this threshold, we replaced the outlying image with an interpolated image, which was the average of the chronologically former and latter ones, and conducted the realignment procedure again. The realigned data were resliced every 3 mm using seventh-degree B-spline interpolation.

Each individual's structural image was matched with the mean functional image generated during realignment. The resultant structural image was spatially normalized to the standard brain space as defined by the Montreal Neurological Institute (MNI) using the extended version of the unified segmentation algorithm with light regularization; this is a generative model that combines tissue segmentation, bias correction, and spatial normalization in a single model (Ashburner and Friston, 2005). The resultant deformation field was applied to each realigned functional image to be spatially normalized with non-linear transformation. All normalized functional images were then smoothed by using an isotropic Gaussian kernel of 9 mm full-width at half maximum (FWHM). Low-frequency noise was removed by high-pass filtering at 1/128 Hz.

In the first-level analysis (i.e., the fixed-effects analysis within a participant), each participant's hemodynamic responses were modeled for the following types of events: initial 2-back trials with encoding alone, other 2-back trials, 6-s memory retrieval phase of retrieval trials, and 4-s *post hoc* period of retrieval trials. These event types were separately set for each group. Each event was modeled with the boxcar function overlaid with a hemodynamic response function. To minimize the effects of head movement, the six realignment parameters obtained from

[1]http://www.fil.ion.ucl.ac.uk/spm

preprocessing were included as a nuisance factor in a general linear model.

These modeled responses were then generated in a general linear model for each participant and used for the inter-subject comparison in a second-level analysis (i.e., the random-effects analysis for a group). To examine the activation of the regions in an unbiased manner, we adopted whole-brain analyses. For statistical analyses, a two-way ANOVA (group × event type) with *t*-tests was performed with three nuisance factors (age, gender, and laterality quotient), where the statistical threshold was set to family-wise error (FWE) corrected $p < 0.05$ for the voxel level. For the anatomical identification of activated regions, essentially we used the Anatomical Automatic Labeling (AAL) method[2] (Tzourio-Mazoyer et al., 2002) and the labeled data as provided by Neuromorphometrics Inc.[3], under academic subscription. In addition to whole-brain analyses, we adopted analyses of each region of interest (ROI) by using the MarsBaR-toolbox[4], in which an ROI was taken from a cluster identified by the "retrieval—2-back" contrast for all participants, which were further extracted with an AAL mask of each region.

RESULTS

Behavioral Results

We first compared the amounts of time required to write down the scheduled appointments (i.e., the duration of schedule recording) among the Note, Tablet, and Phone groups, and we observed a significant difference by a one-way ANOVA ($F_{(2,45)} = 6.5$, $p = 0.003$; **Figure 2A**). The duration was significantly shorter for the Note group compared to the Tablet and Phone groups combined (*t*-test, $t_{(46)} = 3.2$, $p = 0.002$). We also confirmed a significant difference between the Note and Device groups ($t_{(29)} = 3.0$, $p = 0.003$).

Relative to the chance level of 25% accuracy, the accuracy for the retrieval task was reliable and well below the ceiling level (**Figure 2B**). The participants' self-evaluation on confidence was also correct, because the AUC for the Note, Tablet, and Phone groups were 0.77 ± 0.14, 0.77 ± 0.12, and 0.74 ± 0.11, respectively, where group differences were not significant ($F_{(2,45)} = 0.2$, $p = 0.8$). The accuracy or RTs in the retrieval was not significantly different among the three groups (accuracy: $F_{(2,45)} = 0.5$, $p = 0.6$; RTs: $F_{(2,45)} = 0.8$, $p = 0.5$; **Figure 2C**); the accuracy and RTs in the interference and 2-back tasks were also comparable among the three groups ($p > 0.4$). However, we observed significant group differences when we focused on the easier questions of scheduled appointments (see "Materials and Methods" section; **Figure 2B**). According to non-parametric tests for the data showing ceiling effects, the accuracy of the easier questions was significantly higher for the Note group than the Tablet group (Wilcoxon rank-sum test, $W = 179$, $p = 0.04$), and the difference between the

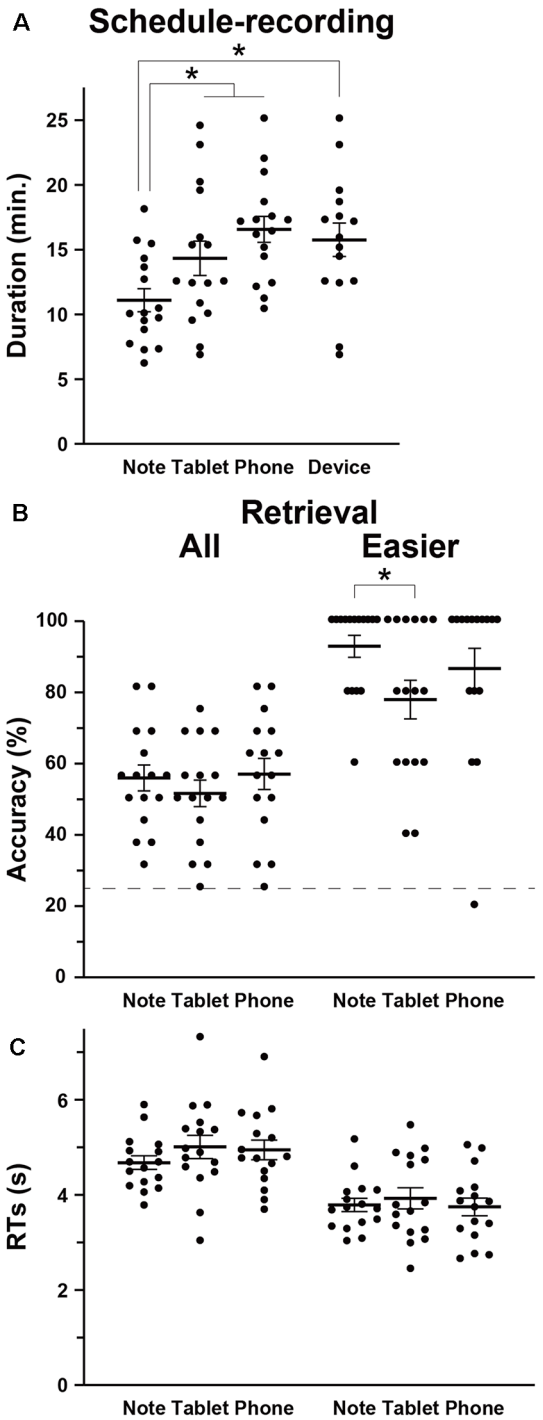

FIGURE 2 | Behavioral data. **(A)** The intergroup differences in the mean duration of schedule recording (see **Figure 1**), together with individual data points overlapped. In addition to the three groups (Note, Tablet, and Phone), we also introduced a *Device* group, which consisted of participants who used mainly notebooks daily and were assigned to either the Tablet or Phone group. **(B)** Accuracy in the retrieval task. The broken line denotes the chance level of 25% accuracy. For the easier (i.e., more straightforward) questions, the Note group showed significantly higher accuracy than the Tablet group. **(C)** Response times (RTs) in the retrieval task. Error bars indicate standard errors of the mean. *$p < 0.05$.

[2]http://www.gin.cnrs.fr/AAL2/
[3]http://Neuromorphometrics.com/
[4]http://marsbar.sourceforge.net/

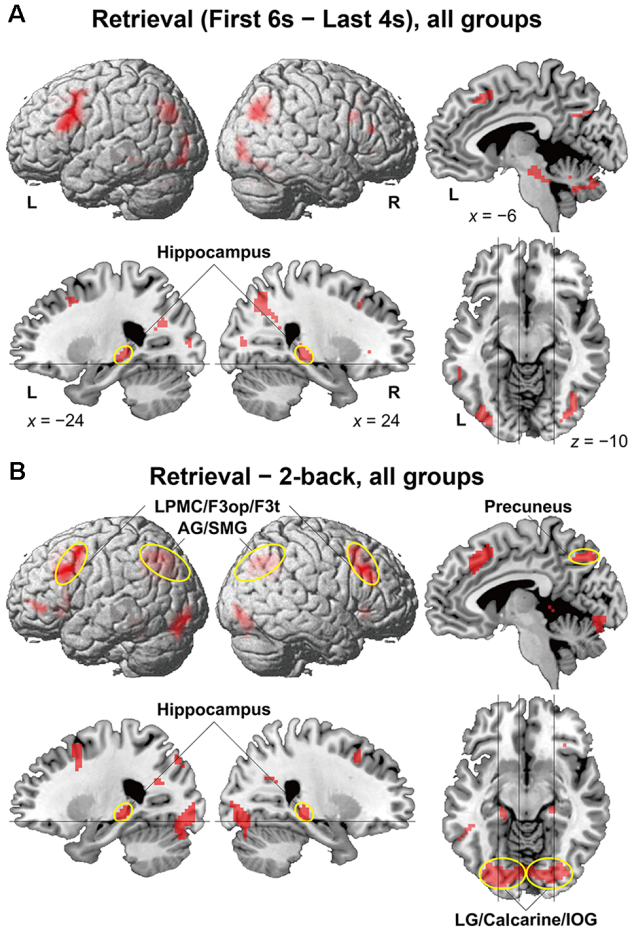

FIGURE 3 | Activated regions for the retrieval task. **(A)** Results of the "First 6 s—Last 4 s" contrast within the retrieval task period are shown for all participants. **(B)** Results of the "retrieval—2-back" contrast are shown for all participants. The lines indicate the locations of the sections. Localized activations were observed bilaterally in the lateral premotor cortex/opercular/triangular parts of the inferior frontal gyrus (LPMC/F3op/F3t), angular/supramarginal gyri (AG/SMG), hippocampus, precuneus, and lingual gyrus/calcarine/inferior occipital gyrus (LG/calcarine/IOG; see **Table 2** for the list of local maxima).

Note and Device groups was marginally significant (W = 164, p = 0.06).

The Tablet and Phone groups (or Device group) took more time for writing down (**Figure 2A**), and this might be due to slower input of characters with such mobile devices (no typing on the computer keyboard). However, at least between the Note and Tablet groups, the use of a stylus pen was just similar to writing with a four-color pen, and the physical layout of a notebook or tablet was equated as much as possible (see "Materials and Methods" section). Moreover, there was ample time for every group to write down all appointments into a monthly calendar. Therefore, *shorter* amounts of time for writing down and *higher* accuracy in easier questions for the Note group suggest that those cognitive processes for the Note group were actually deeper and more solid.

When all participants in the three groups were combined, the accuracy in the retrieval and 2-back tasks were significantly correlated (Pearson's correlation, r = 0.31, $t_{(46)}$ = 2.2, p = 0.03). RTs showed a significant correlation as well (r = 0.33, $t_{(46)}$ = 2.4, p = 0.02). These results confirm consistent immediate- and short-term memory capacities for every participant.

Enhanced Activations in Bilateral Regions for the Note Group

To identify brain regions specifically involved in the memory retrieval process, we directly compared activations between the 6-s memory retrieval phase and the 4-s *post hoc* period from each 10-s period of the retrieval task, denoted as "First 6 s—Last 4 s" contrast. This was because the mean RTs were less than 6 s for all but two participants (see **Figure 2C**). With this stringent contrast during the same stimulus presentation and task, dynamic signal changes induced by such active retrieval processes should be revealed. As shown in **Figure 3A**, localized activations were found bilaterally in the middle frontal gyrus, F3op/F3t, fusiform gyrus, AG/SMG, middle/inferior occipital gyrus (MOG/IOG), pallidum, and hippocampus; we also observed left-lateralized activation in the LPMC and precuneus.

It is still possible that these activations reflect immediate memory processes that were necessary to solve the retrieval task; note the above-mentioned correlation between performances of the two tasks. Thus, we further compared activations in the retrieval task (10-s period) against those in the 2-back task with more demanding immediate memory, which successfully removed common factors in both tasks (**Figure 3B**). The result of activations replicated the above-mentioned regions (**Table 2**), providing appropriate ROIs for further analyses. Additional activations were found bilaterally in the lingual gyrus (LG) and calcarine sulcus; we also observed left-lateralized activation in the orbital part of the inferior frontal gyrus (F3O).

We assessed percent signal changes for these ROIs, and found significant intergroup differences in the posterior hippocampus, precuneus, LG/calcarine/IOG, LPMC/F3op/F3t, and AG/SMG (**Figures 4A–E**). Activations in the first four regions were significantly different between the Note group and the combined Tablet and Phone groups (hippocampus: $t_{(94)}$ = 2.4, p = 0.02; precuneus: $t_{(94)}$ = 2.3, p = 0.03; LG/calcarine/IOG: $t_{(94)}$ = 2.7, p = 0.008; LPMC/F3op/F3t: $t_{(94)}$ = 2.0, p = 0.05), whereas those in the last region were significantly different between the Note and Phone groups ($t_{(62)}$ = 2.2, p = 0.03). Activations in the LG/calcarine/IOG and LPMC/F3op/F3t were also significantly different between the Note and Device groups (LG/calcarine/IOG: $t_{(60)}$ = 2.2, p = 0.03; LPMC/F3op/F3t: $t_{(60)}$ = 2.4, p = 0.02), even when the experience/accustomedness of using paper notebooks was equated. Moreover, we observed a significant positive correlation between the RTs in the retrieval task and the averaged signal changes in the ROIs of LPMC/F3op/F3t and AG/SMG for all participants (r = 0.31, $t_{(46)}$ = 2.2, p = 0.03; **Figure 4F**). This link between behavioral results and brain activations indicates that inner language processes were indeed involved

TABLE 2 | ROIs determined by the contrast of "retrieval—2-back" for all participants.

Brain regions	BA	Side	*x*	*y*	*z*	*Z*	Voxels
LPMC	6/8/9	L	−36	8	47	Inf	1,030
		R	39	17	53	7.1	*
F3op/F3t	44/45	L	−48	20	8	5.9	*
		R	48	29	35	7.5	*
ACC/pre-SMA	32/8	M	−6	29	44	Inf	*
F3t/F3O	45/47	L	−45	41	−4	7.8	71
Insula	13	L	−30	26	−4	7.0	46
		R	33	29	−4	6.8	49
ITG	20	L	−54	−43	−13	6.6	15
FG	37	L	−36	−46	−22	5.4	12
AG/SMG	39/40	L	−33	−70	35	Inf	246
		R	39	−67	38	Inf	180
Precuneus	7	L	−9	−64	41	7.8	97
			−21	−61	26	6.9	*
		R	24	−58	26	6.0	8
LG/Calcarine/IOG	18/19/17	L	−12	−88	−10	Inf	622
		R	9	−85	−10	Inf	*
Cerebellum Crus I/Crus II/VI		R	12	−79	−28	6.2	*
ibid. IV/V		M	−6	−40	−1	5.4	71
Hippocampus		L	−24	−31	−4	7.0	*
		R	24	−28	−4	6.4	31

Stereotactic coordinates (x, y, z) in the MNI space are shown for activation peaks of Z values which were more than 12 mm apart in either direction of the x, y, or z-axis. FWE corrected $p < 0.05$ for the voxel level. The region with an asterisk is included within the same cluster shown in the nearest row above. BA, Brodmann's area; L, left; M, medial; R, right; ACC, anterior cingulate cortex; AG, angular gyrus; F3O, orbital part of the inferior frontal gyrus (F3); F3op, opercular part of the F3; F3t, triangular part of the F3; FG, fusiform gyrus; IOG, inferior occipital gyrus; ITG, inferior temporal gyrus; LG, lingual gyrus; LPMC, lateral premotor cortex; MOG, middle occipital gyrus; pre-SMA, pre-supplementary motor area; SMG, supramarginal gyrus.

in during memory retrieval *via* the function of the language-related regions.

DISCUSSION

Using three groups of participants who performed a schedule-recording task using a paper notebook, electronic tablet, or smartphone, followed by a retrieval task (**Figure 1**), we obtained three major results. First, the duration of schedule recording was significantly shorter for the Note group than the Tablet and Phone groups, and accuracy was much higher for the Note group in easier (i.e., more straightforward) questions (**Figure 2**). Because the input methods were equated as much as possible between the Note and Tablet groups, these results indicate that the cognitive processes for the Note group were actually deeper and more solid. Second, brain activations for all groups during the retrieval phase were localized in the bilateral hippocampus, precuneus, LG/calcarine/IOG, and LPMC/F3op/F3t (**Figure 3**), confirming the involvement of verbalized memory retrieval processes for appointments. Third, activations in these regions were significantly higher for the Note group than those for the Tablet and Phone groups (**Figure 4**). These enhanced activations for the Note group could not be explained by general cognitive loads or task difficulty, because overall task performances were similar among the groups. Brain activations for the Tablet and Phone groups were similar, where the difference in input methods did not affect the results. On the other hand, the Note and Tablet groups showed a clear difference in brain activations even if the physical layout and input methods were controlled. Brain activations were significantly different also between the Note and Device groups, even when accustomedness to paper notebooks or mobile devices was equated for daily usage. The significant superiority in both accuracy and activations for the Note group suggested that the use of a paper notebook promoted the acquisition of rich encoding information and/or spatial information of real papers (see the "Introduction" section) and that this information could be utilized as effective retrieval clues, leading to higher activations in these specific regions.

The hippocampus is crucially involved not only in memory encoding and retrieval processes but also in spatial memory itself. The hippocampal-entorhinal cortex provides spatial representations, as demonstrated by grid cells (Hartley et al., 2014; Moser et al., 2015). It has also been suggested that activations in the human hippocampus encode distances between locations in the real world (Morgan et al., 2011; Howard et al., 2014). In a recent fMRI study using a graph structure of pictures, the adaptation signals in the hippocampal-entorhinal cortex were suppressed for shorter distances on the graph, indicating that non-spatial relationships were also encoded in these regions (Garvert et al., 2017). Other neuroimaging studies have shown that activations in the left posterior hippocampus were enhanced during retrieval compared with the encoding of word pairs (Prince et al., 2005) and that better recollection of proverbs was associated with a larger volume of the bilateral posterior hippocampus (Poppenk and Moscovitch, 2011). The results of the present study are consistent with these previous findings, in that the scheduled appointments included various cues of spatial and structural information in the calendar, which were especially abundant when participants used paper notebooks. Moreover, the retrieval

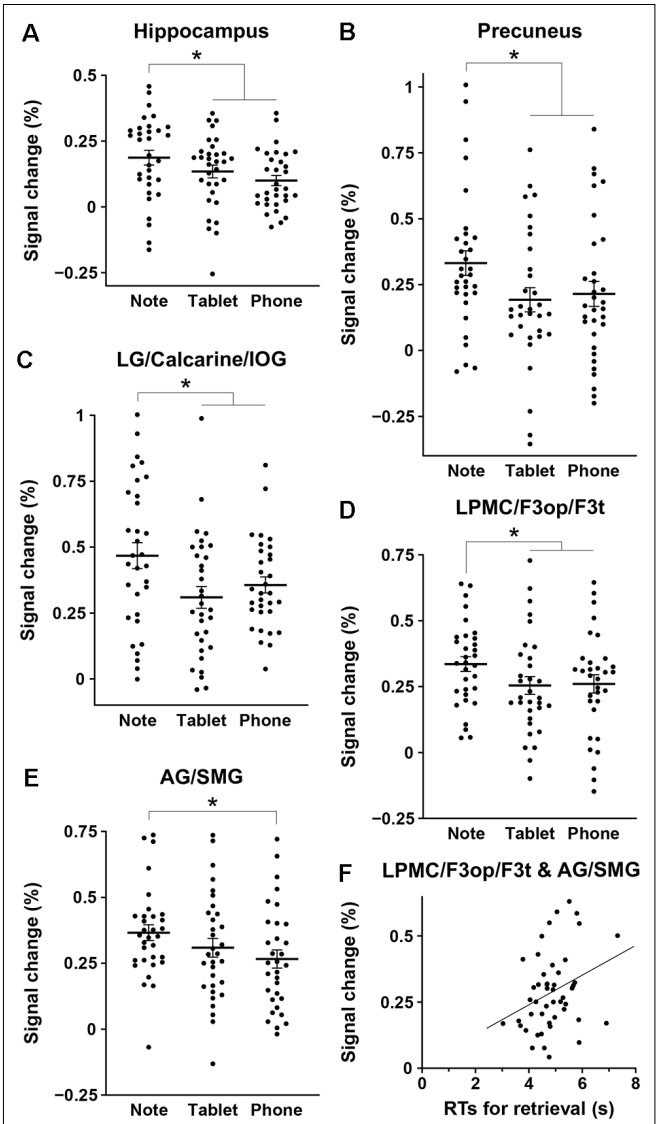

FIGURE 4 | Intergroup differences in brain activations for the retrieval task. **(A–E)** Mean percent signal changes, together with individual data points overlapped, for the three groups in the regions of interest (ROIs) of the hippocampus **(A)**, precuneus **(B)**, LG/calcarine/IOG **(C)**, LPMC/F3op/F3t **(D)**, and AG/SMG **(E)**. The signal changes of an ROI in each hemisphere were treated as independent samples, in reference to those in the 2-back task. Error bars indicate standard errors of the mean. $^*p < 0.05$. **(F)** A significant correlation between the RTs in the retrieval task and the averaged signal changes in the ROIs of the LPMC/F3op/F3t and AG/SMG (the language-related regions) for all participants.

of such encoded information was explicitly required by our retrieval task and was shown to elicit activations in the bilateral posterior hippocampus.

Concerning activation in the visual cortex, a previous study reported that the visual cortex was activated during the retrieval of pictorial visual information without actual visual stimulation (Wheeler et al., 2000). The visual areas play a key role in visual imagery as well, and activations in those regions could be affected by focal attention during imagery (Sakai and Miyashita, 1994). Indeed, a study with fMRI decoding revealed activation in the V1–V3 when participants reported visual imagery of an object during dreaming, about which was inquired afterward (Horikawa et al., 2013). Another study reported that retrieval of visual information was related to activation patterns in the V1–V3, and further showed that the activation patterns in the hippocampus predicted the mnemonic strength (Bosch et al., 2014). As regards the precuneus, a positron emission tomography (PET) study with a paired-word retrieval task showed memory-related activation for both visual and auditory stimuli, indicating a modality-general role of the precuneus (Krause et al., 1999). The internal representation for visual imagery of the encoded calendar provides a plausible account for our results, in which the paper notebook provides richer information than mobile devices.

According to our previous study, the left F3op/F3t, right LPMC, and right F3op/F3t are included in the network for syntax and its supportive system (Network I; Kinno et al., 2014), whereas the left LPMC is critical to the network for syntax and input/output interface (Network II). In the present study, we observed activation in the left F3t/F3O, which is an essential part of the network for syntax and semantics (Network III). Thus all three networks that are crucial for syntactic processing were involved in the retrieval of scheduled appointments. The enhanced activations for the Note group suggest that the use of paper notebooks even influenced natural language processes, possibly reflecting the encoding of specific episodes.

Our present experiments demonstrated that brain activations related to memory, visual imagery, and language during the retrieval of specific information, as well as the deeper encoding of that information, were stronger in participants using a paper notebook than in those using electronic devices. Our results suggest that the use of a paper notebook affects these higher-order brain functions, and this could have important implications for education, particularly in terms of the pros and cons of e-learning. The expanded use of mobile devices or computers could undercut the use of traditional textbooks and paper notebooks, which may in fact provide richer information from the perspective of memory encoding. Further research is needed to elucidate the actual changes in brain activation due to the long-term exposure to mobile devices.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Review Board of the University of Tokyo, Komaba Campus. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

KU and KLS designed the study, analyzed the data, and wrote the manuscript. TI and TY contributed to the initial discussion. KU conducted the experiment. All authors contributed to the article and approved the submitted version.

FUNDING

The authors declare that this study received funding from the Consortium for Applied Neuroscience, NTT Data Institute of Management Consulting, Inc. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication.

ACKNOWLEDGMENTS

We would like to thank the members of the Consortium for Applied Neuroscience for their input on the use of media, Yuma Tsuta for contributions to the design of the experiments, the MR scanning, and the analysis of behavioral results, Naoko Komoro for technical assistance, and Hiromi Matsuda for administrative assistance.

REFERENCES

Aragón-Mendizábal, E., Delgado-Casas, C., Navarro-Guzmán, J., Menacho-Jiménez, I., and Romero-Oliva, M. (2016). A comparative study of handwriting and computer typing in note-taking by university students. *Comunicar* 48, 101–107. doi: 10.1038/s42003-020-1052-8

Ashburner, J., and Friston, K. J. (2005). Unified segmentation. *NeuroImage* 26, 839–851. doi: 10.1016/j.neuroimage.2005.02.018

Bosch, S. E., Jehee, J. F. M., Fernández, G., and Doeller, C. F. (2014). Reinstatement of associative memories in early visual cortex is signaled by the hippocampus. *J. Neurosci.* 34, 7493–7500. doi: 10.1523/JNEUROSCI.0805-14.2014

Broadbent, N. J., Squire, L. R., and Clark, R. E. (2004). Spatial memory, recognition memory and the hippocampus. *Proc. Natl. Acad. Sci. U S A* 101, 14515–14520. doi: 10.1073/pnas.0406344101

Chadwick, M. J., Hassabis, D., Weiskopf, N., and Maguire, E. A. (2010). Decoding individual episodic memory traces in the human hippocampus. *Curr. Biol.* 20, 544–547. doi: 10.1016/j.cub.2010.01.053

Drozdick, L. W., Wahlstrom, D., Zhu, J., and Weiss, L. G. (2012). "The Wechsler adult intelligence scale—fourth edition and the wechsler memory scale—fourth edition," in *Contemporary Intellectual Assessment: Theories, Tests and Issues* eds D. P. Flanagan and P. L. Harrison (New York, NY: Guilford Press), 197–223.

Eichenbaum, H. (2004). Hippocampus: cognitive processes and neural representations that underlie declarative memory. *Neuron* 44, 109–120. doi: 10.1016/j.neuron.2004.08.028

Eichenbaum, H. (2017). Prefrontal-hippocampal interactions in episodic memory. *Nat. Rev. Neurosci.* 18, 547–558. doi: 10.1038/nrn.2017.74

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognit. Lett.* 27, 861–874. doi: 10.1016/j.patrec.2005.10.010

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., and Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: a general linear approach. *Hum. Brain Mapp.* 2, 189–210. doi: 10.1002/hbm.460020402

Garvert, M. M., Dolan, R. J., and Behrens, T. E. J. (2017). A map of abstract relational knowledge in the human hippocampal-entorhinal cortex. *eLife* 6, 1–20. doi: 10.7554/eLife.17086

Goh, W. D., and Lu, S. H. X. (2012). Testing the myth of the encoding-retrieval match. *Mem. Cogn.* 40, 28–39. doi: 10.3758/s13421-011-0133-9

Hartley, T., Lever, C., Burgess, N., and O'keefe, J. (2014). Space in the brain: how the hippocampal formation supports spatial cognition. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369, 1–18. doi: 10.1098/rstb.2012.0510

Horikawa, T., Tamaki, M., Miyawaki, Y., and Kamitani, Y. (2013). Neural decoding of visual imagery during sleep. *Science* 340, 639–642. doi: 10.1126/science.1234330

Howard, L. R., Javadi, A. H., Yu, Y., Mill, R. D., Morrison, L. C., Knight, R., et al. (2014). The hippocampus and entorhinal cortex encode the path and euclidean distances to goals during navigation. *Curr. Biol.* 24, 1331–1340. doi: 10.1016/j.cub.2014.05.001

Kinno, R., Kawamura, M., Shioda, S., and Sakai, K. L. (2008). Neural correlates of noncanonical syntactic processing revealed by a picture-sentence matching task. *Hum. Brain Mapp.* 29, 1015–1027. doi: 10.1002/hbm.20441

Kinno, R., Ohta, S., Muragaki, Y., Maruyama, T., and Sakai, K. L. (2014). Differential reorganization of three syntax-related networks induced by a left frontal glioma. *Brain* 137, 1193–1212. doi: 10.1093/brain/awu013

Krause, B. J., Schmidt, D., Mottaghy, F. M., Taylor, J., Halsband, U., Herzog, H., et al. (1999). Episodic retrieval activates the precuneus irrespective of the imagery content of word pair associates—a PET study. *Brain* 122, 255–263. doi: 10.1093/brain/122.2.255

Lundstrom, B. N., Petersson, K. M., Andersson, J., Johansson, M., Fransson, P., and Ingvar, M. (2003). Isolating the retrieval of imagined pictures during episodic memory: activation of the left precuneus and left prefrontal cortex. *NeuroImage* 20, 1934–1943. doi: 10.1016/j.neuroimage.2003.07.017

Mangen, A., Walgermo, B. R., and Brønnick, K. (2013). Reading linear texts on paper versus computer screen: effects on reading comprehension. *Int. J. Educ. Res.* 58, 61–68. doi: 10.1016/j.ijer.2012.12.002

Miyashita, Y. (2019). Perirhinal circuits for memory processing. *Nat. Rev. Neurosci.* 20, 577–592. doi: 10.1038/s41583-019-0213-6

Morgan, L. K., Macevoy, S. P., Aguirre, G. K., and Epstein, R. A. (2011). Distances between real-world locations are represented in the human hippocampus. *J. Neurosci.* 31, 1238–1245. doi: 10.1523/JNEUROSCI.4667-10.2011

Moscovitch, M., Cabeza, R., Winocur, G., and Nadel, L. (2016). Episodic memory and beyond: the hippocampus and neocortex in transformation. *Annu. Rev. Psychol.* 67, 105–134. doi: 10.1146/annurev-psych-113011-143733

Moser, M.-B., Rowland, D. C., and Moser, E. I. (2015). Place cells, grid cells and memory. *Cold Spring Harb. Perspect. Biol.* 7, 1–15. doi: 10.1101/cshperspect.a021808

Mueller, P. A., and Oppenheimer, D. M. (2014). The pen is mightier than the keyboard: advantages of longhand over laptop note taking. *Psychol. Sci.* 25, 1159–1168. doi: 10.1177/0956797614524581

Nairne, J. S. (2002). The myth of the encoding-retrieval match. *Memory* 10, 389–395. doi: 10.1080/09658210244000216

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

Poppenk, J., and Moscovitch, M. (2011). A hippocampal marker of recollection memory ability among healthy young adults: contributions of posterior and anterior segments. *Neuron* 72, 931–937. doi: 10.1016/j.neuron.2011.10.014

Prince, S. E., Daselaar, S. M., and Cabeza, R. (2005). Neural correlates of relational memory: successful encoding and retrieval of semantic and perceptual associations. *J. Neurosci.* 25, 1203–1210. doi: 10.1523/JNEUROSCI.2540-04.2005

Rugg, M. D., Johnson, J. D., Park, H., and Uncapher, M. R. (2008). "Encoding-retrieval overlap in human episodic memory: a functional neuroimaging perspective," in *Essence of Memory*, Progress in Brain Research 169, eds W. S. Sossin, J.-C. Lacaille, V. F. Castellucci and S. Belleville (Amsterdam: Elsevier), 339–352.

Sakai, K. L. (2005). Language acquisition and brain development. *Science* 310, 815–819. doi: 10.1126/science.1113530

Sakai, K., and Miyashita, Y. (1994). Visual imagery: an interaction between memory retrieval and focal attention. *Trends Neurosci.* 17, 287–289. doi: 10.1016/0166-2236(94)90058-2

Schacter, D. L., Addis, D. R., and Buckner, R. L. (2007). Remembering the past to imagine the future: the prospective brain. *Nat. Rev. Neurosci.* 8, 657–661. doi: 10.1038/nrn2213

Tulving, E. (2002). Episodic memory: from mind to brain. *Annu. Rev. Psychol.* 53, 1–25. doi: 10.1146/annurev.psych.53.100901.135114

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage* 15, 273–289. doi: 10.1006/nimg.2001.0978

Wästlund, E., Reinikka, H., Norlander, T., and Archer, T. (2005). Effects of VDT and paper presentation on consumption and production of information: psychological and physiological factors. *Comput. Human Behav.* 21, 377–394. doi: 10.1016/j.chb.2004.02.007

Wheeler, M. E., Petersen, S. E., and Buckner, R. L. (2000). Memory's echo: vivid remembering reactivates sensory-specific cortex. *Proc. Natl. Acad. Sci. U S A* 97, 11125–11129. doi: 10.1073/pnas.97.20.11125

Conflict of Interest: TI and TY were employed by the company NTT Data Institute of Management Consulting, Inc.

The remaining 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 in Behavioral Neuroscience

ORIGINAL RESEARCH
published: 26 March 2021
doi: 10.3389/fnbeh.2021.631957

Modality-Dependent Brain Activation Changes Induced by Acquiring a Second Language Abroad

*Kuniyoshi L. Sakai *, Tatsuro Kuwamoto, Satoma Yagi and Kyohei Matsuya*

Department of Basic Science, Graduate School of Arts and Sciences, The University of Tokyo, Tokyo, Japan

The dynamic nature of cortical activation changes during language acquisition, including second-language learning, has not been fully elucidated. In this study, we administered two sets of reading and listening tests (Pre and Post) to participants who had begun to learn Japanese abroad. The two sets were separated by an interval of about 2 months of Japanese language training. We compared the results of longitudinal functional MRI experiments between the two time-points and obtained the following major findings. First, the left-dominant language areas, as well as bilateral visual and auditory areas, were activated, demonstrating the synergistic effects of multiple modalities. There was also significant activation in the bilateral hippocampi, indicating the expected involvement of memory-related processes. Second, consistent with the behavioral improvements from Pre to Post, the brain activations *decreased* in the left inferior and middle frontal gyri during the listening tests, as well as in the visual areas (the bilateral inferior and superior parietal lobules, and left inferior and middle occipital gyri) during the reading tests, while activations in the right superior and middle temporal gyri *increased* during the listening tests. These modality-dependent activation changes could not be explained by domain-general cognitive factors, such as habituation or familiarization, because we used completely different test sets for Pre and Post. Third, the posterior hippocampus showed a main effect of the hemisphere, whereas the anterior hippocampus showed a significant main effect of the event (i.e., specific to first listening events), reflecting initial encoding of auditory information alone. In summary, activation changes from Pre to Post indicate functional changes in modality-dependent networks over a short period of staying abroad, which would enable effective acquisition of a second language.

Keywords: language acquisition, syntax, fMRI, hippocampus, learning and memory (neurosciences)

OPEN ACCESS

Edited by:
Alfredo Brancucci,
Foro Italico University of Rome, Italy

Reviewed by:
Anita D'Anselmo,
University of Bologna, Italy
Sharlene D. Newman,
University of Alabama, United States

****Correspondence:***
Kuniyoshi L. Sakai
sakai@sakai-lab.jp

Specialty section:
This article was submitted to Learning and Memory, a section of the journal Frontiers in Behavioral Neuroscience

Received: *21 November 2020*
Accepted: *22 February 2021*
Published: *26 March 2021*

Citation:
Sakai KL, Kuwamoto T, Yagi S and Matsuya K (2021) Modality-Dependent Brain Activation Changes Induced by Acquiring a Second Language Abroad. Front. Behav. Neurosci. 15:631957. doi: 10.3389/fnbeh.2021.631957

INTRODUCTION

Recent advances in human neuroimaging studies have revealed both anatomical and functional changes during second language (L2) acquisition (Chee et al., 2001; Reiterer et al., 2009; Schlegel et al., 2012; Li et al., 2014), which occur in the language-related regions required for first or native languages (L1), and possibly in other regions as well. By using diffusion magnetic resonance imaging (diffusion MRI), we have recently shown that the structural measure (fractional anisotropy) of the left arcuate fasciculus (i.e., dorsal pathway) connecting the left inferior frontal gyrus (IFG) and other language-related regions was significantly correlated with

performance on a syntactic task in high-school students who had studied English as an L2 at school (Yamamoto and Sakai, 2016), and this correlation was clearly dissociated from developmental changes (Yamamoto and Sakai, 2017). Another MRI study performed during 16 weeks of vocabulary training in L2 for university students showed a volume increase in the *right* IFG, as well as a connectivity change in the right hemisphere (Hosoda et al., 2013). Although the issue of hemispheric dominance remains, such anatomical changes during L2 acquisition likely affect the functions of those regions as well.

Each of the language-related regions in the left hemisphere has a specific function irrespective of modalities (i.e., audition or vision) or input/output. For example, the dorsal and ventral regions of the left IFG are specialized in syntax and sentence comprehension, respectively, whereas the left superior temporal regions and angular/supramarginal gyri subserve phonological and lexico-semantic processing, respectively (Sakai, 2005). By using functional MRI (fMRI), we previously compared two groups of high-school students, whose ages of acquisition (AOA) in L2 were 6 years apart, and demonstrated that activations of the left IFG correlated positively with the accuracy of a syntactic task for the *late* learners (mean AOA: 12.6), whereas activations of the left ventral IFG correlated negatively with the accuracy for the *early* learners (mean AOA: 5.6; Sakai et al., 2009). Although some fMRI studies support the idea that AOA affects cortical activations, such that the left IFG activation for grammatical processing in L2 is greater than that in L1 (Wartenburger et al., 2003), other fMRI studies have concluded that the degree of exposure to language affects the left IFG activation, even if the AOA is matched (Perani et al., 2003). To resolve these conflicting claims, we have proposed that the cortical activations may initially increase upon onset of acquisition, then remain at the increased level for some time, and finally fall during the consolidation of linguistic competence (see Figure 3E in Sakai, 2005). If this general law applies to L2 in general, then the language-related regions may show higher, lower, or comparable activation, depending on which developmental phases or aspects are compared. It is also probable that the time course of L2 acquisition depends on the specific linguistic function we focus on.

It is also likely that modality-specific regions, i.e., auditory and visual areas, would also show activation changes during L2 training for listening and reading abilities. In our previous study on newly learned Hangul letters with speech sounds, we showed a positive correlation between accuracy improvements and activation *increases* in the left posterior inferior temporal gyrus (Hashimoto and Sakai, 2004), indicating that activation changes occur even within two consecutive *days* during the initial stage of learning. It would be interesting to examine whether cortical activations continue to increase or rather begin to decrease after several *months* of exposure to newly acquired speech sounds and letters in L2, still at the early stage of training. In our present longitudinal experiments, we conducted reading and listening tests (see **Figure 1**) before and after about 2 months of L2 training for each participant, and examined the differences in brain activation between the two time-points. The reading and listening tests in L2 were similar to those used in a classroom, where auditory stimuli are usually presented twice: first listening for general comprehension, and second listening for more selective focusing on specific questions. The reading and listening tests differed not only in their modalities but in their question contents and presentations (e.g., visual material can be read repeatedly). However, both syntactic and semantic processes were crucially involved in both types of tests, and we should bear in mind that modality-dependent effects examined here may contain such subsidiary processes as well.

During the initial L2 training, language-related regions, as well as auditory and visual areas, are expected to be crucially involved. Because they are the most critical regions of syntactic processing, we predict that the left inferior and middle frontal gyri (L. IFG/MFG) will exhibit activation changes. The L. IFG/MFG have been proposed as grammar centers (Sakai, 2005), and we have recently clarified that their right-side homologs, the right IFG/MFG (R. IFG/MFG), play a supportive role as a part of the syntax-related network I (Kinno et al., 2014; Tanaka et al., 2020). The auditory areas include the bilateral superior and middle temporal gyri (STG/MTG), and the visual areas may include at least the bilateral inferior and superior parietal lobules (IPL/SPL), fusiform gyrus (FG), and inferior and middle occipital gyri (IOG/MOG). Another candidate region involved in L2 acquisition would be the hippocampus due to its role in processing episodic memory (Zeidman and Maguire, 2016), although its role in the process of language acquisition has not been fully elucidated. While an fMRI study suggested that an increased proficiency level of an artificial language was associated with decreased activity in the left posterior hippocampus (Opitz and Friederici, 2003), the left posterior hippocampus is selectively activated for memory retrieval of word associations (Prince et al., 2005). The objectives of the present study were to clarify how each of these cortical regions and hippocampi showed activations depending on different modalities and time-points during L2 acquisition.

MATERIALS AND METHODS

Participants

We recruited 18 participants, who had learned Japanese through EF (EF Education First, Switzerland) courses for the first time at the EF Tokyo campus in Shibuya. During their 6-12 months stay abroad in Japan, the participants were exposed to Japanese in various rich social environments for at least 3 h each day. To focus on activation changes at the early stage of training in a second language, one participant who had studied Japanese for 6 years before the stay was dropped. The other participants had no history of learning Japanese before their stay in Japan. We assessed handedness according to the Edinburgh inventory (Oldfield, 1971), and dropped two participants whose laterality quotients showed left-handedness. The remaining 15 participants (10 males and 5 females) were 21 $\pm$ 3.5 years old [mean $\pm$ standard deviation (SD)], and showed right-handedness (laterality quotients: 79 $\pm$ 26). Their L1s were German (7), Norwegian (4), Spanish (2), French (1), and Dutch (1); they had also learned English as L2 mostly at school.

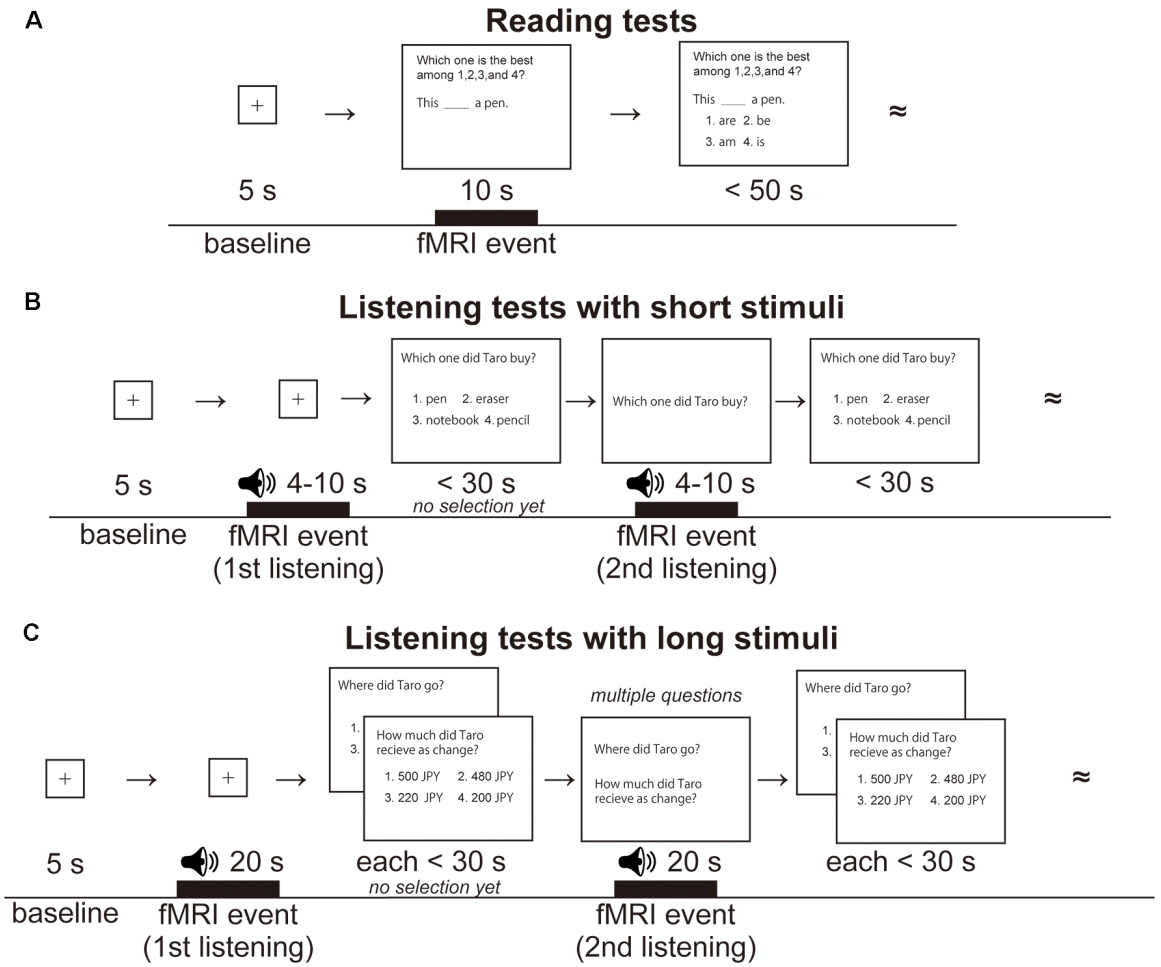

FIGURE 1 | Examples of reading and listening tests for learning Japanese as a second language. **(A)** A single trial of the reading tests. Following the presentation of a fixation point for 5 s (baseline for brain activation), a test sentence was visually presented for 10 s (an fMRI event of *reading*), and then four choices were additionally presented as possible replacements for the underlined portion of the sentence. The participants were required to select one of the choices within 50 s. The double tilde symbols denote continuation of similar trials in that run. Visual stimuli were presented in Japanese, but English equivalents are shown here. **(B)** A single trial of the listening tests with short stimuli. The same speech stimulus was presented twice, for 4–10 s each time (fMRI events labeled *first listening* and *second listening*). After listening to the first stimulus, the participants were visually presented with a question sentence and four possible answers and were given 30 s to proceed without choosing an answer. After listening to the second stimulus, the participants finally chose one of the four answers by pressing a button. Both auditory and visual stimuli were presented in Japanese, but English equivalents are shown here. **(C)** A single trial of the listening tests with long stimuli. The same speech stimulus was presented twice, for 20 s each time (fMRI events labeled *first listening* and *second listening*), together with multiple questions. English equivalents are shown here.

All experiments were performed as per relevant guidelines and regulations, including the Declaration of Helsinki, and the Singapore Statement on Research Integrity. All participants provided their written informed consent to participate in this study after the nature and possible consequences of the study were explained. Approval for these experiments was obtained from the institutional review board of the University of Tokyo, Komaba Campus.

Stimuli and Tasks

We designed two sets of tests (A and B) from EF placement tests in Japanese, each of which consisted of 25 reading and five listening tests, including questions about syntactic structures and sentence comprehension. Auditory stimuli were speech sounds taken from EF placement tests in Japanese. During the EF course of Japanese for beginners, different sets of tests (*Pre* and *Post*) were administered 40–100 days apart (61 ± 18 days), where the Pre sets followed 2 or 3 months of initial courses at the beginning of their stay in the L2 environment. The order of sets A and B was counterbalanced across participants.

For the reading tests, participants read a short sentence with a missing phrase or particle, and chose one from four suggested replacements. All visual stimuli were presented in white against a dark background. For fixation, a small red cross was shown at the center of the screen to initiate eye movements from the same fixed position, and the participants were instructed to return their eyes to this position. During the scans, the participants wore an eyeglass-like MRI-compatible display (resolution = 800 × 600,

framerate = 60 fps; VisuaStim Digital, Resonance Technology Inc., Northridge, CA, USA). For weak-sighted participants, correction lenses were inserted in front of the display.

For each reading test, we presented a fixation point for 5 s, which was used as a baseline for the fMRI analyses (**Figure 1A**). A test sentence with a missing phrase or particle was visually presented for 10 s (an fMRI event of *reading*), and then four choices were presented to fill in the blank. Participants read the stimuli and responded within 50 s while the sentence was presented, which was the maximum time set for beginners. A set of reading tests were presented in three different writing systems, corresponding to the three steps of learning to read in Japanese: five sentences written in the Roman alphabet of Japanese sounds (easiest notation for Europeans), 15 sentences in *kana* (a phonetic alphabet in Japanese, including both *hiragana* and *katakana*), and five sentences in *kana* and *kanji* (with Chinese characters, in a normal writing style). The leading questions (e.g., "Which one is the best among 1, 2, 3, and 4?") and choices were presented in *kana* alone. For each participant, we tested five scanning runs, where each run included five reading tests.

For the listening tests, not only *short* stimuli of a brief Japanese sentence for 4–10 s (**Figure 1B**), but *long* stimuli of multiple sentences for 20 s were administered to examine different aspects of listening abilities (**Figure 1C**). The former involved a basic knowledge about sentence construction and comprehension of a single sentence, while the latter required further syntactic and contextual understanding of connected sentences; we combined both aspects in the subsequent analyses. Only one straightforward question was assigned to each short stimulus, whereas two or three questions were associated with each long stimulus. Test set A included three short stimuli with one question each, and a long stimulus with two questions; test set B included two short stimuli with one question each, and a long stimulus with three questions. For each participant, we tested a scanning run with five listening tests (i.e., five questions).

For each listening test with a short stimulus, we presented a fixation point as a baseline, and then speech sounds were presented for 4–10 s (an fMRI event of *first listening*). The first listening event was the initial pure exposure to speech stimuli without performing a task or seeing any visual material. Next, a question was visually presented in *kana* with four choices to let the participants know what would be asked on the auditory stimuli; the participants were given up to 30 s to read this material, and indicated their readiness to proceed, without choosing an answer, by pressing *any* of the four buttons. The same speech sounds were presented again while the same question set (but without choices) was visually presented (an fMRI event of *second listening*), which simulated a listening test with questions on the display, like a TOEFL (Test of English as a Foreign Language) listening test; this event thus involved multimodal integration. After listening to the second stimulus, the participants finally chose one of the four answers by pressing an appropriate button. Concerning each listening test with a long stimulus (20 s), the differences from the listening test with short stimuli were the duration of fMRI events and the presence of multiple questions, as stated above.

During the scans, the participants wore an MRI-compatible headphone, VisuaStim Digital (Resonance Technology Inc., Northridge, CA, USA), a pair of earmuffs (3M Peltor, St. Paul, MN, USA), and a pair of earplugs (Earasers, Persona Medical, Casselberry, FL, USA) to reduce the high-frequency noises (>1 kHz) of the scanner. Before scanning, we appropriately adjusted the sound pressure level for each participant by presenting sample stimuli. The stimulus presentation and collection of behavioral data [accuracy and response times (RTs)] were controlled with the Presentation software package (Neurobehavioral Systems, Albany, CA, USA).

MRI Data Acquisition and Analyses

For the MRI data acquisition, we used a 3.0 T MRI system with a bird-cage head coil (Signa HDxt; GE Healthcare, Milwaukee, WI, USA), and all tests were conducted in an MRI scanner. We obtained fMRI data using a gradient-echo echo-planar imaging (EPI) sequence [30 axial slices, thickness = 3 mm, slice gap = 0.5 mm, repetition time (TR) = 2 s, echo time (TE) = 30 ms, flip angle (FA) = 78°, field of view (FOV) = 192 × 192 mm^2, in-plane resolution = 3 × 3 mm^2]. In each run, we discarded the initial four volumes that allowed for the rise of the MR signals. We also obtained high-resolution T1-weighted images of the whole brain with a three-dimensional fast spoiled gradient recalled acquisition in the steady-state (3D FSPGR) sequence (136 axial slices, TR = 8.5 ms, TE = 2.6 ms, FA = 25°, FOV = 256 × 256 mm^2, volume resolution = 1 × 1 × 1 mm^3) for normalizing fMRI data.

The fMRI data were analyzed in a standard manner using SPM12 statistical parametric mapping software (Wellcome Trust Center for Neuroimaging[1] (Friston et al., 1995), implemented on MATLAB (Math Works, Natick, MA, USA). The acquisition timing of each slice was corrected using the middle slice (the 15th slice chronologically) as a reference for the EPI data. We realigned the time-series data in multiple runs to the first volume in all runs, and further realigned the data to the mean volume of all runs. The realigned data were resliced using seventh-degree B-spline interpolation so that each voxel of each functional image matched that of the first volume.

After alignment to the AC-PC line, each participant's T1-weighted structural image was coregistered to the mean functional image generated during realignment. The coregistered structural image was spatially normalized to the standard brain space as defined by the Montreal Neurological Institute (MNI), by using the unified segmentation algorithm with medium regularization, which is a generative model that combines tissue segmentation, bias correction, and spatial normalization in the inversion of a single unified model. After spatial normalization, the resultant deformation field was applied to the realigned functional imaging data, which was resampled every 3 mm using seventh-degree B-spline interpolation. All normalized functional images were then smoothed by using an isotropic Gaussian kernel of 9 mm full-width at half maximum.

[1] http://www.fil.ion.ucl.ac.uk/spm/

Low-frequency noise was removed by high-pass filtering at 1/128 Hz.

In the first-level analysis (i.e., the fixed-effects analysis) for each test, we set three types of conditions: reading, first listening, and second listening. Each participant's hemodynamic responses induced by fMRI events were modeled with a boxcar function with a duration as shown in **Figure 1**. The boxcar function was then convolved with a hemodynamic response function. To minimize the effects of head movement, the six realignment parameters obtained from preprocessing were included as a nuisance factor in a general linear model. The images under the fMRI events and those under the baselines were then generated in the general linear model for each participant, and they were used for the "fMRI event – baseline" contrast in the second-level analysis (i.e., the random-effects analysis), which was thresholded at uncorrected $p < 0.0001$ for the voxel level, and at corrected $p < 0.05$ for the cluster level, with family-wise error (FWE) correction across the whole brain. To examine the activation of the regions in an unbiased manner, we adopted whole-brain analyses.

To calculate averaged percent signal changes for the Pre and Post sets, we focused on seven bilateral regions of interest (ROIs), which were selected from the significantly activated regions during the *second listening* events of the *Pre* sets: the inferior and middle frontal gyri (IFG/MFG), superior and middle temporal gyri (STG/MTG), inferior and superior parietal lobules (IPL/SPL), fusiform gyrus (FG), inferior and middle occipital gyri (IOG/MOG), and posterior hippocampus. The ROI of the anterior hippocampus was selected from activations during first listening events of the Post sets, and the following mask was further applied to separate y coordinates from the posterior part: $-50 < x < 0$, $-26 < y < 30$, $-40 < z < 30$ for the left hippocampus, and $0 < x < 50$, $-23 < y < 30$, $-40 < z < 30$ for the right hippocampus. To separate larger clusters spanning multiple regions into each ROI, we used masks combining cortical regions defined by the Anatomical Automatic Labeling method[2]; (Tzourio-Mazoyer et al., 2002), together with the labeled data as provided by Neuromorphometrics Inc.,[3] under academic subscription. The IFG/MFG consisted of three regions [middle frontal gyrus (Frontal_Mid_2), inferior frontal gyrus/opercular part (Frontal_Inf_Oper), and inferior frontal gyrus/triangular part (Frontal_Inf_Tri)], the STG/MTG of three regions [Heschl's gyrus (Heschl), superior temporal gyrus (Temporal_Sup), and middle temporal gyrus (Temporal_Mid)], the IPL/SPL of four regions [superior parietal gyrus (Parietal_Sup), inferior parietal gyrus (Parietal_Inf), supramarginal gyrus (SupraMarginal), and angular gyrus (Angular)], and the IOG/MOG of six regions [calcarine fissure and surrounding cortex (Calcarine), cuneus (Cuneus), lingual gyrus (Lingual), superior occipital gyrus (Occipital_Sup), middle occipital gyrus (Occipital_Mid), and inferior occipital gyrus (Occipital_Inf)]. For each ROI, we extracted the mean percent signal changes from each participant using the MarsBaR-toolbox[4].

RESULTS

Here, we present an outline of our analyses carried out on behavioral data and fMRI data. First, we assessed improvements in the reading and listening tests from Pre to Post sets in terms of the accuracy and response times (RTs). We measured RTs from the onset of choice presentation for the reading tests, and the *second* presentation of each question for the listening tests. Because listening tests were generally difficult for L2 learners, the listening stimuli were always presented twice, as in the case of original placement tests. Note that the first presentation of each question and the choices (options) was just preparatory without choosing an answer (see the "Stimuli and Tasks" section). This procedure was identical between the short and long stimuli (see **Figure 1**), and the RTs would become consistent if the participants could reach a correct answer by the second presentation; we thus averaged behavioral data for these stimuli together.

Next, we examined cortical activation during the presentation of Pre or Post Sets. By observing overall activation patterns for the events of reading, first listening, and second listening, we identified the most critical regions that showed significant activations, including sensory areas and language areas, as well as memory-related hippocampi. We finally conducted ROI analyses to statistically compare signal changes among various conditions: hemispheres [left, right], sets [Pre, Post], and events [reading, first listening, second listening]. We used t-tests for direct comparisons between individual conditions, whereas a repeated measures ANOVA (rANOVA) was used when the main effect of hemisphere or event was expected.

Behavioral Results

The accuracy and RTs are shown in **Figure 2**. With respect to the accuracy, there was a significant improvement in the reading tests from Pre to Post sets (paired t-test, $t_{(14)} = 3.1$, $p = 0.0073$; **Figure 2A**), while there was no difference in the listening tests ($t_{(14)} = 0$, $p = 1$). In regard to the RTs, we observed a significant improvement in listening tests, i.e., a reduction in RTs from Pre to Post ($t_{(14)} = -3.1$, $p = 0.0073$; **Figure 2B**), while there was no significant difference in the reading tests ($t_{(14)} = -1.1$, $p = 0.29$). These behavioral results indicate that there were improvements in both tests during the course of about 2 months.

Regarding the individually variable intervals between the Pre and Post sets, there was no significant correlation between the intervals and significant performance improvements from Pre to Post sets: the accuracy changes in the reading tests ($r = 0.16$, $p = 0.6$), or the RT changes in the listening tests ($r = 0.28$, $p = 0.3$).

Overall Cortical Activation During the Presentation of Pre or Post Sets

Significant activation was observed in both hemispheres with some notable differences for the reading, first listening, and

[2]http://www.gin.cnrs.fr/AAL2/

[3]http://Neuromorphometrics.com/

[4]http://marsbar.sourceforge.net/

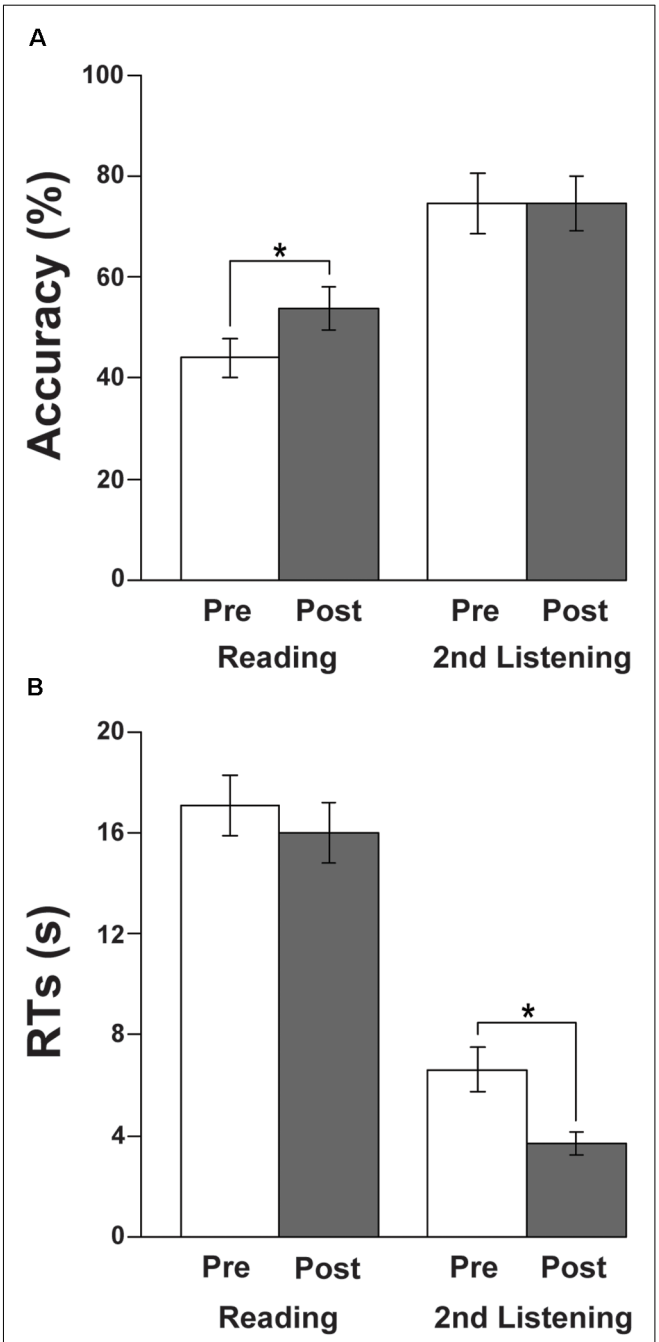

FIGURE 2 | Behavioral results. **(A)** Accuracy of the reading and listening tests. Note the significant improvement in the reading tests from Pre to Post sets during the course of learning Japanese. **(B)** Response times (RTs) of the reading and listening tests. Note the significant improvement (i.e., reduction in RTs) in the listening tests from Pre to Post. The error bars denote the SEM (standard error of the mean). *$p < 0.05$ (paired t-test).

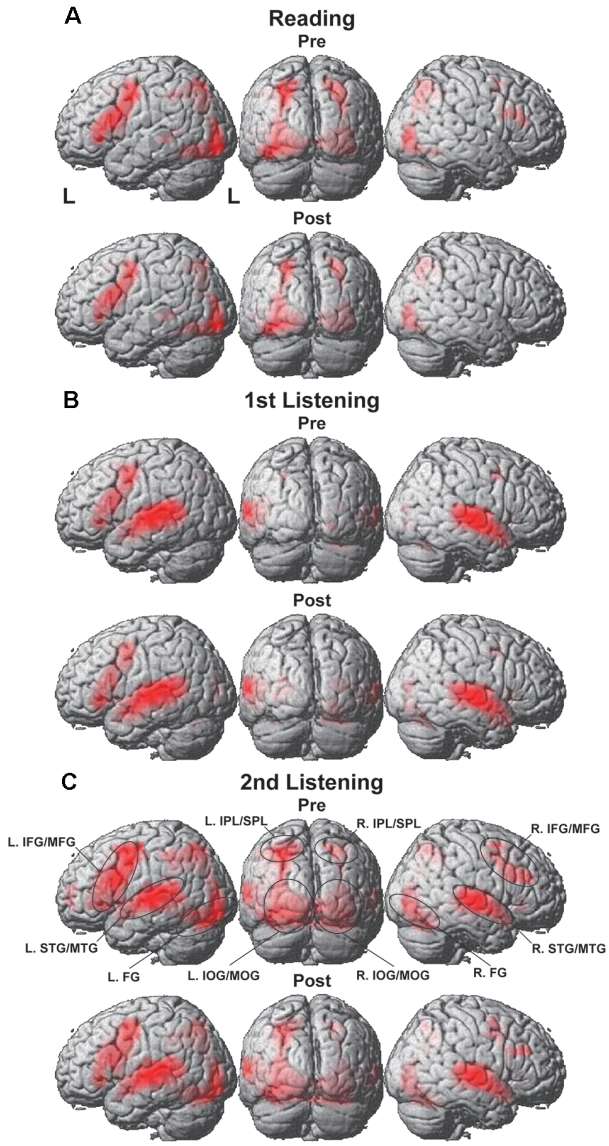

FIGURE 3 | Overall cortical activation during the presentation of the Pre or Post sets. **(A)** Reading events, **(B)** first listening events, and **(C)** second listening events. Activations were projected onto the left (L) lateral, back, and right lateral surfaces of a standard brain (thresholded at uncorrected $p < 0.0001$ for the voxel level, and at corrected $p < 0.05$ for the cluster level). Regions of interest (ROIs) are denoted with black ellipses. Visual and audio areas were activated bilaterally, while the language-related frontal activations were left-dominant. IFG/MFG, the inferior and middle frontal gyri; STG/MTG, the superior and middle temporal gyri; IPL/SPL, the inferior and superior parietal lobules; FG, the fusiform gyrus; and IOG/MOG, the inferior and middle occipital gyri.

second listening events. During the reading events (**Figure 3A**), prominent activation was observed in the left-dominant IFG/MFG, as well as in the bilateral IPL/SPL, FG, and IOG/MOG, i.e., the language areas and visual areas. This activation pattern was similar for the Pre and Post sets (see **Figure 3** for lateral and back views, and **Figure 4** for parasagittal planes). After including the individual intervals between the Pre and Post sets as a nuisance factor in a general linear model, we obtained the same activation patterns for the Post sets (**Supplementary Figure 1**).

During the first listening events (**Figure 3B**), the left-dominant activation in the IFG/MFG remained, while the visual areas were completely replaced by the bilateral

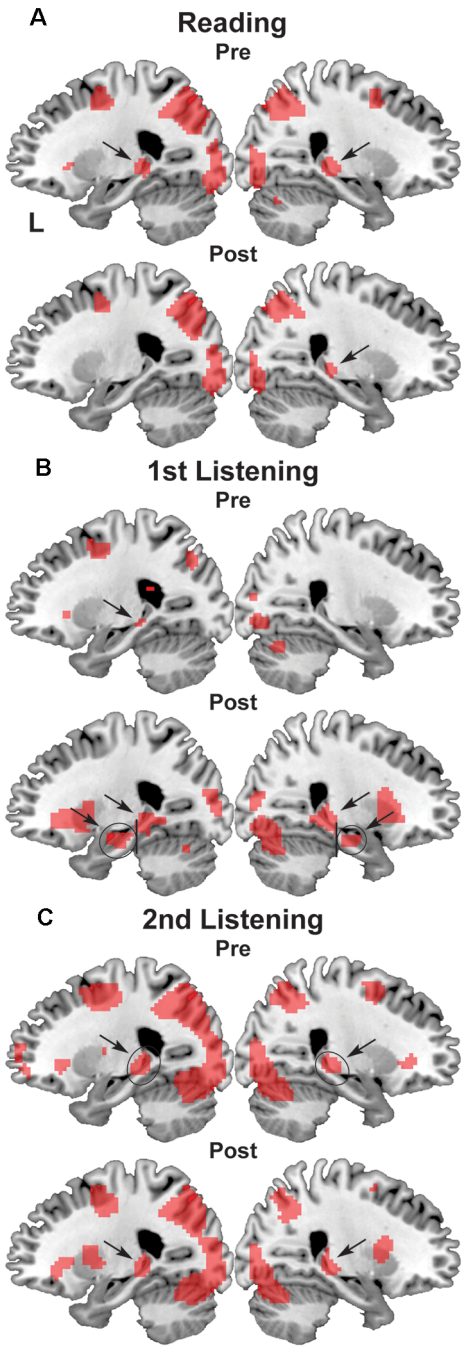

FIGURE 4 | Activation in the bilateral hippocampus. **(A)** Reading events, **(B)** first listening events, and **(C)** second listening events. On the parasagittal planes ($x = \pm 24$), each arrow denotes the posterior or anterior hippocampus. ROIs are denoted with black ellipses.

STG/MTG, a part of the auditory areas. Moreover, overall cortical activation was greatly increased during the *second* listening events (**Figure 3C**), in which the same speech sounds were presented again together with visual stimuli (see **Figure 1**). In addition to the bilateral IFG/MFG and STG/MTG, visual areas of the bilateral IPL/SPL, FG, and IOG/MOG were also markedly activated during the second listening events in both the Pre and Post sets, much more than during the first listening events. These results indicate the synergistic effects of multiple modalities (i.e., vision, audition, and language), in that the sentence comprehension of auditory stimuli could be enhanced by questions and possible answers provided by visual information.

We also found significant activation in the bilateral hippocampi, which was evident on the sagittal sections of $x = \pm 24$ (**Figure 4**). While the posterior hippocampus [MNI coordinates of local maxima: (−21, −34, −1) and (24, −28, −4)] showed consistent activation for all events (see **Figure 4C**), the anterior hippocampus [MNI coordinates of local maxima: (−27, −16, −16) and (24, −13, −19)] exhibited selective activation during the first listening events (see **Figure 4B**). This result indicates the critical role of the anterior hippocampus for the initial encoding of auditory information (see **Figure 1**).

Lateralization and Changes in Activation Between Pre and Post Sets

To quantify and compare activation changes from Pre to Post for each type of event, as well as activations between ROIs of both hemispheres, we calculated mean signal changes in each ROI, separately for the Pre and Post sets (**Figure 5**). First, we observed a significant *reduction* in the L. IFG/MFG activations from Pre to Post during the *second listening* events (paired *t*-test; $t_{(14)} = -2.3$, $p = 0.041$; **Figure 5A**), though this effect was not evident in the reading or first listening events (reading: $t_{(14)} = -0.79$, $p = 0.44$; first listening: $t_{(14)} = -0.95$, $p = 0.36$). The R. IFG/MFG did not show significant changes during any of the events. Moreover, activations in the L. IFG/MFG were significantly larger than those in the R. IFG/MFG during all events except the Pre sets in the reading events. This lateralization is consistent with the critical role of the L. IFG/MFG as a grammar center. Regarding some brain regions without significant activation changes during this short period, those regions may not be critically involved in the acquisition processes, although negative results cannot exclude possible involvement.

In contrast, the R. STG/MTG showed a significant *increase* in signal changes from Pre to Post during the *second listening* events ($t_{(14)} = 2.3$, $p = 0.038$), whereas this increase was marginal during the first listening events ($t_{(14)} = 1.8$, $p = 0.089$; **Figure 5B**), indicating that the auditory areas became more sensitive during the course of acquiring a new language, at least at this early stage. This activation increase was in marked contrast with the results in the visual areas—namely, each side of the IPL/SPL showed a significant *decrease* in signal changes from Pre to Post during the *reading* events (L. IPL/SPL: $t_{(14)} = -3.8$, $p < 0.005$; R. IPL/SPL: $t_{(14)} = -3.0$, $p = 0.0097$; **Figure 5C**). The L. IOG/MOG also showed a similarly significant tendency during the reading events ($t_{(14)} = -2.3$, $p = 0.040$; **Figure 5E**). On the other hand, the L. FG activation was consistently stronger than the R. FG activation during the *reading* events, just like the IOG/MOG, while this tendency was reversed during the *first listening* events (**Figure 5D**). These results suggest the existence of differential

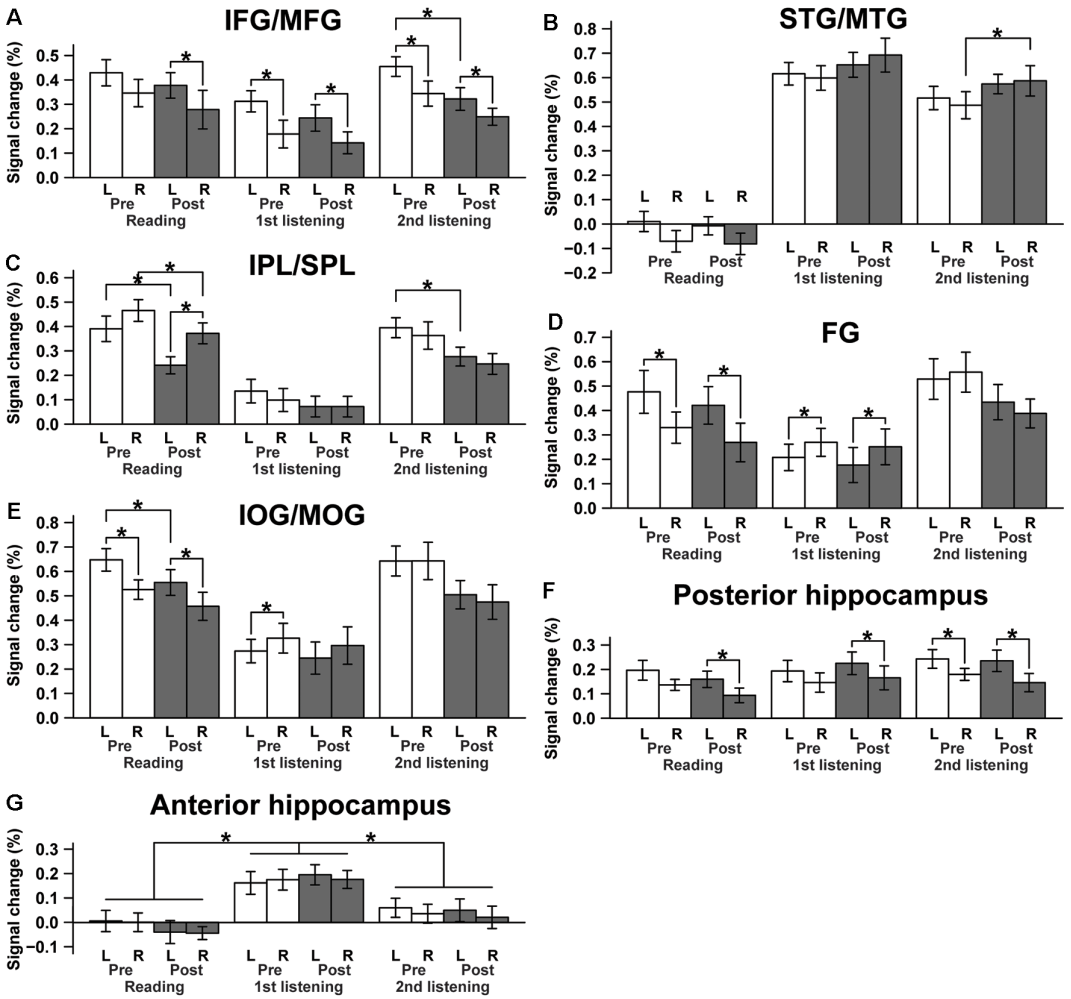

FIGURE 5 | Lateralization and changes in activations from Pre to Post for each ROI. **(A)** The IFG/MFG, **(B)** the STG/MTG, **(C)** the IPL/SPL, **(D)** the FG, **(E)** the IOG/MOG, **(F)** the posterior hippocampus, and **(G)** the anterior hippocampus. *$p < 0.05$ (paired t-test).

control mechanisms of activation changes depending on the modalities and hemispheric regions.

On the other hand, the posterior and anterior hippocampi showed different activation patterns, exhibiting differences related to hemispheres or events rather than Pre/Post sets. Therefore, we used an rANOVA with three factors (hemisphere [left, right] × set [Pre, Post] × event [reading, first listening, second listening]). The posterior hippocampus showed a significant main effect of the hemisphere ($F_{(1,14)} = 14, p < 0.0005$) with larger activations in the left hippocampus (**Figure 5F**). In contrast, the anterior hippocampus showed a significant main effect of the event ($F_{(2,28)} = 15$, $p < 0.0005$) with increased activations during the *first listening* events alone, but without a hemispheric or set difference ($p > 0.5$), i.e., no significant learning effects (**Figure 5G**). A *post hoc* analysis further showed marked differences between the first listening events and others (reading, $t_{(14)} = 4.9$, $p < 0.0005$; second listening, $t_{(14)} = 6.0$, $p < 0.0005$). These results suggest different functional roles of the posterior and anterior hippocampi.

DISCUSSION

By administering a series of reading and listening tests (**Figure 1**) to participants who had begun to learn Japanese abroad, we obtained the following major findings. First, the left-dominant language areas, as well as bilateral visual and auditory areas, were activated (**Figure 3**), demonstrating the synergistic effects of multiple modalities. There was also significant activation in the bilateral hippocampi (**Figure 4**), indicating the expected involvement of memory-related processes. Second, consistent with the behavioral improvements from Pre to Post (**Figure 2**), the brain activations *decreased* in the L. IFG/MFG during the listening tests, as well as in the visual areas (the bilateral IPL/SPL and L. IOG/MOG) during the reading tests, while activations in the R. STG/MTG *increased* during the listening tests (**Figure 5**). These modality-dependent activation changes could not be explained by domain-general cognitive factors, such as habituation or familiarization, because we used completely different test sets for Pre and Post. Third, the posterior

hippocampus showed a main effect of the hemisphere, whereas the anterior hippocampus showed a significant main effect of the event (i.e., specific to first listening events), reflecting initial encoding of auditory information alone. In summary, activation changes from Pre to Post indicate functional changes in modality-dependent networks over a short period of staying abroad, which would enable effective acquisition in an L2.

The relevance of these findings to the second language acquisition was evident from the involvement of sensory areas and language-related regions. Moreover, the decreased activations in the L. IFG/MFG are consistent with our previous hypothesis (Sakai, 2005) that such decreases reflect the saving of extraneous energy to process syntactic structures in L2. Although a further longitudinal evaluation of our participants was not possible, brain activations during the period of consolidation for the acquired linguistic knowledge would directly verify the hypothesis. The decreased activations in the visual areas may also be explained by reduced visual attention to letters and words when skills or proficiency in reading texts improved. On the other hand, the increased activations in the R. STG/MTG would be due to an improved ability to catch the sound patterns in L2, which further facilitate auditory attention to speech sounds as well. Further studies are required to clarify when activations in the bilateral STG/MTG turn to a decrease or become more focal during the subsequent development of L2 acquisition.

A previous study suggests that hemispheric asymmetry or lateralization of L2 processing depends on the similarity between L1 and L2 (D'Anselmo et al., 2013). The activation differences found in the present study might be affected by such linguistic factors, where Japanese belongs to a type of language different from the Indo-European family as L1 for the participants. Moreover, the use of a complicated writing system in Japanese (*kana*, *kanji*, and the Roman alphabet) would have required additional visual processes of identifying letters for inexperienced participants. However, our previous fMRI results of acquiring English as L2 for Japanese participants, i.e., in the reverse direction between languages from L1 to L2, have already established the basic commonality between L1 and L2 for both anatomical and functional properties in the language areas (Sakai et al., 2004, 2009; Tatsuno and Sakai, 2005; Nauchi and Sakai, 2009).

The anterior hippocampus, which is typically located at MNI coordinates close to those in the present study (22, −20, −18), has been suggested to have multiple cognitive functions: perception, imagination, and episodic memory (e.g., recall of scenes and events; Zeidman and Maguire, 2016). These proposed functional roles are consistent with the event-related activations in the present study. Indeed, perceptual *encoding* was particularly required for the initial exposure to auditory information, together with the ability to imagine what was being described in the presented speech and conversation, which were required most for the *first listening* events among the events we tested. It might be possible that increased activations during the first listening, but not during the second listening events, suggest domain-general responses to novel auditory/visual stimuli, but the complete absence of activations in the anterior hippocampus during the reading events (see **Figure 5G**) excluded this possibility. On the other hand, an fMRI study on associative memory showed that activations in the left posterior hippocampus were increased more during retrieval, while the left anterior hippocampus was activated more during encoding (Prince et al., 2005). The region was centered at Talairach coordinates of (−19, −41, 6), which were also close to those in the present study. Our results are consistent with this interesting double dissociation, in that the left posterior hippocampus was activated irrespective of the events we tested, but only when specific linguistic information had to be retrieved to solve the tests. This finding suggests the possible network for these subregions of the hippocampus and the modality-specific cortical regions, indicating that dynamic encoding and retrieval processes are involved in the acquisition of a new language.

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 Institutional Review Board of The University of Tokyo, Komaba Campus. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

KLS and TK designed the study. TK conducted the experiment. KLS, TK, SY, and KM analyzed the data and wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The authors declare that this study received funding from EF Japan. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

ACKNOWLEDGMENTS

We would like to thank Christopher McCormick in EF for the initial setup and encouragement, Sang-Chul (Sange) Lee, Kris Kullengren, and M. K. Shah in EF for coordinating our collaboration, Keita Umejima and Naoko Komoro for technical assistance, and Hiromi Matsuda for administrative assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnbeh.2021.631957/full#supplementary-material.

REFERENCES

Chee, M. W. L., Hon, N., Lee, H. L., and Soon, C. S. (2001). Relative language proficiency modulates BOLD signal change when bilinguals perform semantic judgments. *NeuroImage* 13, 1155–1163. doi: 10.1006/nimg.2001.0781

D'Anselmo, A., Reiterer, S., Zuccarini, F., Tommasi, L., and Brancucci, A. (2013). Hemispheric asymmetries in bilinguals: tongue similarity affects lateralization of second language. *Neuropsychologia* 51, 1187–1194. doi: 10.1016/j.neuropsychologia.2013.03.016

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., and Frackowiak, R. S. J. (1995). Statistical parametric maps in functional imaging: a general linear approach. *Hum. Brain Mapp.* 2, 189–210. doi: 10.1002/hbm.460020402

Hashimoto, R., and Sakai, K. L. (2004). Learning letters in adulthood: direct visualization of cortical plasticity for forming a new link between orthography and phonology. *Neuron* 42, 311–322. doi: 10.1016/s0896-6273(04)00196-5

Hosoda, C., Tanaka, K., Nariai, T., Honda, M., and Hanakawa, T. (2013). Dynamic neural network reorganization associated with second language vocabulary acquisition: a multimodal imaging study. *J. Neurosci.* 33, 13663–13672. doi: 10.1523/JNEUROSCI.0410-13.2013

Kinno, R., Ohta, S., Muragaki, Y., Maruyama, T., and Sakai, K. L. (2014). Differential reorganization of three syntax-related networks induced by a left frontal glioma. *Brain* 137, 1193–1212. doi: 10.1093/brain/awu013

Li, P., Legault, J., and Litcofsky, K. A. (2014). Neuroplasticity as a function of second language learning: anatomical changes in the human brain. *Cortex* 58, 301–324. doi: 10.1016/j.cortex.2014.05.001

Nauchi, A., and Sakai, K. L. (2009). Greater leftward lateralization of the inferior frontal gyrus in second language learners with higher syntactic abilities. *Hum. Brain Mapp.* 30, 3625–3635. doi: 10.1002/hbm.20790

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

Opitz, B., and Friederici, A. D. (2003). Interactions of the hippocampal system and the prefrontal cortex in learning language-like rules. *NeuroImage* 19, 1730–1737. doi: 10.1016/s1053-8119(03)00170-8

Perani, D., Abutalebi, J., Paulesu, E., Brambati, S., Scifo, P., Cappa, S. F., et al. (2003). The role of age of acquisition and language usage in early, high-proficient bilinguals: an fMRI study during verbal fluency. *Hum. Brain Mapp.* 19, 170–182. doi: 10.1002/hbm.10110

Prince, S. E., Daselaar, S. M., and Cabeza, R. (2005). Neural correlates of relational memory: successful encoding and retrieval of semantic and perceptual associations. *J. Neurosci.* 25, 1203–1210. doi: 10.1523/JNEUROSCI.2540-04.2005

Reiterer, S., Pereda, E., and Bhattacharya, J. (2009). Measuring second language proficiency with EEG synchronization: how functional cortical networks and hemispheric involvement differ as a function of proficiency level in second language speakers. *Second Lang. Res.* 25, 77–106. doi: 10.1177/0267658308098997

Sakai, K. L. (2005). Language acquisition and brain development. *Science* 310, 815–819. doi: 10.1126/science.1113530

Sakai, K. L., Miura, K., Narafu, N., and Muraishi, M. (2004). Correlated functional changes of the prefrontal cortex in twins induced by classroom education of second language. *Cereb. Cortex* 14, 1233–1239. doi: 10.1093/cercor/bhh084

Sakai, K. L., Nauchi, A., Tatsuno, Y., Hirano, K., Muraishi, M., Kimura, M., et al. (2009). Distinct roles of left inferior frontal regions that explain individual differences in second language acquisition. *Hum. Brain Mapp.* 30, 2440–2452. doi: 10.1002/hbm.20681

Schlegel, A. A., Rudelson, J. J., and Tse, P. U. (2012). White matter structure changes as adults learn a second language. *J. Cogn. Neurosci.* 24, 1664–1670. doi: 10.1162/jocn_a_00240

Tanaka, K., Kinno, R., Muragaki, Y., Maruyama, T., and Sakai, K. L. (2020). Task-induced functional connectivity of the syntax-related networks for patients with a cortical glioma. *Cereb. Cortex Commun.* 1, 1–15. doi: 10.1093/texcom/tgaa061

Tatsuno, Y., and Sakai, K. L. (2005). Language-related activations in the left prefrontal regions are differentially modulated by age, proficiency and task demands. *J. Neurosci.* 25, 1637–1644. doi: 10.1523/JNEUROSCI.3978-04.2005

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage* 15, 273–289. doi: 10.1006/nimg.2001.0978

Wartenburger, I., Heekeren, H. R., Abutalebi, J., Cappa, S. F., Villringer, A., and Perani, D. (2003). Early setting of grammatical processing in the bilingual brain. *Neuron* 37, 159–170. doi: 10.1016/s0896-6273(02)01150-9

Yamamoto, K., and Sakai, K. L. (2016). The dorsal rather than ventral pathway better reflects individual syntactic abilities in second language. *Front. Hum. Neurosci.* 10, 1–18. doi: 10.3389/fnhum.2016.00295

Yamamoto, K., and Sakai, K. L. (2017). Differential signatures of second language syntactic performance and age on the structural properties of the left dorsal pathway. *Front. Psychol.* 8, 1–13. doi: 10.3389/fpsyg.2017.00829

Zeidman, P., and Maguire, E. A. (2016). Anterior hippocampus: the anatomy of perception, imagination and episodic memory. *Nat. Rev. Neurosci.* 17, 173–182. doi: 10.1038/nrn.2015.24

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 in Behavioral Neuroscience

REVIEW
published: 09 June 2021
doi: 10.3389/fnbeh.2021.640659

The Roles of the Cortical Motor Areas in Sequential Movements

Machiko Ohbayashi[1,2]*

[1] Department of Neurobiology, University of Pittsburgh School of Medicine, Pittsburgh, PA, United States, [2] Systems Neuroscience Center, Center for the Neural Basis of Cognition, University of Pittsburgh, Pittsburgh, PA, United States

The ability to learn and perform a sequence of movements is a key component of voluntary motor behavior. During the learning of sequential movements, individuals go through distinct stages of performance improvement. For instance, sequential movements are initially learned relatively fast and later learned more slowly. Over multiple sessions of repetitive practice, performance of the sequential movements can be further improved to the expert level and maintained as a motor skill. How the brain binds elementary movements together into a meaningful action has been a topic of much interest. Studies in human and non-human primates have shown that a brain-wide distributed network is active during the learning and performance of skilled sequential movements. The current challenge is to identify a unique contribution of each area to the complex process of learning and maintenance of skilled sequential movements. Here, I bring together the recent progress in the field to discuss the distinct roles of cortical motor areas in this process.

Keywords: sequential movements, motor skills, learning, motor cortex, premotor cortex, SMA

OPEN ACCESS

Edited by:
Yuji Naya,
Peking University, China

Reviewed by:
Hidetoshi Amita,
Kyoto University, Japan
Milena Raffi,
University of Bologna, Italy

***Correspondence:**
Machiko Ohbayashi
machiko@pitt.edu

Specialty section:
This article was submitted to Learning and Memory, a section of the journal Frontiers in Behavioral Neuroscience

Received: *11 December 2020*
Accepted: *19 April 2021*
Published: *09 June 2021*

Citation:
Ohbayashi M (2021) The Roles of the Cortical Motor Areas in Sequential Movements. Front. Behav. Neurosci. 15:640659. doi: 10.3389/fnbeh.2021.640659

INTRODUCTION

The production of sequential movements is a fundamental aspect of voluntary behavior. Many of our daily actions, such as playing a musical instrument, handwriting, typing, etc., depend on attaining a high level of skill in the performance of sequential movements. The performance of sequential movements can be acquired and improved to the expert level through extensive practice (Rosenbaum, 2010). Such performance can be maintained as a motor skill. How the brain binds elementary movements together into skilled sequential movements has been a fundamental problem of systems neuroscience.

The neural basis of sequential movements has been extensively studied in human and non-human primates. Human imaging studies have shown that a brain-wide distributed network, which is composed of the presupplementary motor area (pre-SMA), supplementary motor area (SMA), dorsal premotor cortex (PMd), primary motor cortex (M1), primary somatosensory cortex, superior parietal lobule, thalamus, basal ganglia, and the cerebellum, subserves the learning and performance of skilled sequential movements (e.g., Shibasaki et al., 1993; Grafton et al., 1994, 1995; Karni et al., 1995; Hikosaka et al., 1996, 2002; Sakai et al., 1998; Ungerleider et al., 2002; Dayan and Cohen, 2011; Hardwick et al., 2013). Studies in non-human primates showed that neural activity of these areas reflected aspects of sequences (e.g., Mushiake et al., 1991; Mushiake and Strick, 1993, 1995; Tanji and Shima, 1994, 1996b; Clower and Alexander, 1998; Nakamura et al., 1998; Miyachi et al., 2002; Lee and Quessy, 2003; Matsuzaka et al., 2007; Picard et al., 2013; Ohbayashi et al., 2016). Nevertheless, the results sometimes appear

contradictory. This could be due to the various cognitive demands associated with different formulations of sequential tasks. Several features of sequential movement tasks can be identified, such as guidance (i.e., guided by external *vs.* internal cues), memory (i.e., guided by short-term *vs.* long-term memory), movement outcome (i.e., configuration *vs.* spatial position), and movement flow (i.e., with temporal separation *vs.* continuous). It was suggested that spatial and non-spatial sequences may be learned and controlled by different cortical circuits (Tanji, 2001; Ohbayashi et al., 2016). Furthermore, neural activity has been shown to change as a result of practice on motor skill tasks (e.g., Grafton et al., 1994; Karni et al., 1995; Sakai et al., 1998; Coynel et al., 2010).

In this review, I will focus on the roles of the SMA, PMd, and M1 in skilled sequential movements, i.e., those acquired through repetitive practice and internally generated from long-term memory. I will especially focus on the spatial sequence tasks as this type of task was used in non-human primate studies after extensive practice (**Table 1**). The current challenge is to identify a unique contribution of each area to the complex process of acquisition and retention of sequential movements. Interventional studies in non-human primates could represent a valuable complement to neuroimaging studies. These methods can critically address the causal relationship between the activity in a brain area and behavior. I will aim to integrate recent discoveries regarding the cortical control of skilled sequential movements at multiple levels of complexity by highlighting interventional (e.g., inactivation) studies in non-human primates.

LEARNING OF SKILLED SEQUENTIAL MOVEMENTS

An important characteristic of learning skilled sequential movements is that individuals seem to go through several learning stages (Fitts and Posner, 1967; Hikosaka et al., 2002; Doyon et al., 2003; Doyon and Benali, 2005; Rosenbaum, 2010; Dayan and Cohen, 2011; Schmidt and Lee, 2011). An improvement of performance can be detected as changes in the speed and accuracy during learning (**Figure 1**). During an initial learning stage of skilled sequential movements, a subject improves performance of sequential movements relatively fast. The subject tends to make a large number of errors and highly variable movements with a lack of consistency from trial to trial, but achieves large performance improvements (Fitts and Posner, 1967; Rosenbaum, 2010; Dayan and Cohen, 2011; Schmidt and Lee, 2011). Later, the subject improves the performance more slowly over multiple sessions of practice by making fewer and smaller errors. The durations of the learning stages are highly specific to the tasks, subjects, and definitions. For example, the fast stage of learning to perform a sequential finger opposition task was defined as an initial within-session improvement phase in a human study (Karni et al., 1995, 1998), whereas the fast stage of learning to play an advanced piano piece may last months. Monkeys performing a sequential reaching task rapidly improved their response speed over about 50 days (Matsuzaka et al., 2007). Despite timing differences, the learning curves on different skill tasks follow the same patterns of initially fast, then slowing performance improvements with further practice. Through extensive and repetitive training, subjects can further improve their performance to the expert level. Then, the skill will become almost automatic with very small variability and small improvement (Fitts and Posner, 1967; Rosenbaum, 2010; Schmidt and Lee, 2011).

The progress in the learning of sequential movements is associated with a shift in functional MRI (fMRI) activation from the anterior regions to the posterior regions of the brain (Grafton et al., 1994; Sakai et al., 1998; Coynel et al., 2010). The change in fMRI activation is shown to be associated with improvement in the task performance during learning (Bassett et al., 2015; Reddy et al., 2018). This suggests that the extent of contribution of each area may change during learning. Hikosaka et al. (2002) proposed that a subject learns the spatial features of sequences during the fast learning stage and then learns the motor features of the sequences during the slow learning stage. In the following sections, I will discuss the contributions of the SMA, PMd, and M1 to the learning and performance of spatial sequence tasks and how the skilled sequential movements are maintained after extensive practice.

SUPPLEMENTARY MOTOR AREA

Classically, the preparation for and the generation of sequential movements have been thought to depend on the supplementary motor area (SMA) and the pre-SMA (Roland et al., 1980; Brinkman, 1984; Goldberg, 1985; Dick et al., 1986; Halsband, 1987; Halsband et al., 1993; Tanji and Shima, 1994, 1996a,b; Grafton et al., 1995; Shima et al., 1996; Tanji et al., 1996; Gerloff et al., 1997; Picard and Strick, 1997, 2001; Nakamura et al., 1998; Shima and Tanji, 1998, 2000; Tanji, 2001; Hikosaka et al., 2002). Human patients with lesions that include these areas had deficits in performing self-initiated movements, sequential movements, and/or speech (Goldberg, 1985; Dick et al., 1986; Halsband et al., 1993). In agreement with the reports of human patients, studies in non-human primates clearly demonstrated the contributions of the SMA and pre-SMA to the learning or performance of sequence tasks composed of non-spatial movements (Brinkman, 1984; Halsband, 1987; Tanji and Shima, 1994, 1996a,b; Shima et al., 1996; Tanji et al., 1996; Shima and Tanji, 1998, 2000, 2006; Tanji, 2001; **Table 2**). Neural recordings in monkeys demonstrated that neurons in the SMA and the pre-SMA respond preferentially to a specific order of movements rather than a single movement (Tanji and Shima, 1994, 1996a; Shima and Tanji, 2000). The inactivation of these areas in monkeys demonstrated the contributions of the SMA and pre-SMA in the performance of a sequence composed of non-spatial movements (Shima and Tanji, 1998). When the SMA or pre-SMA was bilaterally inactivated by injecting muscimol ($GABA_A$ agonist), the inactivation disrupted the monkey's performance of sequences of arm movements guided by memory, leaving the execution of simpler, single movements unaffected (Shima and Tanji, 1998).

On the other hand, non-human primate studies using spatial sequence tasks suggested that spatial and non-spatial sequences may be learned and controlled by different cortical circuits (Tanji, 2001; Ohbayashi et al., 2016). Even though neural activity reflected a specific order of movements in a sequence in both types of tasks, the results of an inactivation study using a spatial sequence task were different from the results using the non-spatial sequence task. Neurons in the SMA and the pre-SMA respond preferentially to a specific order of movements rather than a single movement during the performance of internally generated spatial sequence tasks (Clower and Alexander, 1998; Nakamura et al., 1998; Lee and Quessy, 2003). The activity of the SMA neurons reflected a particular serial position in a sequence (Lee and Quessy, 2003). The pre-SMA is particularly active for the learning of new sequences of movements, but not the production of movement components (e.g., reaching) (Hikosaka et al., 1996). Furthermore, the 2-deoxyglucose (2DG) signals of these areas after extensive practice (>12 months) reflected the effect of long-term training on spatial sequences (Picard and Strick, 1997, 2003). The 2DG signal is suggested to be associated with presynaptic activity at both excitatory and inhibitory synapses and reflects the metabolic activity of synapses (discussed in Picard and Strick, 2003). In the studies, the monkeys were trained on remembered sequential movements or visually guided reaching for years (>12 months). After extensive practice, both the SMA and the pre-SMA displayed substantial uptakes of 2DG in association with visually guided reaching movements (Picard and Strick, 2003). On the other hand, 2DG incorporation in the SMA and pre-SMA was relatively low in the case of remembered sequential reaching movements (Picard and Strick, 1997). The differential metabolic activities of the pre-SMA and SMA in the two tasks suggested that these areas may be reorganized with overtraining on the remembered sequences after extensive practice. Therefore, the results from neural recording and 2DG showed that neurons in both the SMA and pre-SMA may play roles in the learning and performance of spatial sequence tasks.

Nevertheless, an inactivation study using spatial sequences provided results different from the inactivation study using

TABLE 1 | Non-human primate studies on memory-guided sequential reaching tasks.

Study	Area	Training duration	Method	Main findings
Lee and Quessy, 2003	SMA	1 day	Neural recording	About one-third of the neurons in the SMA displayed gradual changes in their activity across different trials when a particular movement sequence was repeatedly performed.
Mushiake et al., 1991	SMA, PMd, M1	3–5 months	Neural recording	More than a half of the SMA neurons were preferentially active during memory-guided sequential reaching (55% and 65% during the pre-movement and movement periods). More than a half of the PM neurons were preferentially active in visually guided reaching (55% and 64% during the pre-movement and movement periods). M1 neurons showed similar activity regardless of whether it was guided by memory or visual stimulus.
Clower and Alexander, 1998	SMA	–	Neural recording	Neurons in the SMA and pre-SMA reflected the numerical order of the specific movement component of a sequence.
Nakamura et al., 1998	SMA	8 months–2 years	Neural recording	Neurons in the SMA and pre-SMA respond preferentially to a specific order of movements.
Nakamura et al., 1999	SMA	8 months–2 years	Muscimol injection	Inactivation of the SMA did not disrupt the learning or performance of the sequential reaching guided by memory.
Picard and Strick, 1997	SMA	∼39 months	2DG	2DG uptakes in the SMA and pre-SMA were relatively low in the remembered sequential reaching movements.
Picard and Strick, 2003	SMA	12–17 months	2DG	2DG uptakes in the SMA and pre-SMA were substantial in the visually guided reaching.
Ohbayashi et al., 2016	PMd	>50 days	Neural recording, muscimol injection	Injection of muscimol disrupted the performance of the memory-guided sequential reaching, but not the visually guided reaching. Forty-three percent of the neurons were differentially active during the memory-guided sequential reaching and visually guided reaching.
Lu and Ashe, 2005	M1	∼6 months	Neural recording, muscimol injection	Neurons exhibited anticipatory activity related to specific sequences. After muscimol injection, the number of errors in the sequential movements increased.
Ohbayashi, 2020	M1	>100 days	Anisomycin injection, muscimol injection	Anisomycin injection disrupted the performance of the memory-guided sequential reaching, but not the visually guided reaching. Muscimol injection disrupted the performance of both the memory-guided sequential reaching and visually guided reaching.
Matsuzaka et al., 2007	M1	>2 years	Neural recording	∼40% of the task-related neurons were differentially active during the memory-guided sequential reaching and visually guided reaching.
Picard et al., 2013	M1	∼1–6 years	2DG, neural recording	2DG uptake was lower in monkeys that performed sequential reaching guided by memory compared with the 2DG uptake in monkeys that performed visually guided reaching. 2DG uptake was lower in monkeys that were trained for a longer duration.

SMA, supplementary motor area; PMd, dorsal premotor cortex; M1, primary motor cortex; 2DG, 2-deoxyglucose.–: Training duration was not provided in the manuscript.

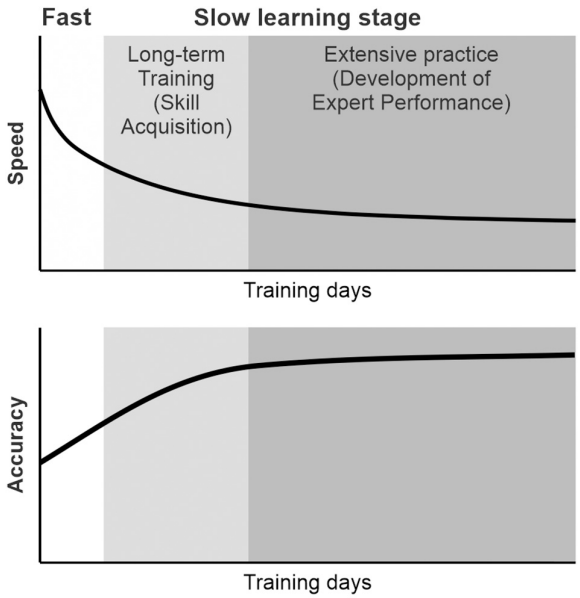

FIGURE 1 | Schematic diagram of the learning of skilled sequential movements. During learning, individuals seem to go through several learning stages. The subject improved the performance of sequential movements relatively fast initially and, later, more slowly over multiple sessions of practice. Through repetitive practice, subjects improve their speed and accuracy of sequential movements. The performance of sequential movements can be improved to the expert level through extensive practice and can be maintained as a motor skill.

non-spatial sequences. Hikosaka's group trained monkeys to learn a spatial sequence of reaching movements to targets (Nakamura et al., 1999). In their sequence task, reaching movements in space were required, so that the selection of spatial end points in sequential reaching was a critical factor to control. When the pre-SMA was bilaterally inactivated by injecting muscimol, the inactivation disrupted the monkeys' learning of a new sequence of movements, but not the performance of the memorized sequence of movements. Interestingly, local inactivation of the SMA did not significantly disrupt the learning or the performance of the sequential reaching task (Nakamura et al., 1999). The result suggests that the SMA may not be critically involved in the performance of this type of spatial sequences at the tested stage of learning, even though neurons exhibited the sequence-related activity. Clearly, the inactivation results highlight the most unique contribution of the targeted area to the sequence task.

Taken these findings together, the pre-SMA seems to be more critically involved in the cognitive aspects of acquiring a novel sequence of movements compared to the SMA (Hikosaka et al., 1996; Shima et al., 1996; Nakamura et al., 1998, 1999; Shima and Tanji, 2000, 2006). The SMA seems to be involved in the temporal organization of multiple non-spatial movements into a sequence (Boecker et al., 1998; Shima and Tanji, 1998; Tanji, 2001; Nachev et al., 2008; Orban et al., 2010; Wiener et al., 2010; Cona and Semenza, 2017). On the other hand, for the spatial sequence tasks, even though neural activity of the SMA neurons reflected aspects of sequences, its role is still debatable and needs to be further investigated. The results of the spatial sequence task suggested that the effect of muscimol injection in the SMA could be compensated by another motor area, possibly the PMd, which is anatomically connected with both the SMA and M1. In the next section, I will discuss the role of the PMd in the performance of internally generated sequential movement tasks.

DORSAL PREMOTOR CORTEX

The dorsal premotor cortex (PMd) has been regarded as the area for the visual guidance of motor behavior in many studies (Kalaska and Crammond, 1995; Johnson et al., 1996; Wise et al., 1997; Hoshi and Tanji, 2007; Averbeck et al., 2009). Moreover, the PMd is suggested to be involved in the cognitive aspects of visually guided motor tasks, such as mental rehearsal and a decision making (Cisek and Kalaska, 2002, 2004, 2005; Pesaran et al., 2008). Considerable evidence suggests that the PMd is specifically involved in the guidance of movements based on memorized arbitrary sensorimotor associations (Passingham, 1988; Mitz et al., 1991; Kurata and Hoffman, 1994). Firstly, lesions or the inactivation of the PMd produces deficits on tasks that rely on the associations between an arbitrary visual cue (e.g., color or shape of a visual stimulus) and a movement (Halsband and Passingham, 1982; Passingham, 1988; Kurata and Hoffman, 1994). For example, Kurata and Hoffman trained monkeys to learn the visuo-motor association task in which the monkeys were required to move their wrist to the right or the left direction based on the color of a conditional cue (Kurata and Hoffman, 1994). Then, they locally inactivated the PMd by injecting a small amount of muscimol at sites where the preparatory neural activity was recorded during the performance of the conditional visuo-motor association task. The local inactivation of PMd disrupted the monkeys' performance of the visuo-motor association task. Secondly, neurons in the PMd show a sustained activity that is specifically related to the performance of these visuo-motor association tasks (Kurata and Wise, 1988; Mitz et al., 1991; Kurata and Hoffman, 1994). The PMd neurons showed sustained activity after the presentation of the arbitrary visual cue during the movement preparation period (Kurata and Wise, 1988; Kurata and Hoffman, 1994). Although these findings are in line with the proposal that the PMd plays a crucial role in the visual guidance of movements in general, they specifically point to the important contribution of the PMd to memory-guided movements in which selection, preparation, and execution of movements are guided by memorized visuo-motor associations (Halsband and Passingham, 1982; Wise, 1985; di Pellegrino and Wise, 1993; Kurata and Hoffman, 1994).

Moreover, human imaging studies consistently reported the activity of the PMd during the performance of sequential movements (Dayan and Cohen, 2011; Hardwick et al., 2013). The studies indicated that the PMd may be a structure of key importance for sequence learning and may contribute to sequence learning by selecting appropriate responses. This idea was verified by a study using non-human primates. The role of the PMd in internally guided sequential reaching

was studied using neural recordings and local inactivation (Ohbayashi et al., 2016). Monkeys were trained to perform two types of reaching tasks (**Figures 2A–C**). In one task, the movements were instructed by spatial visual cues (random task, visually guided reaching; **Figure 2B**), whereas in the other task, sequential movements were internally generated from memory after extended practice (repeating task, internally generated sequential movements guided by memory; **Figure 2C**; Ohbayashi et al., 2016; Ohbayashi and Picard, 2020). After more than 50 days of training on the tasks, the group examined neural activity in the arm area of the PMd, which was identified by intracortical microstimulation. About 40% of the neurons displayed responses that were enhanced in one task compared with the other (i.e., differential neurons). Approximately half of the differential neurons displayed enhanced activity during the repeating task, internally generated sequential movements. In the same study, the PMd was locally and transiently inactivated by injecting a small amount of muscimol into the arm representation area of the PMd after more than 50 days of training (**Figures 3A–E**). The inactivation of the PMd had a marked effect only on the performance of sequential movements that were guided by memory, but not on the performance of visually guided reaching (**Figures 3D,E**; Ohbayashi et al., 2016). Even though comparable numbers of neurons displayed enhanced activity during the internally guided sequential reaching and visually guided reaching, movement performance during the visually guided reaching was unaffected by the PMd inactivation. Furthermore, the monkeys made two types of errors after the inactivation of the PMd: errors of accuracy and errors in direction. Accuracy errors reveal an execution deficit: the monkeys reached in the correct direction for the next target in the sequence, but the movement end points were outside of the correct target. Direction errors indicate a deficit in the selection of the next target in a sequence: the monkeys reached in the direction opposite to the correct target. The inactivation results provide a clear demonstration of the importance of the PMd in the performance of internally generated sequential movements. Similarly, an inactivation of the left PMd of a human subject using transcranial magnetic stimulation (TMS) disrupted the performance of internally generated sequential movements (Wymbs and Grafton, 2013). In the study, human subjects practiced sequence production task using either a button box or a laptop keyboard with their right hand. After 30 days of practice, when the left PMd was stimulated, the error rate during the retrieval of practiced sequences increased.

Taken together, the results suggest that, although the PMd neurons are active during both visually guided and internally generated sequential movements, the PMd plays an important role in the internal generation of sequential movements. The inactivation results demonstrated that the PMd is involved in guiding sequential movements based on internal instructions after practice. With practice on sequential movements, the animal could learn arbitrary motor–motor associations of elements in the sequence and perform the practiced sequence in a seamless and predictive manner. Therefore, one possible interpretation is that the PMd inactivation disrupted the arbitrary motor–motor associations in the same way as lesions of the premotor cortex disrupt an animal's performance of arbitrary sensorimotor associations (Halsband and Passingham, 1982; Wise, 1985; Passingham, 1988; di Pellegrino and Wise, 1993; Kurata and Hoffman, 1994). This is consistent with human imaging studies in which performance of the serial reaction time task (SRTT) variants elicited the bilateral PMd activity (e.g., Hardwick et al., 2013). Hardwick et al. (2013) suggested that the left PMd of humans is "a critical node in the motor learning network" for sequential movements. Further studies are necessary to explore the role of the PMd during early learning and after extensive practice, as well as in different types of sequential movements such as non-spatial sequence tasks.

PRIMARY MOTOR CORTEX

The primary motor cortex (M1) controls muscle activity through its projections to the spinal cord, and its contribution to patterning muscle activity has been extensively studied (Evarts, 1981). Growing evidence suggests that M1 is involved in both the learning and maintenance of motor skills (e.g., Shibasaki et al., 1993; Karni et al., 1995, 1998; Ungerleider et al., 2002; Floyer-Lea and Matthews, 2004). For example, human imaging

TABLE 2 | Non-human primate studies on non-spatial sequence tasks guided by memory.

Study	Area	Training duration	Method	Main findings
Tanji and Shima, 1994	SMA	–	Neural recording	A group of neurons exhibited activity related to a predetermined sequence of movements.
Shima et al., 1996	Pre-SMA	–	Neural recording	A group of pre-SMA cells were active when a monkey was required to switch a sequence of movements to perform the next one.
Shima and Tanji, 1998	SMA, pre-SMA	–	Muscimol injection	An inactivation of either the SMA or the pre-SMA disrupted the performance of memorized sequential movements.
Shima and Tanji, 2000	SMA, pre-SMA	–	Neural recording	Neurons in both the SMA and pre-SMA exhibited activity at different phases in the task.
Shima and Tanji, 2006	Pre-SMA	–	Neural recording	A group of pre-SMA neurons represented odd-numbered trials within the sequential movements; others represented even-numbered trials.

pre-SMA, presupplementary motor area; SMA, supplementary motor area.–: Training duration was not provided in the manuscript.

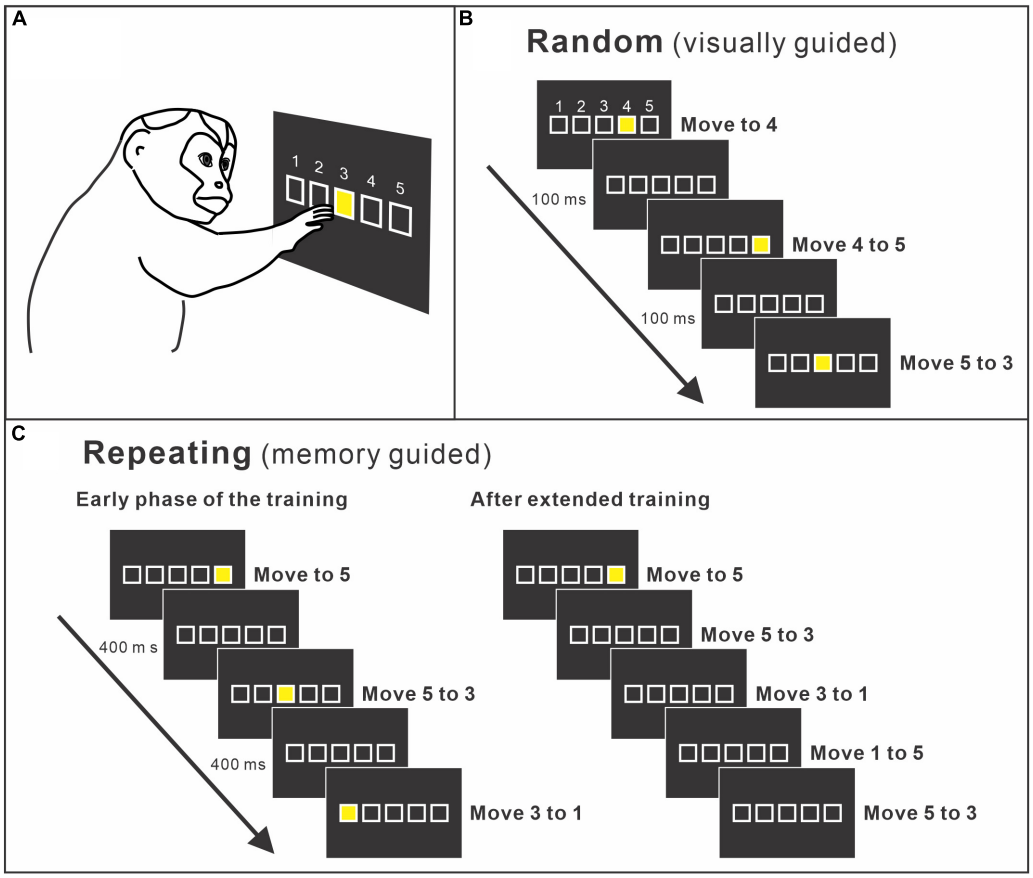

FIGURE 2 | (A) Apparatus of the training setup. A monkey sits in front of a touch-sensitive monitor. To make a correct response, the monkey is required to contact a yellow target cue displayed on the touch monitor. The *yellow target* is presented at one of five squares displayed on the touch monitor. Squares are arranged in a horizontal row and identified as numbers *1* to *5* from *left* to *right*. **(B)** Random task. A new target cue is presented in pseudo-random order in one of the five squares. A new target is presented 100 ms after the monkey made a correct response or immediately after an error. Therefore, the monkey performs visually guided reaching from a target to the next target. **(C)** Repeating task. Targets are presented according to a predetermined sequence (*left*). As the monkey learns the sequence, the monkey started to touch the target in the sequence before the presentation of the visual cue (*right*). After extended practice, the monkeys perform the task without help of visual cues (modified from Ohbayashi and Picard, 2020).

studies have shown that the fMRI blood oxygen level-dependent (BOLD) signal in M1 is modulated by the learning of sequential movement tasks. Karni et al. (1995) reported that after 3 weeks of practice on finger opposition sequences, the extent of M1 activation evoked during the performance of a trained sequence was significantly larger compared with the extent of activation evoked by the control task. The change in the BOLD signal in M1 persisted for several months. Moreover, the effects of prolonged and repetitive practice on the functional organization and cortical structure in M1 have been studied with musicians (i.e., the experts of sequential movements). The functional activation in M1 during the performance of sequential tasks is reduced or becomes more focused in professional musicians compared to non-musicians or amateurs (Hund-Georgiadis and von Cramon, 1999; Jancke et al., 2000; Krings et al., 2000; Haslinger et al., 2004; Meister et al., 2005). The reduced activation after years of extensive training is considered as evidence for the increased efficacy of the motor system and the need for a smaller number of active neurons to perform a highly trained set of sequential movements (Jancke et al., 2000; Krings et al., 2000; Haslinger et al., 2004; Meister et al., 2005). These suggested that the M1 of musicians is reorganized after years of extensive practice on sequential movements.

The view that M1 is reorganized after extensive practice on sequential movements has also been supported by studies focused on the anatomical and functional changes of musicians' M1. The volume of M1 is reported to be larger in professional musicians compared to that in amateurs or non-musicians (Amunts et al., 1997; Gaser and Schlaug, 2003a,b; Draganski and May, 2008; Herholz and Zatorre, 2012; Zatorre et al., 2012; Sampaio-Baptista and Johansen-Berg, 2017; Wenger et al., 2017). The motor representations of the body parts used for skilled performance are enlarged in professional musicians compared with non-musicians (Elbert et al., 1995; Schwenkreis et al., 2007). The structural changes were proposed to be supported by processes occurring at the synapse level, including intracortical remodeling of dendritic spines and axonal terminals, glial hypertrophy, and synaptogenesis (Anderson et al., 1994; Draganski and May,

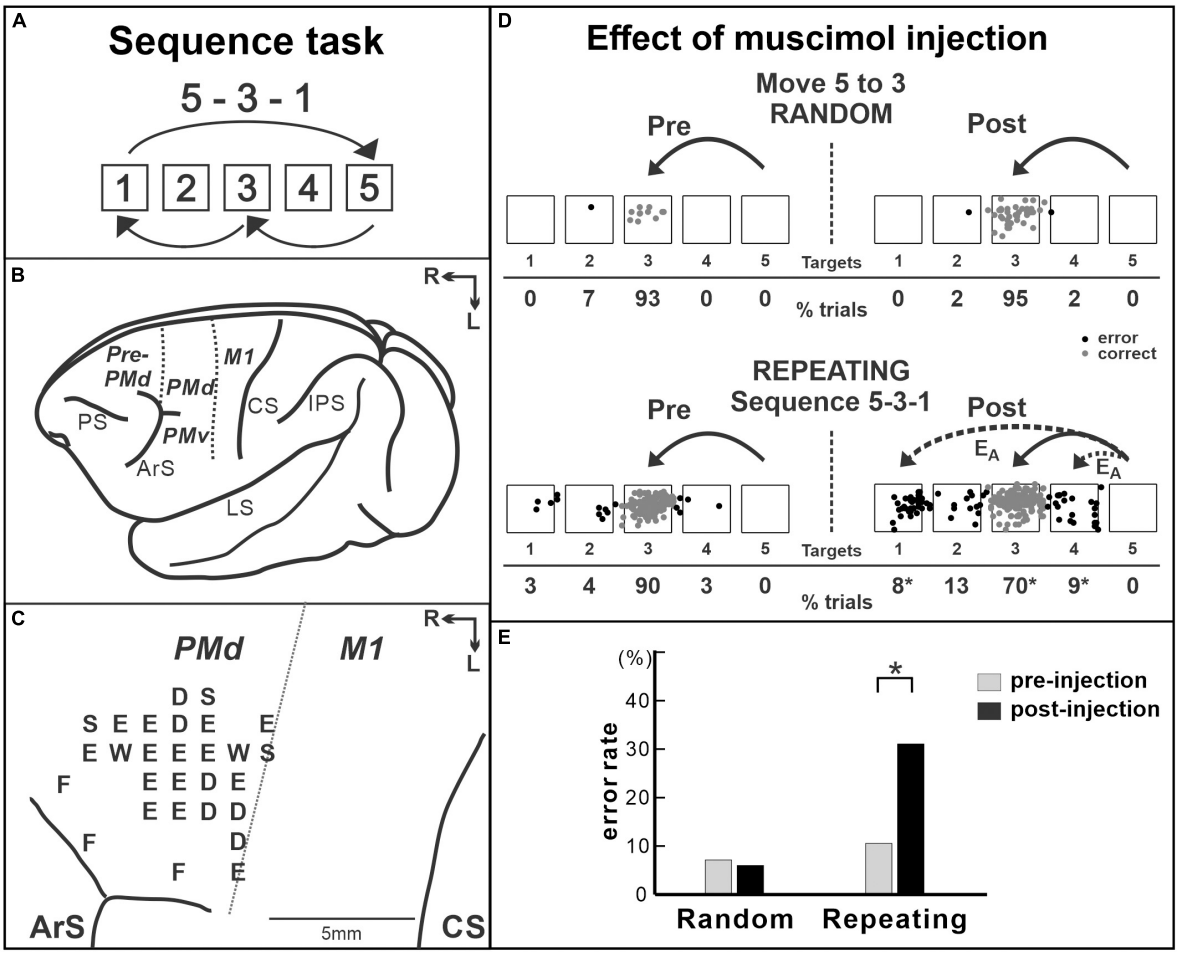

FIGURE 3 | Effect of muscimol injection in the PMd on the performance of internally generated sequential movements. **(A)** Sequence 5-3-1 of the repeating task. **(B)** Lateral view of a *Cebus* brain. *Dashed lines* indicate the M1–PMd border and the pre-PMd–PMd border. *PS*, principal sulcus; *ArS*, arcuate sulcus; *CS*, central sulcus; *IPS*, intra parietal sulcus; *LS*, lateral sulcus; *pre-PMd*, pre-dorsal premotor cortex; *PMd*, dorsal premotor cortex; *M1*, primary motor cortex; *R*, rostral; *L*, lateral. **(C)** Intracortical stimulation map of a *Cebus* monkey. *Letters* indicate the movements evoked at each site. *S*, shoulder; *E*, elbow; *W*, wrist; *D*, digit; *F*, face. Injections were done at sites in which intracortical stimulation evoked shoulder or elbow movements (i.e., arm representation) in the PMd. **(D)** Reaching end points of movements from target 5 to target 3 before and after muscimol injection in the PMd. *Left*: pre-injection; *right*: post-injection. *Top*: random task; *bottom*: repeating task. The monkey was performing sequence 5-3-1 during the repeating task. E_A: accuracy errors, a reach performed in the correct direction (e.g., to the left), but to an end point outside of the correct target. *Gray dots*: correct response; *black dots*: error response. The percentages of trials ending in each target are given *below the targets*. *$p < 0.05$. **(E)** Error rates of the random task (*left*) and the repeating task (*Right*) in the injection session in **(D)**. After muscimol injection, the number of errors increased dramatically in the repeating task, but not in the random task (modified from Ohbayashi et al., 2016. Copyright 2016 Society for Neuroscience).

2008; Herholz and Zatorre, 2012; Zatorre et al., 2012). These studies suggested that increased synaptic efficacy as a result of extensive practice may contribute to changes in structural volume. Similarly, the plasticity of the white matter structure was correlated with skill practice, such as the number of practice hours (Bengtsson et al., 2005; Han et al., 2009). Bengtsson et al. (2005) discussed that increased myelination, caused by neural activity in fiber tracts during training, could be a mechanism underlying the observed increased volume of white matter. Taken together, extensive practice on sequential movements is suggested to lead to the increased synaptic efficacy in M1 through the remodeling of dendritic spines and axonal terminals, synaptogenesis, increased myelination, and glial hypertrophy. The change of fMRI activation in the M1 of humans, decreased 2DG signal in the M1 of non-human primates, and the enlarged volume of the M1 in musicians may all reflect the reorganization in M1 with extensive practice.

Recent fMRI studies suggested that M1's contribution to structured and higher-order aspects of sequential movements may be limited when the training duration was short (Yokoi and Diedrichsen, 2019). In the study, human subjects practiced higher-order sequences that are composed of chunks of short sequences of keyboard pressing. Then, the authors examined whether this hierarchical structure was reflected in the brain activity patterns of the participants using fMRI data (Yokoi and Diedrichsen, 2019). The authors concluded that single-finger movements were represented in M1 and higher-order sequences were represented throughout the frontoparietal regions of the

cortex after 1 week of training. The neural basis for the acquisition and retention of long and high-ordered sequences after extensive practice should be further investigated in future studies.

Neurophysiological studies in non-human primates showed that the neural activity in M1 is modulated by the sequence components. When a monkey performs sequential movements, the neural activity of M1 neurons reflects aspects of the sequential movements (Hatsopoulos et al., 2003; Lu and Ashe, 2005). The effect of extensive practice on the neural and metabolic activities in the M1 of monkeys was examined after 1–6 years of training on a sequential reaching task (Matsuzaka et al., 2007; Picard et al., 2013). In these studies, the monkeys were trained to perform the internally generated sequential reaching task and visually guided reaching task for 1–6 years (**Figure 2**). Then, the neural and metabolic activities were compared between these two conditions to elucidate the characteristics specific to the extensively trained sequential movements. After extensive training on the two tasks (∼2 years), Matsuzaka et al. (2007) recorded the activity of single neurons in the proximal arm representation of M1. In this experiment design, the movements were performed either in the context of an internally generated trained sequence or of a visually guided reaching on the same experiment day (e.g., movement from target 5 to target 3 in **Figures 2B,C**). Therefore, the comparison of activity for movements performed in two different contexts (i.e., internally generated sequence or visually guided reaching) revealed changes of activity associated with training, even though the activity patterns of the neurons before training were unknown (i.e., not recorded). Forty percent of the task-related neurons in M1 were differentially active during the performance of the visually guided and internally generated sequential reaching. The majority of differentially active neurons had enhanced activity for the trained, internally generated sequential reaching (Matsuzaka et al., 2007). Similarly, the uptake of 2DG was examined in the arm area of M1 after extensive training of the sequential reaching task (Picard et al., 2013). Uptake of 2DG is suggested to be associated with presynaptic activity at both excitatory and inhibitory synapses (discussed in Picard and Strick, 2003; Picard et al., 2013). They found that the uptake of 2DG was low in monkeys that performed highly practiced, internally generated sequences of movements compared with the 2DG uptake in monkeys that performed visually guided reaching (Picard et al., 2013). Surprisingly, the low uptake of 2DG was not matched by low neural activity in the same area. Neural activity in arm M1 during the internally generated movements was comparable to that observed during the visually guided movements. Therefore, there was a marked dissociation between the metabolic and neural activities in M1. These observations imply an increase of the synaptic efficacy in M1 after extensive practice, which led to M1's contribution to the planning and generation of sequential movements.

M1 is critical for implementing motor output, so that it has been challenging to test its involvement in the acquisition or the maintenance of motor sequences. Lesions or inactivation of M1 will abolish the motor commands to the spinal cord that generates muscle activity. A few studies reported that, when M1 was inactivated, subjects made more errors in the performance of trained sequential movements compared with that before the inactivation (Lu and Ashe, 2005; Cohen et al., 2009; Censor et al., 2014). However, because M1 is critically involved in motor execution, an advanced approach is required to further understand how M1 contributes to internally generated sequential movements without the confound of basic motor deficits. This was achieved in a recent study by selectively manipulating protein synthesis in the M1 of non-human primates in order to disrupt information storage in this cortical area (**Figure 4**; Ohbayashi, 2020). In the study, the monkeys were trained on two tasks: internally generated sequential movements (repeating task, guided by memory; **Figures 2C**, **4A**) and reaching movements guided by visual cues (random task, visually guided reaching as a control task; **Figure 2B**; Ohbayashi, 2020). After the monkeys practiced each sequence for more than 100 training days and started to perform the memorized sequential movements predictively, the protein synthesis inhibitor anisomycin was injected into the arm representation of M1 to test M1's involvement in the maintenance of sequential movements after extensive practice (**Figure 4B**). Anisomycin injections had a significant effect on the performance of the sequential movements guided by memory during the repeating task. The injections resulted in a significant increase in the number of errors (**Figures 4C,D**) and a significant decrease in the number of predictive responses, an indication of sequence learning, during the repeating task. Moreover, the monkeys made errors reaching in the direction opposite to the correct target (**Figure 4C**, bottom). This type of error suggests a deficit in selecting the movement component in the sequence. In contrast, performance of the visually guided movements during the random task was not significantly disrupted. Interestingly, inactivation of M1 using muscimol injection disrupted the performance on both the random and repeating tasks, suggesting that the inactivation of M1 caused a deficit of motor production (Ohbayashi, 2020). Differences in the effects between anisomycin injection and muscimol injection suggest that the anisomycin injection disrupted the performance of internally generated sequential movements by interfering with the information storage in this area. This observation emphasizes the importance of M1 for the generation of sequential movements guided by memory. The results suggest that, although M1 is critical for movement production, it also is involved in the maintenance of skilled sequential movements (Ohbayashi, 2020).

Protein synthesis inhibitors have been widely used in rodents to study the neural basis of learning and memory. The studies have been conducted in rodents, especially extensively in the context of fear conditioning (Davis and Squire, 1984; Nader et al., 2000a,b; Kandel, 2001; Dudai, 2004, 2012; Dudai and Eisenberg, 2004; Kelleher et al., 2004; Rudy, 2008a,b). *De novo* protein synthesis, during or shortly after the initial training, is shown to be essential in the consolidation of long-term memory (Davis and Squire, 1984). Moreover, when a protein synthesis inhibitor (e.g., anisomycin) was given during the retrieval, the performance of retrieved task was disrupted (Nader et al., 2000a; Nader, 2003; Lee et al., 2008). The studies suggested that the neural trace may be destabilized upon retrieval through protein degradation and then rebounded through protein synthesis

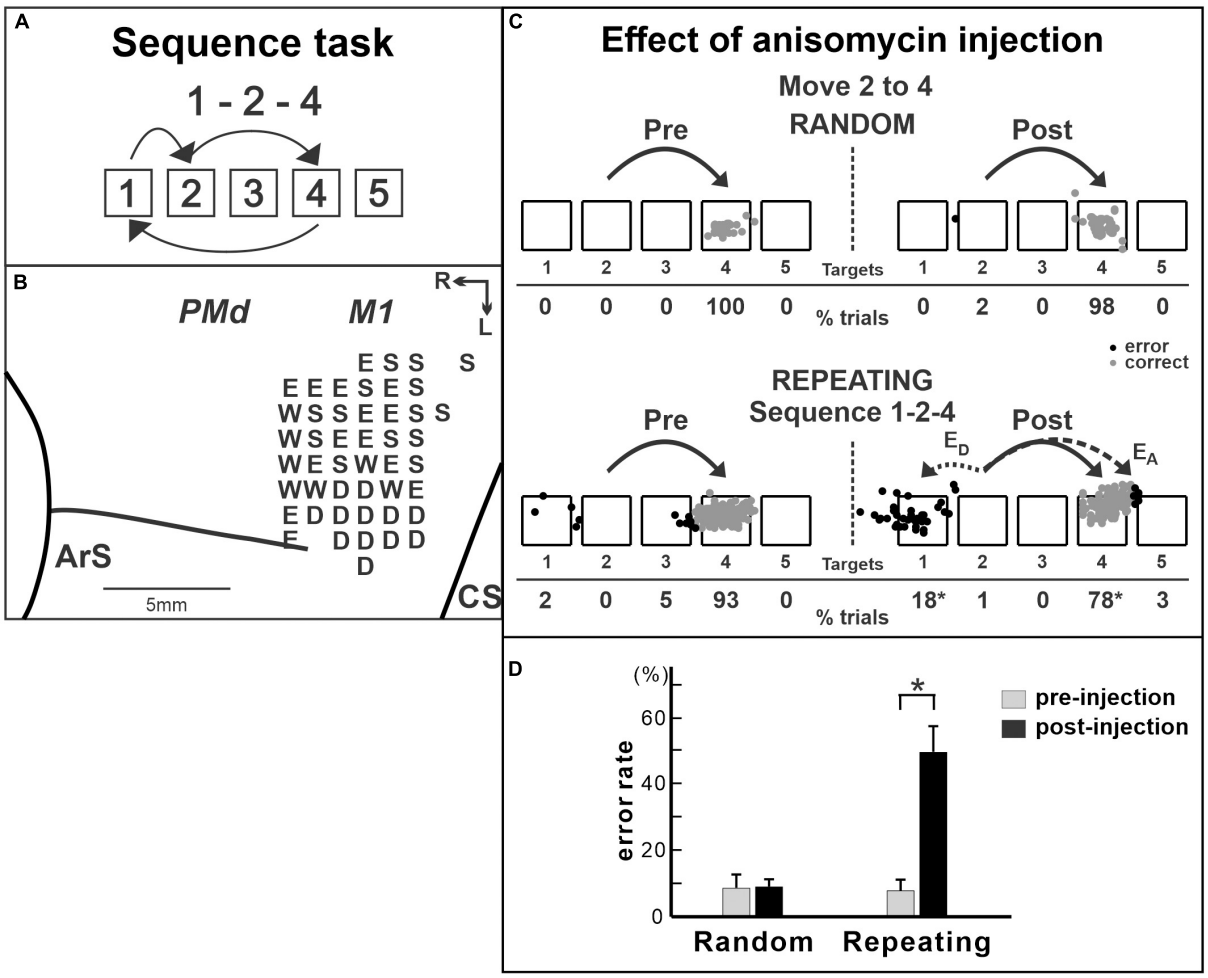

FIGURE 4 | Effect of anisomycin injection in M1 on the performance of internally generated sequential movements. **(A)** Sequence 1-2-4 of the repeating task. **(B)** Intracortical stimulation map of a *Cebus* monkey. *Letters* indicate the movements evoked at each site. *S*, shoulder; *E*, elbow; *W*, wrist; *D*, digit. Injections were done at sites in which intracortical stimulation evoked shoulder or elbow movements (i.e., arm representation). **(C)** Reaching end points of trials from target 2 to target 4 before and after anisomycin injection. *Left*: pre-injection; *right*: post-injection. *Top*: random task; *bottom*: repeating task. The monkey was performing sequence 1-2-4 during the repeating task. E_A: accuracy errors, a reach performed in the correct direction (e.g., to the right), but to an end point outside of the correct target; E_D: direction errors, a reach performed in the direction opposite to the correct target. *Gray dots*: correct response; *black dots*: error response. The percentages of trials ending in each target are given *below the targets*. $^*p < 0.05$. **(D)** Averaged error rates of six injection sessions in the random task (*left*) and the repeating task (*right*). After anisomycin injection, the number of errors increased dramatically in the repeating task, but not in the random task (modified from Ohbayashi, 2020).

during the "reconsolidation" process (Nader et al., 2000a,b; Sara, 2000; Nader, 2003; Lee et al., 2008; Rudy, 2008a,b; Dudai, 2012). Thus, the injection of the protein synthesis inhibitor disrupted the task performance as the inhibitor prevented the synthesis of the proteins needed to reconsolidate the memory trace (Nader et al., 2000a; Lee et al., 2008; Dudai, 2012). The series of studies also proposed that the destabilized trace may be bidirectionally modified to be weakened or strengthened, so that the neural trace can be "updated" (Sara, 2000; Dudai and Eisenberg, 2004; Rudy, 2008b; Dudai, 2012). Although it is unclear whether these proposals can be generalized to other forms of memory, they may inform us of the way by which anisomycin injected in M1 interfered with the performance of the well-practiced sequential movements (discussed in Ohbayashi, 2020). The neural basis for motor skill improvement needs to be further investigated in future studies.

The rodent studies provide valuable insights into the reorganization of the motor cortex during motor skill learning, even though the rodent motor system and the range of motor skills differ from those of human and non-human primates (Dum and Strick, 1991; He et al., 1993, 1995; Rathelot et al., 2016). Early in the learning of the reach–grasp task, the expressions of transcription factors (e.g., an immediate early gene, *c-fos*) increase within the rodent's motor cortex and remain elevated in the plateau phase of the learning curve relative to control animals (Kleim et al., 1996). The increase of gene expression precedes both the changes in synapse number and motor map reorganization (Kleim et al., 1996). The injection of protein

synthesis inhibitors into the motor cortex of rodents disrupted the maintenance (Kleim et al., 2003) or the learning (Luft et al., 2004) of the skilled forelimb reaching. In these studies, the rats were trained to reach and grasp for a food pellet placed outside the cage (Kleim et al., 2003; Luft et al., 2004). The injection of anisomycin into the motor cortex after the training disrupted the performance of the skilled forelimb task and caused reductions in the synapse number and size in the motor cortex *in vivo* (Kleim et al., 2003). The injection of anisomycin into the rodents' motor cortex during the learning disrupted the learning of the motor skill task (Luft et al., 2004). Two photon imaging or electron microscopy studies have shown that skill training leads to the rapid formation of enduring postsynaptic dendritic spines and an increase in synaptic density in neurons in the motor cortex (Kleim et al., 2003; Xu et al., 2009; Yang et al., 2009; Yu and Zuo, 2011). When the newly modified dendritic spines during the training of rotarod tasks were optically manipulated to shrink in the motor cortex, the rodents' performance of the trained task was disrupted (Hayashi-Takagi et al., 2015). The study suggested that the structural plasticity of spines plays a critical role in the learning of motor skills in the motor cortex of rodents.

Taken together, these observations support the view that M1 is involved in skilled sequential movements, especially after extensive practice. The neural activity, metabolic activity, and the structural organization in M1 were influenced by extensive practice on the motor skill tasks. Further studies will expand our understanding of how M1 contributes to the continuous improvement of skilled sequential movements during repetitive practice as well as its contribution to fast learning.

COLLABORATION OF CORTICAL MOTOR AREAS

Studies on anatomical connectivity provided valuable insights into the interaction between multiple areas. The functional distinction of pre-SMA and SMA is supported by the differences in the anatomical connections between these areas (Luppino et al., 1990, 1993; Bates and Goldman-Rakic, 1993, reviewed in Picard and Strick, 2001). Firstly, only the SMA has direct projections to the M1 and the spinal cord (Muakkassa and Strick, 1979; Dum and Strick, 1991, 1996; He et al., 1995; Wang et al., 2001). Secondly, the pre-SMA does not have substantial connections with M1 (Tokuno and Tanji, 1993; Galea and Darian-Smith, 1994; Lu et al., 1994; Hatanaka et al., 2001; Dum and Strick, 2005). Instead, the pre-SMA is densely interconnected with regions of the prefrontal cortex as well as from the rostral cingulate motor area and pre-PMd (F7) (Luppino et al., 1990, 1993, 2003; Bates and Goldman-Rakic, 1993; Lu et al., 1994; Takada et al., 2004; Wang et al., 2005). Moreover, the pre-SMA does not appear to be densely interconnected with the SMA (Luppino et al., 1990, 1993, 2003; Wang et al., 2001). These observations suggest that the SMA is a part of the cortical motor areas and that the pre-SMA can be functionally considered as a region of the prefrontal areas (Bates and Goldman-Rakic, 1993; Luppino et al., 1993; Lu et al., 1994; Picard and Strick, 2001). This view is consistent with the observations of the inactivation studies described above showing that the pre-SMA is involved in cognitive aspects such as the early learning of movement sequences, whereas the SMA is primarily involved in the performance of memorized movement sequences.

The anatomical connections of the M1, PMd, PMv (ventral premotor cortex), and the SMA of monkeys were precisely studied by Dum and Strick's group (Dum and Strick, 2005). The anatomy results showed that the digit representations of the PMd, PMv, and M1 are densely interconnected with each other. Thus, these three cortical areas form a network for the control of hand movements (Dum and Strick, 2005). The projections from the digit representation in the SMA to the PMd and the PMv are stronger than the SMA projections to M1 (Dum and Strick, 2005). This suggests that the SMA may influence through connections with the premotor areas rather than through M1. Overall, the laminar origins of neurons that interconnect the PMd, PMv, and M1 are typical of "lateral" interactions. Dum and Strick commented that "from an anatomical perspective, this cortical network lacks a clear hierarchical organization" (Dum and Strick, 2005). The strong, reciprocal interconnections suggest that these areas may act in concert with each other to produce commands for movements.

In fact, a subset of neurons in each premotor area exhibits activity for relatively simple movements as M1 neurons do (Kurata and Tanji, 1986; Shima et al., 1991; Cadoret and Smith, 1997). Furthermore, in non-human primate studies, aspects of practiced sequences were reflected in the neural or the metabolic activity in all the SMA, PMd, and M1 (Picard and Strick, 1997, 2003; Matsuzaka et al., 2007; Picard et al., 2013; Ohbayashi et al., 2016; Ohbayashi, 2020). On the other hand, the injection of chemical agents in these areas showed that each premotor area is differentially involved in sequential movements. Inactivation of the SMA did not have an effect on the learning and performance of internally generated spatial sequences (Nakamura et al., 1999). Nevertheless, both the muscimol injection in the PMd and the anisomycin injection in M1 selectively disrupted the performance of internally generated sequences, but not the visually guided reaching (Ohbayashi et al., 2016; Ohbayashi, 2020). Moreover, both injections caused deficits in target selection, in which a monkey reached to the direction opposite to the correct target (Ohbayashi et al., 2016; Ohbayashi, 2020). Together with the dense anatomical connection between the PMd and M1, as described above, these suggested a possibility that anisomycin injection disrupted the interaction from the PMd to M1, which resulted in the deficit in the performance of internally generated spatial sequences. These suggest that the PMd functions as a major source of input to M1 to guide the performance of internally generated spatial sequences after practice. More experiments are required to tease out the exact nature of interactions between M1 and the premotor areas in the learning and performance of sequential movements.

SUMMARY

The performance of sequential movements can be improved to the expert level and maintained as a motor skill through

extensive practice. Functional imaging studies in humans show that a brain-wide network subserves the performance of skilled sequential movements. Interventional studies in non-human primates advanced our understanding of its neural basis. The results of interventional studies suggested that each motor area in the network makes a distinct contribution to skilled sequential movements. The SMA is involved in the temporal organization of multiple non-spatial movements into a sequence and the execution of the sequential actions. Its role in spatial sequences is still debatable and needs to be further investigated. The PMd may act as a key structure for the learning of sequential movements by contributing to the selection of appropriate responses. Specifically, the PMd may be critical for the acquisition and maintenance of arbitrary motor–motor associations. In M1, the neural activity, metabolic activity, and structural organization were shown to be modified by extensive practice on sequential movements. M1's involvement in sequential movements after extensive practice was verified by an interventional study using an inhibitor for protein synthesis. These studies suggest that the PMd functions as a major source of input to M1 to guide the performance of internally generated sequences. Together, the PMd and M1 may be parts of the key structures for the learning and maintenance of internally generated sequential movements. The involvements of these areas along the dimensions of time (i.e., learning stages) and sequence category (e.g., spatial and non-spatial) need to be further explored in future experiments.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This work was supported by the National Institutes of Health grant R21NS101499 to MO and the Brain Sciences Project of the CNSI and NINS BS291006 to MO.

ACKNOWLEDGMENTS

I thank Nathalie Picard for discussions, suggestions, and proofreading; Richard Dum and Toshihiro Hayashi for discussions.

REFERENCES

Amunts, K., Schlaug, G., Jancke, L., Steinmetz, H., Schleicher, A., Dabringhaus, A., et al. (1997). Motor cortex and hand motor skills: structural compliance in the human brain. *Hum. Brain Mapp.* 5, 206–215. doi: 10.1002/(SICI)1097-019319975:3<206::AID-HBM5<3.0.CO;2-7

Anderson, B. J., Li, X., Alcantara, A. A., Isaacs, K. R., Black, J. E., and Greenough, W. T. (1994). Glial hypertrophy is associated with synaptogenesis following motor-skill learning, but not with angiogenesis following exercise. *Glia* 11, 73–80. doi: 10.1002/glia.440110110

Averbeck, B. B., Battaglia-Mayer, A., Guglielmo, C., and Caminiti, R. (2009). Statistical analysis of parieto-frontal cognitive-motor networks. *J. Neurophysiol.* 102, 1911–1920. doi: 10.1152/jn.00519.2009

Bassett, D. S., Yang, M., Wymbs, N. F., and Grafton, S. T. (2015). Learning-induced autonomy of sensorimotor systems. *Nat. Neurosci.* 18, 744–751. doi: 10.1038/nn.3993

Bates, J. F., and Goldman-Rakic, P. S. (1993). Prefrontal connections of medial motor areas in the rhesus monkey. *J. Comp. Neurol.* 336, 211–228. doi: 10.1002/cne.903360205

Bengtsson, S. L., Nagy, Z., Skare, S., Forsman, L., Forssberg, H., and Ullen, F. (2005). Extensive piano practicing has regionally specific effects on white matter development. *Nat. Neurosci.* 8, 1148–1150. doi: 10.1038/nn1516

Boecker, H., Dagher, A., Ceballos-Baumann, A. O., Passingham, R. E., Samuel, M., Friston, K. J., et al. (1998). Role of the human rostral supplementary motor area and the basal ganglia in motor sequence control: investigations with H2 15O PET. *J. Neurophysiol.* 79, 1070–1080. doi: 10.1152/jn.1998.79.2.1070

Brinkman, C. (1984). Supplementary motor area of the monkey's cerebral cortex: short- and long-term deficits after unilateral ablation and the effects of subsequent callosal section. *J. Neurosci.* 4, 918–929. doi: 10.1523/JNEUROSCI.04-04-00918.1984

Cadoret, G., and Smith, A. M. (1997). Comparison of the neuronal activity in the SMA and the ventral cingulate cortex during prehension in the monkey. *J. Neurophysiol.* 77, 153–166. doi: 10.1152/jn.1997.77.1.153

Censor, N., Horovitz, S. G., and Cohen, L. G. (2014). Interference with existing memories alters offline intrinsic functional brain connectivity. *Neuron* 81, 69–76. doi: 10.1016/j.neuron.2013.10.042

Cisek, P., and Kalaska, J. F. (2002). Simultaneous encoding of multiple potential reach directions in dorsal premotor cortex. *J. Neurophysiol.* 87, 1149–1154. doi: 10.1152/jn.00443.2001

Cisek, P., and Kalaska, J. F. (2004). Neural correlates of mental rehearsal in dorsal premotor cortex. *Nature* 431, 993–996. doi: 10.1038/nature03005

Cisek, P., and Kalaska, J. F. (2005). Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action. *Neuron* 45, 801–814. doi: 10.1016/j.neuron.2005.01.027

Clower, W. T., and Alexander, G. E. (1998). Movement sequence-related activity reflecting numerical order of components in supplementary and presupplementary motor areas. *J. Neurophysiol.* 80, 1562–1566. doi: 10.1152/jn.1998.80.3.1562

Cohen, N. R., Cross, E. S., Wymbs, N. F., and Grafton, S. T. (2009). Transient disruption of M1 during response planning impairs subsequent offline consolidation. *Exp. Brain Res.* 196, 303–309. doi: 10.1007/s00221-009-1838-x

Cona, G., and Semenza, C. (2017). Supplementary motor area as key structure for domain-general sequence processing: a unified account. *Neurosci. Biobehav. Rev.* 72, 28–42. doi: 10.1016/j.neubiorev.2016.10.033

Coynel, D., Marrelec, G., Perlbarg, V., Pelegrini-Issac, M., van de Moortele, P. F., Ugurbil, K., et al. (2010). Dynamics of motor-related functional integration during motor sequence learning. *Neuroimage* 49, 759–766. doi: 10.1016/j.neuroimage.2009.08.048

Davis, H. P., and Squire, L. R. (1984). Protein synthesis and memory: a review. *Psychol. Bull.* 96, 518–559. doi: 10.1037/0033-2909.96.3.518

Dayan, E., and Cohen, L. G. (2011). Neuroplasticity subserving motor skill learning. *Neuron* 72, 443–454. doi: 10.1016/j.neuron.2011.10.008

di Pellegrino, G., and Wise, S. P. (1993). Effects of attention on visuomotor activity in the premotor and prefrontal cortex of a primate. *Somatosens. Mot. Res.* 10, 245–262. doi: 10.3109/08990229309028835

Dick, J. P., Benecke, R., Rothwell, J. C., Day, B. L., and Marsden, C. D. (1986). Simple and complex movements in a patient with infarction of the right supplementary motor area. *Mov. Disord.* 1, 255–266. doi: 10.1002/mds.870010405

Doyon, J., and Benali, H. (2005). Reorganization and plasticity in the adult brain during learning of motor skills. *Curr. Opin. Neurobiol.* 15, 161–167. doi: 10.1016/j.conb.2005.03.004

Doyon, J., Penhune, V., and Ungerleider, L. G. (2003). Distinct contribution of the cortico-striatal and cortico-cerebellar systems to motor skill learning. *Neuropsychologia* 41, 252–262. doi: 10.1016/s0028-3932(02)00158-6

Draganski, B., and May, A. (2008). Training-induced structural changes in the adult human brain. *Behav. Brain Res.* 192, 137–142. doi: 10.1016/j.bbr.2008.02.015

Dudai, Y. (2004). The neurobiology of consolidations, or, how stable is the engram? *Annu. Rev. Psychol.* 955, 51–86. doi: 10.1146/annurev.psych.55.090902.142050

Dudai, Y. (2012). The restless engram: consolidations never end. *Annu. Rev. Neurosci.* 35, 227–247. doi: 10.1146/annurev-neuro-062111-150500

Dudai, Y., and Eisenberg, M. (2004). Rites of passage of the engram: reconsolidation and the lingering consolidation hypothesis. *Neuron* 44, 93–100. doi: 10.1016/j.neuron.2004.09.003

Dum, R. P., and Strick, P. L. (1991). The origin of corticospinal projections from the premotor areas in the frontal lobe. *J. Neurosci.* 11, 667–689. doi: 10.1523/JNEUROSCI.11-03-00667.1991

Dum, R. P., and Strick, P. L. (1996). Spinal cord terminations of the medial wall motor areas in macaque monkeys. *J. Neurosci.* 16, 6513–6525. doi: 10.1523/JNEUROSCI.16-20-06513.1996

Dum, R. P., and Strick, P. L. (2005). Frontal lobe inputs to the digit representations of the motor areas on the lateral surface of the hemisphere. *J. Neurosci.* 25, 1375–1386. doi: 10.1523/JNEUROSCI.3902-04.2005

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., and Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science* 270, 305–307. doi: 10.1126/science.270.5234.305

Evarts, E. V. (1981). "Chapter 23: role of motor cortex in voluntary movements in primates," in *Handbook of Physiology, The Nervous System, Motor Control*, Vol. 2, ed. V. B. Brook (Baltimore, MD: Williams and Wilkins).

Fitts, P. M., and Posner, M. I. (1967). *Human Performance*. Belmont, CA: Brooks/Cole Publishers.

Floyer-Lea, A., and Matthews, P. M. (2004). Changing brain networks for visuomotor control with increased movement automaticity. *J. Neurophysiol.* 92, 2405–2412. doi: 10.1152/jn.01092.2003

Galea, M. P., and Darian-Smith, I. (1994). Multiple corticospinal neuron populations in the macaque monkey are specified by their unique cortical origins, spinal terminations, and connections. *Cereb. Cortex* 4, 166–194. doi: 10.1093/cercor/4.2.166

Gaser, C., and Schlaug, G. (2003a). Brain structures differ between musicians and non-musicians. *J. Neurosci.* 23, 9240–9245. doi: 10.1523/JNEUROSCI.23-27-09240.2003

Gaser, C., and Schlaug, G. (2003b). Gray matter differences between musicians and non-musicians. *Ann. N. Y. Acad. Sci.* 999, 514–517. doi: 10.1196/annals.1284.062

Gerloff, C., Corwell, B., Chen, R., Hallett, M., and Cohen, L. G. (1997). Stimulation over the human supplementary motor area interferes with the organization of future elements in complex motor sequences. *Brain* 120(Pt 9), 1587–1602. doi: 10.1093/brain/120.9.1587

Goldberg, G. (1985). Supplementary motor area structure and function – review and hypotheses. *Behav. Brain Sci.* 8, 567–588. doi: 10.1017/S0140525x00045167

Grafton, S. T., Hazeltine, E., and Ivry, R. (1995). Functional mapping of sequence learning in normal humans. *J. Cogn. Neurosci.* 7, 497–510. doi: 10.1162/jocn.1995.7.4.497

Grafton, S. T., Woods, R. P., and Tyszka, M. (1994). Functional imaging of procedural motor learning: relating cerebral blood flow with individual subject performance. *Hum. Brain Mapp.* 1, 221–234. doi: 10.1002/hbm.460010307

Halsband, U. (1987). "Higher disturbances of movement in monkeys (*Macaca mulatta*)," in *Motor Control*, eds G. N. Gantchev, B. Dimitrov, and P. Gatev (New York, NY: Plenum Press), 79–85. doi: 10.1007/978-1-4615-7508-5_14

Halsband, U., Ito, N., Tanji, J., and Freund, H. J. (1993). The role of premotor cortex and the supplementary motor area in the temporal control of movement in man. *Brain* 116(Pt 1), 243–266. doi: 10.1093/brain/116.1.243

Halsband, U., and Passingham, R. (1982). The role of premotor and parietal cortex in the direction of action. *Brain Res.* 240, 368–372. doi: 10.1016/0006-8993(82)90239-6

Han, Y., Yang, H., Lv, Y. T., Zhu, C. Z., He, Y., Tang, H. H., et al. (2009). Gray matter density and white matter integrity in pianists' brain: a combined structural and diffusion tensor MRI study. *Neurosci. Lett.* 459, 3–6. doi: 10.1016/j.neulet.2008.07.056

Hardwick, R. M., Rottschy, C., Miall, R. C., and Eickhoff, S. B. (2013). A quantitative meta-analysis and review of motor learning in the human brain. *Neuroimage* 67, 283–297. doi: 10.1016/j.neuroimage.2012.11.020

Haslinger, B., Erhard, P., Altenmuller, E., Hennenlotter, A., Schwaiger, M., Grafin Von Einsiedel, H., et al. (2004). Reduced recruitment of motor association areas during bimanual coordination in concert pianists. *Hum. Brain Mapp.* 22, 206–215. doi: 10.1002/hbm.20028

Hatanaka, N., Nambu, A., Yamashita, A., Takada, M., and Tokuno, H. (2001). Somatotopic arrangement and corticocortical inputs of the hindlimb region of the primary motor cortex in the macaque monkey. *Neurosci. Res.* 40, 9–22. doi: 10.1016/s0168-0102(01)00210-3

Hatsopoulos, N. G., Paninski, L., and Donoghue, J. P. (2003). Sequential movement representations based on correlated neuronal activity. *Exp. Brain Res.* 149, 478–486. doi: 10.1007/s00221-003-1385-9

Hayashi-Takagi, A., Yagishita, S., Nakamura, M., Shirai, F., Wu, Y. I., Loshbaugh, A. L., et al. (2015). Labelling and optical erasure of synaptic memory traces in the motor cortex. *Nature* 525, 333–338. doi: 10.1038/nature15257

He, S. Q., Dum, R. P., and Strick, P. L. (1993). Topographic organization of corticospinal projections from the frontal lobe: motor areas on the lateral surface of the hemisphere. *J. Neurosci.* 13, 952–980. doi: 10.1523/JNEUROSCI.13-03-00952.1993

He, S. Q., Dum, R. P., and Strick, P. L. (1995). Topographic organization of corticospinal projections from the frontal lobe: motor areas on the medial surface of the hemisphere. *J. Neurosci.* 15, 3284–3306. doi: 10.1523/JNEUROSCI.15-05-03284.1995

Herholz, S. C., and Zatorre, R. J. (2012). Musical training as a framework for brain plasticity: behavior, function, and structure. *Neuron* 76, 486–502. doi: 10.1016/j.neuron.2012.10.011

Hikosaka, O., Nakamura, K., Sakai, K., and Nakahara, H. (2002). Central mechanisms of motor skill learning. *Curr. Opin. Neurobiol.* 12, 217–222. doi: 10.1016/s0959-4388(02)00307-0

Hikosaka, O., Sakai, K., Miyauchi, S., Takino, R., Sasaki, Y., and Putz, B. (1996). Activation of human presupplementary motor area in learning of sequential procedures: a functional MRI study. *J. Neurophysiol.* 76, 617–621. doi: 10.1152/jn.1996.76.1.617

Hoshi, E., and Tanji, J. (2007). Distinctions between dorsal and ventral premotor areas: anatomical connectivity and functional properties. *Curr. Opin. Neurobiol.* 17, 234–242. doi: 10.1016/j.conb.2007.02.003

Hund-Georgiadis, M., and von Cramon, D. Y. (1999). Motor-learning-related changes in piano players and non-musicians revealed by functional magnetic-resonance signals. *Exp. Brain Res.* 125, 417–425. doi: 10.1007/s002210050698

Jancke, L., Shah, N. J., and Peters, M. (2000). Cortical activations in primary and secondary motor areas for complex bimanual movements in professional pianists. *Brain Res. Cogn. Brain Res.* 10, 177–183. doi: 10.1016/s0926-6410(00)00028-8

Johnson, P. B., Ferraina, S., Bianchi, L., and Caminiti, R. (1996). Cortical networks for visual reaching: physiological and anatomical organization of frontal and parietal lobe arm regions. *Cereb. Cortex* 6, 102–119. doi: 10.1093/cercor/6.2.102

Kalaska, J. F., and Crammond, D. J. (1995). Deciding not to GO: neuronal correlates of response selection in a GO/NOGO task in primate premotor and parietal cortex. *Cereb. Cortex* 5, 410–428. doi: 10.1093/cercor/5.5.410

Kandel, E. R. (2001). The molecular biology of memory storage: a dialogue between genes and synapses. *Science* 294, 1030–1038. doi: 10.1126/science.1067020

Karni, A., Meyer, G., Jezzard, P., Adams, M. M., Turner, R., and Ungerleider, L. G. (1995). Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature* 377, 155–158. doi: 10.1038/377155a0

Karni, A., Meyer, G., Rey-Hipolito, C., Jezzard, P., Adams, M. M., Turner, R., et al. (1998). The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex. *Proc. Natl. Acad. Sci. U.S.A.* 95, 861–868. doi: 10.1073/pnas.95.3.861

Kelleher, R. J. III, Govindarajan, A., and Tonegawa, S. (2004). Translational regulatory mechanisms in persistent forms of synaptic plasticity. *Neuron* 44, 59–73. doi: 10.1016/j.neuron.2004.09.013

Kleim, J. A., Bruneau, R., Calder, K., Pocock, D., Vandenberg, P. M., Macdonald, E., et al. (2003). Functional organization of adult motor cortex is dependent upon continued protein synthesis. *Neuron* 40, 167–176. doi: 10.1016/s0896-6273(03)00592-0

Kleim, J. A., Lussnig, E., Schwarz, E. R., Comery, T. A., and Greenough, W. T. (1996). Synaptogenesis and Fos expression in the motor cortex of the adult rat after motor skill learning. *J. Neurosci.* 16, 4529–4535. doi: 10.1523/JNEUROSCI.16-14-04529.1996

Krings, T., Topper, R., Foltys, H., Erberich, S., Sparing, R., Willmes, K., et al. (2000). Cortical activation patterns during complex motor tasks in piano players and control subjects. A functional magnetic resonance imaging study. *Neurosci. Lett.* 278, 189–193. doi: 10.1016/s0304-3940(99)00930-1

Kurata, K., and Hoffman, D. S. (1994). Differential effects of muscimol microinjection into dorsal and ventral aspects of the premotor cortex of monkeys. *J. Neurophysiol.* 71, 1151–1164. doi: 10.1152/jn.1994.71.3.1151

Kurata, K., and Tanji, J. (1986). Premotor cortex neurons in macaques: activity before distal and proximal forelimb movements. *J. Neurosci.* 6, 403–411. doi: 10.1523/JNEUROSCI.06-02-00403.1986

Kurata, K., and Wise, S. P. (1988). Premotor cortex of rhesus monkeys: set-related activity during two conditional motor tasks. *Exp. Brain Res.* 69, 327–343. doi: 10.1007/BF00247578

Lee, D., and Quessy, S. (2003). Activity in the supplementary motor area related to learning and performance during a sequential visuomotor task. *J. Neurophysiol.* 89, 1039–1056. doi: 10.1152/jn.00638.2002

Lee, S. H., Choi, J. H., Lee, N., Lee, H. R., Kim, J. I., Yu, N. K., et al. (2008). Synaptic protein degradation underlies destabilization of retrieved fear memory. *Science* 319, 1253–1256. doi: 10.1126/science.1150541

Lu, M. T., Preston, J. B., and Strick, P. L. (1994). Interconnections between the prefrontal cortex and the premotor areas in the frontal lobe. *J. Comp. Neurol.* 341, 375–392. doi: 10.1002/cne.903410308

Lu, X., and Ashe, J. (2005). Anticipatory activity in primary motor cortex codes memorized movement sequences. *Neuron* 45, 967–973. doi: 10.1016/j.neuron.2005.01.036

Luft, A. R., Buitrago, M. M., Ringer, T., Dichgans, J., and Schulz, J. B. (2004). Motor skill learning depends on protein synthesis in motor cortex after training. *J. Neurosci.* 24, 6515–6520. doi: 10.1523/JNEUROSCI.1034-04.2004

Luppino, G., Matelli, M., Camarda, R., and Rizzolatti, G. (1993). Corticocortical connections of area F3 (SMA-proper) and area F6 (pre-SMA) in the macaque monkey. *J. Comp. Neurol.* 338, 114–140. doi: 10.1002/cne.903380109

Luppino, G., Matelli, M., and Rizzolatti, G. (1990). Cortico-cortical connections of two electrophysiologically identified arm representations in the mesial agranular frontal cortex. *Exp. Brain Res.* 82, 214–218. doi: 10.1007/BF00230855

Luppino, G., Rozzi, S., Calzavara, R., and Matelli, M. (2003). Prefrontal and agranular cingulate projections to the dorsal premotor areas F2 and F7 in the macaque monkey. *Eur. J. Neurosci.* 17, 559–578. doi: 10.1046/j.1460-9568.2003.02476.x

Matsuzaka, Y., Picard, N., and Strick, P. L. (2007). Skill representation in the primary motor cortex after long-term practice. *J. Neurophysiol.* 97, 1819–1832. doi: 10.1152/jn.00784.2006

Meister, I., Krings, T., Foltys, H., Boroojerdi, B., Muller, M., Topper, R., et al. (2005). Effects of long-term practice and task complexity in musicians and nonmusicians performing simple and complex motor tasks: implications for cortical motor organization. *Hum. Brain Mapp.* 25, 345–352. doi: 10.1002/hbm.20112

Mitz, A. R., Godschalk, M., and Wise, S. P. (1991). Learning-dependent neuronal activity in the premotor cortex: activity during the acquisition of conditional motor associations. *J. Neurosci.* 11, 1855–1872. doi: 10.1523/JNEUROSCI.11-06-01855.1991

Miyachi, S., Hikosaka, O., and Lu, X. (2002). Differential activation of monkey striatal neurons in the early and late stages of procedural learning. *Exp. Brain Res.* 146, 122–126. doi: 10.1007/s00221-002-1213-7

Muakkassa, K. F., and Strick, P. L. (1979). Frontal lobe inputs to primate motor cortex: evidence for four somatotopically organized 'premotor' areas. *Brain Res.* 177, 176–182. doi: 10.1016/0006-8993(79)90928-4

Mushiake, H., Inase, M., and Tanji, J. (1991). Neuronal activity in the primate premotor, supplementary, and precentral motor cortex during visually guided and internally determined sequential movements. *J. Neurophysiol.* 66, 705–718. doi: 10.1152/jn.1991.66.3.705

Mushiake, H., and Strick, P. L. (1993). Preferential activity of dentate neurons during limb movements guided by vision. *J. Neurophysiol.* 70, 2660–2664. doi: 10.1152/jn.1993.70.6.2660

Mushiake, H., and Strick, P. L. (1995). Pallidal neuron activity during sequential arm movements. *J. Neurophysiol.* 74, 2754–2758. doi: 10.1152/jn.1995.74.6.2754

Nachev, P., Kennard, C., and Husain, M. (2008). Functional role of the supplementary and pre-supplementary motor areas. *Nat. Rev. Neurosci.* 9, 856–869. doi: 10.1038/nrn2478

Nader, K. (2003). Memory traces unbound. *Trends Neurosci.* 26, 65–72. doi: 10.1016/S0166-2236(02)00042-5

Nader, K., Schafe, G. E., and Le Doux, J. E. (2000a). Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature* 406, 722–726. doi: 10.1038/35021052

Nader, K., Schafe, G. E., and Le Doux, J. E. (2000b). The labile nature of consolidation theory. *Nat. Rev. Neurosci.* 1, 216–219. doi: 10.1038/35044580

Nakamura, K., Sakai, K., and Hikosaka, O. (1998). Neuronal activity in medial frontal cortex during learning of sequential procedures. *J. Neurophysiol.* 80, 2671–2687. doi: 10.1152/jn.1998.80.5.2671

Nakamura, K., Sakai, K., and Hikosaka, O. (1999). Effects of local inactivation of monkey medial frontal cortex in learning of sequential procedures. *J. Neurophysiol.* 82, 1063–1068. doi: 10.1152/jn.1999.82.2.1063

Ohbayashi, M. (2020). Inhibition of protein synthesis in M1 of monkeys disrupts performance of sequential movements guided by memory. *Elife* 9:e53038. doi: 10.7554/eLife.53038

Ohbayashi, M., and Picard, N. (2020). Sequential reaching task for the study of motor skills in monkeys. *Bio Protoc.* 10:e3719. doi: 10.21769/BioProtoc.3719

Ohbayashi, M., Picard, N., and Strick, P. L. (2016). Inactivation of the dorsal premotor area disrupts internally generated, but not visually guided sequential movements. *J. Neurosci.* 36, 1971–1976. doi: 10.1523/JNEUROSCI.2356-15.2016

Orban, P., Peigneux, P., Lungu, O., Albouy, G., Breton, E., Laberenne, F., et al. (2010). The multifaceted nature of the relationship between performance and brain activity in motor sequence learning. *Neuroimage* 49, 694–702. doi: 10.1016/j.neuroimage.2009.08.055

Passingham, R. E. (1988). Premotor cortex and preparation for movement. *Exp. Brain Res.* 70, 590–596. doi: 10.1007/BF00247607

Pesaran, B., Nelson, M. J., and Andersen, R. A. (2008). Free choice activates a decision circuit between frontal and parietal cortex. *Nature* 453, 406–409. doi: 10.1038/nature06849

Picard, N., Matsuzaka, Y., and Strick, P. L. (2013). Extended practice of a motor skill is associated with reduced metabolic activity in M1. *Nat. Neurosci.* 16, 1340–1347. doi: 10.1038/nn.3477

Picard, N., and Strick, P. L. (1997). Activation on the medial wall during remembered sequences of reaching movements in monkeys. *J. Neurophysiol.* 77, 2197–2201. doi: 10.1152/jn.1997.77.4.2197

Picard, N., and Strick, P. L. (2001). Imaging the premotor areas. *Curr. Opin. Neurobiol.* 11, 663–672. doi: 10.1016/s0959-4388(01)00266-5

Picard, N., and Strick, P. L. (2003). Activation of the supplementary motor area (SMA) during performance of visually guided movements. *Cereb. Cortex* 13, 977–986. doi: 10.1093/cercor/13.9.977

Rathelot, J. A., Nwankwo, A., and Strick, P. L. (2016). *Origin of Descending Commands from the Cerebral Cortex to Hand Motoneurons in the Rat. Abstract Retrieved from Abstracts in Society of Neuroscience Database Accession No. 534.01*. San Diego, CA: Society for Neuroscience. Available online at: https://www.abstractsonline.com/pp8/index.html#!/4071/

Reddy, P. G., Mattar, M. G., Murphy, A. C., Wymbs, N. F., Grafton, S. T., Satterthwaite, T. D., et al. (2018). Brain state flexibility accompanies motor-skill acquisition. *Neuroimage* 171, 135–147. doi: 10.1016/j.neuroimage.2017.12.093

Roland, P. E., Larsen, B., Lassen, N. A., and Skinhoj, E. (1980). Supplementary motor area and other cortical areas in organization of voluntary movements in man. *J. Neurophysiol.* 43, 118–136. doi: 10.1152/jn.1980.43.1.118

Rosenbaum, D. A. (2010). *Human Motor Control*. Amsterdam: Elsevier Inc. Publishers.

Rudy, J. W. (2008a). Destroying memories to strengthen them. *Nat. Neurosci.* 11, 1241–1242. doi: 10.1038/nn1108-1241

Rudy, J. W. (2008b). *The Neurobiology of Learning and Memory*. Sunderland, MA: Sinauer Associates, Inc. Publishers.

Sakai, K., Hikosaka, O., Miyauchi, S., Takino, R., Sasaki, Y., and Putz, B. (1998). Transition of brain activation from frontal to parietal areas in visuomotor

sequence learning. *J. Neurosci.* 18, 1827–1840. doi: 10.1523/jneurosci.18-05-01827.1998

Sampaio-Baptista, C., and Johansen-Berg, H. (2017). White matter plasticity in the adult brain. *Neuron* 96, 1239–1251. doi: 10.1016/j.neuron.2017.11.026

Sara, S. J. (2000). Strengthening the shaky trace through retrieval. *Nat. Rev. Neurosci.* 1, 212–213. doi: 10.1038/35044575

Schmidt, R. A., and Lee, T. D. (2011). *Motor Control and Learning: A Behavioral Emphasis*. Champaign, IL: Human Kinetics. Publishers.

Schwenkreis, P., El Tom, S., Ragert, P., Pleger, B., Tegenthoff, M., and Dinse, H. R. (2007). Assessment of sensorimotor cortical representation asymmetries and motor skills in violin players. *Eur. J. Neurosci.* 26, 3291–3302. doi: 10.1111/j.1460-9568.2007.05894.x

Shibasaki, H., Sadato, N., Lyshkow, H., Yonekura, Y., Honda, M., Nagamine, T., et al. (1993). Both primary motor cortex and supplementary motor area play an important role in complex finger movement. *Brain* 116(Pt 6), 1387–1398. doi: 10.1093/brain/116.6.1387

Shima, K., Aya, K., Mushiake, H., Inase, M., Aizawa, H., and Tanji, J. (1991). Two movement-related foci in the primate cingulate cortex observed in signal-triggered and self-paced forelimb movements. *J. Neurophysiol.* 65, 188–202. doi: 10.1152/jn.1991.65.2.188

Shima, K., Mushiake, H., Saito, N., and Tanji, J. (1996). Role for cells in the presupplementary motor area in updating motor plans. *Proc. Natl. Acad. Sci. U.S.A.* 93, 8694–8698. doi: 10.1073/pnas.93.16.8694

Shima, K., and Tanji, J. (1998). Both supplementary and presupplementary motor areas are crucial for the temporal organization of multiple movements. *J. Neurophysiol.* 80, 3247–3260. doi: 10.1152/jn.1998.80.6.3247

Shima, K., and Tanji, J. (2000). Neuronal activity in the supplementary and presupplementary motor areas for temporal organization of multiple movements. *J. Neurophysiol.* 84, 2148–2160. doi: 10.1152/jn.2000.84.4.2148

Shima, K., and Tanji, J. (2006). Binary-coded monitoring of a behavioral sequence by cells in the pre-supplementary motor area. *J. Neurosci.* 26, 2579–2582. doi: 10.1523/JNEUROSCI.4161-05.2006

Takada, M., Nambu, A., Hatanaka, N., Tachibana, Y., Miyachi, S., Taira, M., et al. (2004). Organization of prefrontal outflow toward frontal motor-related areas in macaque monkeys. *Eur. J. Neurosci.* 19, 3328–3342. doi: 10.1111/j.0953-816X.2004.03425.x

Tanji, J. (2001). Sequential organization of multiple movements: involvement of cortical motor areas. *Annu. Rev. Neurosci.* 24, 631–651. doi: 10.1146/annurev.neuro.24.1.631

Tanji, J., and Shima, K. (1994). Role for supplementary motor area cells in planning several movements ahead. *Nature* 371, 413–416. doi: 10.1038/371413a0

Tanji, J., and Shima, K. (1996a). Contrast of neuronal activity between the supplemental motor area and other cortical motor areas. *Adv. Neurol.* 70, 95–103.

Tanji, J., and Shima, K. (1996b). Supplementary motor cortex in organization of movement. *Eur. Neurol.* 36(Suppl. 1), 13–19. doi: 10.1159/000118878

Tanji, J., Shima, K., and Mushiake, H. (1996). Multiple cortical motor areas and temporal sequencing of movements. *Brain Res. Cogn. Brain Res.* 5, 117–122. doi: 10.1016/s0926-6410(96)00047-x

Tokuno, H., and Tanji, J. (1993). Input organization of distal and proximal forelimb areas in the monkey primary motor cortex: a retrograde double labeling study. *J. Comp. Neurol.* 333, 199–209. doi: 10.1002/cne.903330206

Ungerleider, L. G., Doyon, J., and Karni, A. (2002). Imaging brain plasticity during motor skill learning. *Neurobiol. Learn. Mem.* 78, 553–564. doi: 10.1006/nlme.2002.4091

Wang, Y., Isoda, M., Matsuzaka, Y., Shima, K., and Tanji, J. (2005). Prefrontal cortical cells projecting to the supplementary eye field and presupplementary motor area in the monkey. *Neurosci. Res.* 53, 1–7. doi: 10.1016/j.neures.2005.05.005

Wang, Y., Shima, K., Sawamura, H., and Tanji, J. (2001). Spatial distribution of cingulate cells projecting to the primary, supplementary, and pre-supplementary motor areas: a retrograde multiple labeling study in the macaque monkey. *Neurosci. Res.* 39, 39–49. doi: 10.1016/s0168-0102(00)00198-x

Wenger, E., Brozzoli, C., Lindenberger, U., and Lovden, M. (2017). Expansion and renormalization of human brain structure during skill acquisition. *Trends Cogn. Sci.* 21, 930–939. doi: 10.1016/j.tics.2017.09.008

Wiener, M., Turkeltaub, P., and Coslett, H. B. (2010). The image of time: a voxel-wise meta-analysis. *Neuroimage* 49, 1728–1740. doi: 10.1016/j.neuroimage.2009.09.064

Wise, S. P. (1985). The primate premotor cortex: past, present, and preparatory. *Annu. Rev. Neurosci.* 8, 1–19. doi: 10.1146/annurev.ne.08.030185.000245

Wise, S. P., Boussaoud, D., Johnson, P. B., and Caminiti, R. (1997). Premotor and parietal cortex: corticocortical connectivity and combinatorial computations. *Annu. Rev. Neurosci.* 20, 25–42. doi: 10.1146/annurev.neuro.20.1.25

Wymbs, N. F., and Grafton, S. T. (2013). Contributions from the left PMd and the SMA during sequence retrieval as determined by depth of training. *Exp. Brain Res.* 224, 49–58. doi: 10.1007/s00221-012-3287-1

Xu, T., Yu, X., Perlik, A. J., Tobin, W. F., Zweig, J. A., Tennant, K., et al. (2009). Rapid formation and selective stabilization of synapses for enduring motor memories. *Nature* 462, 915–919. doi: 10.1038/nature08389

Yang, G., Pan, F., and Gan, W. B. (2009). Stably maintained dendritic spines are associated with lifelong memories. *Nature* 462, 920–924. doi: 10.1038/nature08577

Yokoi, A., and Diedrichsen, J. (2019). Neural organization of hierarchical motor sequence representations in the human neocortex. *Neuron* 103, 1178–1190.e7. doi: 10.1016/j.neuron.2019.06.017

Yu, X., and Zuo, Y. (2011). Spine plasticity in the motor cortex. *Curr. Opin. Neurobiol.* 21, 169–174. doi: 10.1016/j.conb.2010.07.010

Zatorre, R. J., Fields, R. D., and Johansen-Berg, H. (2012). Plasticity in gray and white: neuroimaging changes in brain structure during learning. *Nat. Neurosci.* 15, 528–536. doi: 10.1038/nn.3045

Conflict of Interest: The author declares 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 in Human Neuroscience

ORIGINAL RESEARCH
published: 18 June 2021
doi: 10.3389/fnhum.2021.663789

Transcranial Magnetic Stimulation Over the Right Posterior Superior Temporal Sulcus Promotes the Feature Discrimination Processing

Qihui Zhou[1], *Penghui Song*[1], *Xueming Wang*[1], *Hua Lin*[1] *and Yuping Wang*[1,2,3]*

[1] *Department of Neurology, Xuanwu Hospital, Capital Medical University, Beijing, China,* [2] *Collaborative Innovation Center for Brain Disorders, Institute of Sleep and Consciousness Disorders, Beijing Institute of Brain Disorders, Capital Medical University, Beijing, China,* [3] *Beijing Key Laboratory of Neuromodulation, Beijing Municipal Science and Technology Commission, Beijing, China*

Attention is the dynamic process of allocating limited resources to the information that is most relevant to our goals. Accumulating studies have demonstrated the crucial role of frontal and parietal areas in attention. However, the effect of posterior superior temporal sulcus (pSTS) in attention is still unclear. To address this question, in this study, we measured transcranial magnetic stimulation (TMS)-induced event-related potentials (ERPs) to determine the extent of involvement of the right pSTS in attentional processing. We hypothesized that TMS would enhance the activation of the right pSTS during feature discrimination processing. We recruited 21 healthy subjects who performed the dual-feature delayed matching task while undergoing single-pulse sham or real TMS to the right pSTS 300 ms before the second stimulus onset. The results showed that the response time was reduced by real TMS of the pSTS as compared to sham stimulation. N270 amplitude was reduced during conflict processing, and the time-varying network analysis revealed increased connectivity between the frontal lobe and temporo-parietal and occipital regions. Thus, single-pulse TMS of the right pSTS enhances feature discrimination processing and task performance by reducing N270 amplitude and increasing connections between the frontal pole and temporo-parietal and occipital regions. These findings provide evidence that the right pSTS facilitates feature discrimination by accelerating the formation of a dynamic network.

Keywords: transcranial magnetic stimulation, posterior superior temporal sulcus, dual-feature matching task, event-related potentials, electroencephalography

OPEN ACCESS

Edited by:
Yuji Naya,
Peking University, China

Reviewed by:
Adam Steel,
Dartmouth College, United States
Masaki Takeda,
Kochi University of Technology, Japan

***Correspondence:**
Yuping Wang
wangyuping01@sina.cn

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

Received: *03 February 2021*
Accepted: *17 May 2021*
Published: *18 June 2021*

Citation:
Zhou Q, Song P, Wang X, Lin H and Wang Y (2021) Transcranial Magnetic Stimulation Over the Right Posterior Superior Temporal Sulcus Promotes the Feature Discrimination Processing. Front. Hum. Neurosci. 15:663789. doi: 10.3389/fnhum.2021.663789

INTRODUCTION

Attention is the key cognitive function that selects currently relevant pieces of information at the expense of irrelevant ones, thereby facilitating the selection of features (Moore and Zirnsak, 2017). Classical studies of attention have identified that different attentional processing methods are carried out by two distinct attention systems: a dorsal system comprising the intraparietal sulcus and frontal eye fields, which are involved in top-down attention control, and a ventral feature-based system comprising the temporo-parietal junction for unexpected stimuli (Katsuki and Constantinidis, 2014; Majerus et al., 2018).

It is widely accepted that frontal and parietal cortices are the main brain areas related to attention control (Moore and Zirnsak, 2017), but recent evidence suggests that temporal cortex is also involved in the modulation of motion and color discrimination (Bogadhi et al., 2018; Stemmann and Freiwald, 2019) and is a critical structure in the cortical control of covert selective attention (Bogadhi et al., 2019). The posterior superior temporal sulcus (pSTS), as an essential part of temporal cortex, is thought to be involved in multiple neural processes (Hein and Knight, 2008; Beauchamp, 2011). Thus, clarifying the role of the pSTS in feature discrimination can provide a novel insight into the attention control system.

Transcranial magnetic stimulation (TMS) is a non-invasive brain stimulation method used to alter brain activity (Barker and Jalinous, 1985; Hallett, 2007). The TMS-induced spread of synchronized neural activity in the target area to connected brain regions can be observed by functional brain recording techniques (Siebner et al., 2009). Electroencephalography (EEG) has been increasingly employed in recent years to measure the cortical responses evoked by focal TMS at high temporal resolution (Bergmann et al., 2016). Combining TMS with EEG (TMS–EEG) enables the measurement of brain-wide cortical responses to TMS (Ilmoniemi and Kicic, 2010) and has permitted studies to examine the brain states and the dynamics across the cortical areas with excellent temporal resolution (Pellicciari et al., 2017). In addition, TMS–EEG provides precise information on the spatiotemporal order of activation of cortical areas, which can reveal causal interactions within functional brain networks (Rogasch and Fitzgerald, 2013).

Event-related potentials (ERPs) reflect brain functioning at a high temporal resolution (Luck, 2014). P100 (P1) and N150 (N1) are associated with early visual perceptual processing (Taylor, 2002; Luck, 2014), while P300 (P3) is related to information recognition and inhibition processing (Polich, 1998). N270 is a negative ERP component with a peak latency of around 270 ms that is elicited by conflicting modalities [e.g., visual and auditory (Wang et al., 2002), color and shape (Wang et al., 2004) and position mismatch (Yang and Wang, 2002)]. N270 amplitude is modulated by attention (Mao and Wang, 2008; Zhang et al., 2013). Thus, N270 is generally regarded as an electrophysiological marker of attention (Wang et al., 2001; Scannella et al., 2016). The dual-feature delayed matching task is associated with conflict processing (Wang et al., 2003; Zhang et al., 2003). Previous studies have revealed that conflict processing is modulated by attention (Carter et al., 2000), so we could use the task to explore the attentional modulation of pSTS to the conflict processing.

In this study, we used a single-pulse TMS–EEG procedure combined with a dual-feature delayed matching task to explore the role of the right pSTS in feature discrimination. We hypothesized that TMS applied over the right pSTS would modulate the brain dynamic network and facilitate the processing of the color-shape feature discrimination.

MATERIALS AND METHODS

Participants

We recruited 21 healthy right-handed subjects [11 females; mean age (±SD): 24.2 ± 1.9 years] for the study. All subjects had a normal or corrected-to-normal vision with no color blindness and no history of any major diseases or neurological or mental disorders. Written informed consent was obtained from all participants, and the experiment was approved by the Ethics Committee of the Xuanwu Hospital, Capital Medical University.

Experimental Procedures

All of the participants underwent a MRI scan and TMS. T1-weighted anatomic MRI was performed using a 3T MRI scanner (Siemens Medical Solutions, Erlangen, Germany) before the TMS to identify the target site for stimulation. Each participant completed two TMS sessions on 2 different days (for sham and real TMS sessions). Each TMS session comprised of four blocks, namely, two color and two shape task blocks; and each block consisted of 80 trials and lasted approximately 8 min.

During the two TMS sessions, participants sat on a comfortable chair in a dimly lit and quiet room with their gaze fixated on the computer monitor, completing both color and shape tasks in each TMS session. The order of tasks and TMS sessions was counterbalanced across subjects.

The resting motor threshold (RMT) of each participant was measured before the start of each session with an EEG cap. The RMT was defined as the lowest stimulus intensity that induced at least five motor evoked potentials of 50 μV over ten consecutive pulses in the first dorsal interosseous muscle.

Experimental Task

Each participant performed a dual-feature delayed matching task (Wang et al., 2004; **Figure 1**). A stimulus presentation system (STIM; Neurosoft Labs, Charlotte, NC, the United States) was used to present each target in the task, which was defined by the combination of a specific color (red, green, yellow, or white) and shape (pentagon, ellipse, triangle, hexagon, octagon, arrow, cross, circular, pentagram, or rhombus).

Each trial began with a central fixation point presented for 2,000 ms. The first (S1) and second (S2) stimuli were presented for 500 ms each with an interstimulus interval of 500 ms. A single sham or real TMS pulse was delivered 300 ms before the onset of S2 in each trial. The interval between the end of the previous S2 and the onset of the next S1 was 5 s. Stimuli were presented in the center of the screen on a black background. There were four types of stimulus pair: C–S– [same color, same shape (match)]; C + S– [different color, same shape (task-relevant mismatch in color task)]; C–S + [same color, different shape (task-irrelevant mismatch in color task)]; and C + S+ [different color, different shape (conjunction mismatch)]. The four types of stimulus pairs were randomly presented in sequence, and each type had the same probability.

The color task (selective attention to color, Ac) required participants to judge whether the color of S2 was identical to that of S1 while ignoring their shape. The shape task (selective

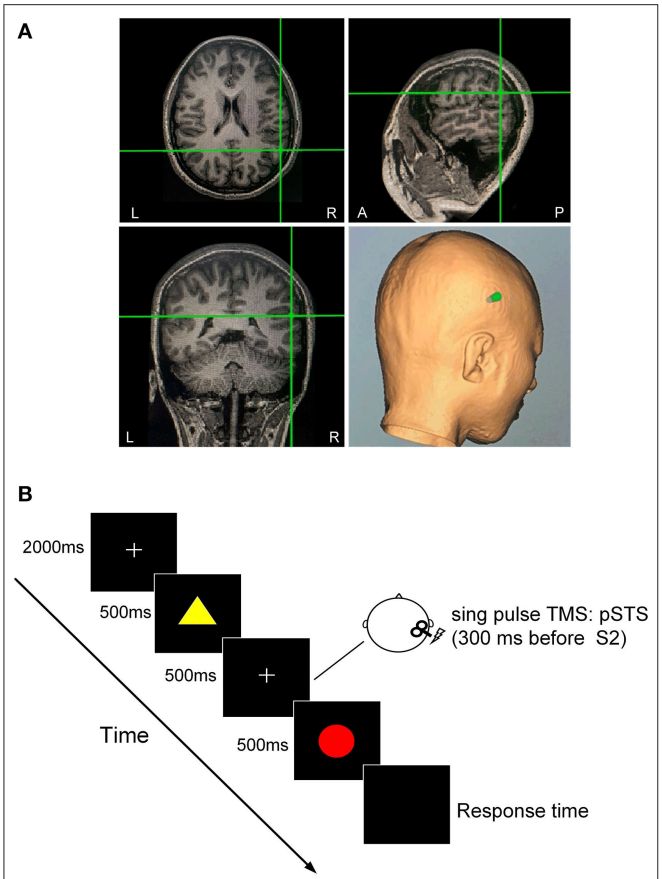

FIGURE 1 | Dual-feature delayed matching task. **(A)** TMS-targeted areas of the right pSTS in one subject. The TMS coil orientation was determined based on the anatomical MRI scan. **(B)** For each trial, two presentations of the first (S1) and second (S2) stimuli were separated by 500 ms, during which a sham or real TMS pulse was applied over the right pSTS 300 ms before the onset of S2 presentation. At the end of the trial, the subject pressed the left or the right button of a mouse to judge whether the color of S2 was identical to that of S1while ignoring the shape of the stimulus in the color task, or whether the shape of S2 was identical to that of S1 while ignoring the color of the stimulus in the shape task.

attention to shape, As) required participants to discriminate whether S1 and S2 had identical shapes while ignoring their color. Upon the presentation of S2, participants were required to press the left or right button of a mouse with their left or right thumb to indicate their response as quickly and accurately as possible. The response hand was counterbalanced in each subject.

TMS Site Localization

The TMS target sites in the right pSTS were individually identified using T1-weighted anatomic MRI image of the participants with Brainsight frameless stereotactic neuronavigation system (SofTaxic Navigator System, EMS Italy). For each participant, the right pSTS area was localized using anatomical features: the inflection point in the pSTS where it angles upward toward the parietal lobe (Beauchamp et al., 2008). We manually selected the inflection as the location of pSTS on the structural MRI image of an individual. A near-IR navigation system was used to track the position of the stimulation coil to identify the right pSTS in the head of the participant.

TMS–EEG

Single-pulse TMS was delivered to the right pSTS using a monophasic Magstim stimulator (Magstim Company Ltd., London, the United Kingdom) with a figure-of-eight coil (outer winding diameter of 70 mm). The stimulation intensity was applied at 90% of the RMT. In the sham TMS condition, the intensity was the same but the surface of the coil was perpendicular to the head of the subject.

Continuous EEG recording from 21 scalp electrodes in an elastic cap that were positioned according to the international 10–20 system (Greentek Ltd, Wuhan, China) was performed with a magnetic field-compatible EEG amplifier (Yunshen, Beijing, China). The apex nasi was used as a reference. Skin impedance was maintained below 5 kΩ. The signal was digitized at a sampling rate of 1,024 Hz. Earplugs with white noise were used at all times by the participants to attenuate the sound of TMS.

Data Analysis

Behavioral Data Analysis

Behavioral data were extracted from the perceptual reports. Trials with incorrect responses or a reaction time (RT) $<$200 or $>$1,500 ms were excluded from the analysis. The results from the two tasks were averaged across the 21 participants. The correct rate and mean RT data were subjected to repeated-measures analysis of variance (ANOVA) with condition (real/sham), task (color/shape), and type (match/task-relevant/task-irrelevant/conjunction mismatch) as within-subject factors.

ERP Data Analysis

The EEG data were analyzed with MATLAB vR2015b (MathWorks, Natick, MA, the United States) using customized scripts and EEGLAB Toolbox (Delorme and Makeig, 2004). The preprocessing consisted of two rounds of independent component analysis (ICA) (Hyvärinen and Oja, 2000): the first to remove the components containing large-amplitude TMS-induced artifacts and the second to remove any remaining artifacts (e.g., blinking). TMS-induced artifacts were removed by discarding the signal for 60 ms after each pulse. The EEG epochs were extracted using a time window of 2,000 ms (1,000 ms before and 1,000 ms after TMS) and baseline-corrected to the 200 ms before the first visual stimulus onset.

The amplitude of P1, N1, N270, and P3 was measured from the averaged waveforms at notable electrodes. By the visual inspection of the ERPs, the mean amplitudes of P1 (60–110 ms), N1 (112–200 ms), N270 (220–320 ms), and P3 (322–500 ms) in corresponding notable electrodes were analyzed. Electrophysiological parameters were analyzed by means of a four-way ANOVA with conditions and task, type, and electrode sites as factors. The Greenhouse–Geisser epsilon correction for non-sphericity was applied where appropriate. *Post-hoc*

paired *t*-tests were passed through Bonferroni correction for multiple comparisons.

Time-Varying Network Analysis

EEG data analysis was divided into preprocessing and time-varying network analysis. The latter required several segmentations to enable the construction of a reliable network that captured the brain architectures and networks. In this study, we used TMS disturbances as stimulus labels. For each labeled disturbance event, the time point corresponding to the peak of the label was set as time 0; data corresponding to 0.5 s before and 1 s after time 0 were extracted (total segment length = 1.5 s).

To reduce the calculation load in the time-varying network analysis, eight times downsampling was applied to obtain a sampling frequency of 32 Hz. The time-varying network was calculated by the adaptive directed transfer function (ADTF) method and a time-varying multivariate adaptive autoregressive (tv-MVAAR) model (Zhang et al., 2017) to observe dynamic information processing during TMS disturbance. The two-sample *t*-test was used to compare time-varying networks under the sham and real TMS conditions, and false discovery rate correction was applied for multiple comparisons. $P < 0.05$ were regarded as statistically significant.

Time-Varying Multivariate Adaptive Autoregressive Model

For each artifact-free segment, the tv-MVAAR model is defined as

$$X(t) = \sum_{i=1}^{p} A(i,t)\, X(t-1) + E(t) \tag{1}$$

where *X(t)* represents the data vector of EEG signal and *E(t)* represents the multivariate independent white noise. *A(i,t)* represents the matrix of tv-MVAAR model coefficients, which is estimated by the Kalman filter algorithm. *p* represents the order of the model that is automatically determined by the Akaike Information Criterion (AIC) within the range of 2–20 as,

$$AIC(p) = In\left[\det(\chi)\right] + 2M^2 p/N \tag{2}$$

where *M* is the number of the electrodes, *p* is the optimal order of the model, *N* represents the number of the time points of each time series, and χ represents the corresponding covariance matrix.

Adaptive Directed Transfer Function

Parameters *A(f,t)* and *H(f,t)* in the frequency domain are defined as follows:

$$A(f,t) = \sum_{k=0}^{p} A_k(t) e^{-j2\pi ftk} \tag{3}$$

$$A(f,t)\, X(f,t) = E(f,t) \tag{4}$$

$$X(f,t) = A^{-1}(f,t)\, E(f,t) = H(f,t)E(f,t) \tag{5}$$

where A_k is the matrix of the tv-MVAAR model coefficients, and *X(f,t)* and *E(f,t)* are the Fourier transformations of *X(t)* and *E(t)* in the frequency domain, respectively.

Moreover, the normalized ADTF describing the directed flow from the *j*th to the *i*th node is defined in Equation (6), and the final integrated ADTF is defined in Equation (7) within the frequency band of interest (i.e., 0.5–14.5 Hz in this work) as follows:

$$\gamma_{ij}^2(f,t) = \frac{\left|H_{ij}(f,t)\right|^2}{\sum_{m=1}^{n}\left|H_{im}(f,t)\right|^2} \tag{6}$$

$$Q_{ij}^2(t) = \frac{\sum_{k=f_1}^{f_2}\gamma_{ij}^2(k,t)}{f_2 - f_1} \tag{7}$$

The normalized total information outflow of the *j*th node is further estimated in Equation (8) as:

$$Q_j^2(t) = \frac{\sum_{k=1}^{n} Q_{kj}^2(t)}{n-1}, \quad for\ k \neq j \tag{8}$$

where *n* is the total number of nodes. When each node (*n*) has been calculated for each sample time point (*t*), a directional edge (*i* to *j*) can be displayed. From Equation 8, we can derive an outflow that denotes the time-varying of each node across different time points.

RESULTS

Behavior Data

The correct rate and mean RTs in the dual-feature delayed matching task are shown in **Table 1**. The 2 × 2 × 4 ANOVA of RTs revealed the main effects of condition [$F_{(1,20)} = 26.87$, $P < 0.001$] and stimulus type [$F_{(3,60)} = 15.47$, $P < 0.001$]. A *post-hoc* analysis showed that RT was significantly shorter in the real TMS condition than in the sham TMS condition with all types ($Ps < 0.05$). In both conditions, RTs were significantly longer for task-relevant and conjunction mismatch types than for the match type in each task ($P < 0.05$). The overall correct response rates for each type and task did not differ between the two conditions ($P > 0.05$).

ERP data

Grand-Averaged ERP Waveforms

Grand-averaged ERP waveforms difference between the real and sham conditions for the four types in each task are shown in **Figures 2**, **3**. The ERPs with the four types consisted of P1 and N1 components at the posterior scalp and the P3 component over the whole scalp. The amplitude of P1 in the real TMS condition was increased than that in the sham TMS condition, while the amplitude of N270 was reduced in mismatch types. Visual inspection of the grand-averaged data showed mean amplitudes measured at 60–110 ms for P1, 112–200 ms for N1, 220–320 ms for N270, and 322–500 ms for P3. The components were similar in the sham and real TMS conditions.

P1 Component

In the time window between 60 and 110 ms, after the onset of S2, the condition showed a main effect on the amplitude [$F_{(1,20)} = 10.32$, $P = 0.004$] as well as interaction with electrode area [$F_{(3,60)} = 10.08$, $P < 0.001$). The pairwise comparisons indicated that the amplitude at electrode P4 was more positive in the real TMS condition than in the sham condition ($P < 0.05$; **Table 2**).

TABLE 1 | Mean reaction time and correct rate in two conditions for each task and type.

	Type†	Sham TMS		Real TMS	
		Ac	As	Ac	As
Reaction time (ms)	Match	573.8 ± 72.8	555.6 ± 90.2	527.4 ± 81.8*	524.4 ± 85.5*
	Task-relevant	616.4 ± 66.7[a]	638.9 ± 64.2[a]	572.9 ± 84.6*[a]	602.8 ± 77.7*[a]
	Task-irrelevant	559.3 ± 71.1[b]	574.5 ± 81.0[b]	533.8 ± 87.5*	547.4 ± 77.6*[b]
	Conjunction	623.6 ± 62.5[ac]	614.7 ± 75.7[ac]	578.2 ± 88.9*[a]	599.3 ± 85.8[ac]
Correct rate (%)	Match	98.3 ± 2.1	99.3 ± 1.8	99.0 ± 1.7	99.2 ± 1.4
	Task-relevant	99.2 ± 1.4	98.6 ± 2.0	99.0 ± 1.8	98.6 ± 2.0
	Task-irrelevant	99.0 ± 1.5	99.1 ± 1.6	99.1 ± 1.6	98.1 ± 2.0
	Conjunction	98.4 ± 2.0	99.2 ± 1.4	99.3 ± 1.1	98.3 ± 2.6

Data represent mean ± SD.

**Significantly different from sham condition in the same task and type ($P < 0.05$).*

[a]Significantly different from match type in the same condition and task ($P < 0.05$).

[b]Significantly different from the relevant type in the same condition and task ($P < 0.05$).

[c]Significantly different from the irrelevant type in the same condition and task ($P < 0.05$).

†Match, Ac/C–S– or As/C–S–; task-relevant, Ac/C + S– or As/C–S +; task-irrelevant, Ac/C–S+ or As/C+S–; conjunction, Ac/C + S + or As/C + S+; C–S–, color and shape are the same in stimulus pair; C–S+, the same color with different shape; C + S–, different color with the same shape; C+S+, color and shape both are different.

Ac, attention to color; As, attention to shape; TMS, transcranial magnetic stimulation.

N1 Component

In the time window between 112 and 200 ms after the onset of S2, the stimulus type showed main effects on the amplitude [$F_{(3,60)} = 4.63$, $P = 0.006$]. Type also showed an interaction with electrode area [$F_{(9,180)} = 3.78$, $P < 0.001$]. The *post-hoc* analysis showed that the conjunction type was more positive than match and irrelevant types at electrode P4 ($P < 0.05$). However, no significant difference of the amplitude between the real and sham TMS conditions was found [$F_{(1,20)} = 3.25$, $P = 0.087$; **Table 2**).

N270 Component

In the time window between 220 and 320 ms, ANOVA revealed the significant main effects of condition [$F_{(1,20)} = 20.40$, $P < 0.001$] and type [$F_{(1.35,27.05)} = 11.02$, $P = 0.001$] on the amplitude. Type showed an interaction with electrode [$F_{(5.27,105.30)} = 2.92$, $P = 0.015$]. The *post-hoc* analysis showed that the amplitude of the mismatch types was more negative than that of the match type at electrodes C3 and C4 ($P < 0.05$).

Condition also showed an interaction with type [$F_{(1.50,29.96)} = 4.85$, $P = 0.023$]. The pairwise comparisons revealed that amplitude was significantly more positive in the real TMS condition than in the sham TMS condition with all three mismatch types in each task (all *Ps* < 0.05). However, in the match type, there was no significant difference in amplitude between the sham and real TMS conditions (**Table 2**).

P3 Component

In the time window between 322 and 500 ms, the mean amplitude of P300 did not differ between the sham and real TMS conditions [$F_{(1,20)} = 0.16$, $P = 0.693$]. The TMS condition did not interact with task [$F_{(1,20)} = 0.24$, $P = 0.629$], type [$F_{(3,60)} = 0.74$, $P = 0.533$], or electrode [$F_{(1.66,33.25)} = 1.41$, $P = 0.257$; **Table 2**].

Topography

Viewing the topography of the two conditions, it showed that the amplitude of P1 at different scalp areas in the real TMS condition was more positive than that in the sham TMS condition in both tasks. The N1, N270, and P3 distributions have showed no distinct differences between the sham and real TMS conditions in both tasks (**Figures 2**, **3**).

Time-Varying Network

Changes in the time-varying network differed significantly between sham and real TMS conditions (**Figure 4**). TMS of the right pSTS induced changes in the time-varying networks of different types. The increased and decreased connections by real TMS appeared at the timing of TMS and are maintained throughout S2 presentation. Here, we focused on the conflict processing stage, which was around 270 ms after the presentation of S2.

In the color task, the connections between the bilateral frontal poles and the temporo-occipital region were significantly increased with each type compared to the sham TMS condition (**Figure 4A**). Meanwhile, the connections between the bilateral frontal regions and the temporo-parietal region, central regions, were decreased in each type (**Figure 4B**).

In the shape task, the connections between the bilateral frontal poles and the temporo-occipital region, parietal regions, were increased in each type (**Figure 4C**). Meanwhile, compared to the

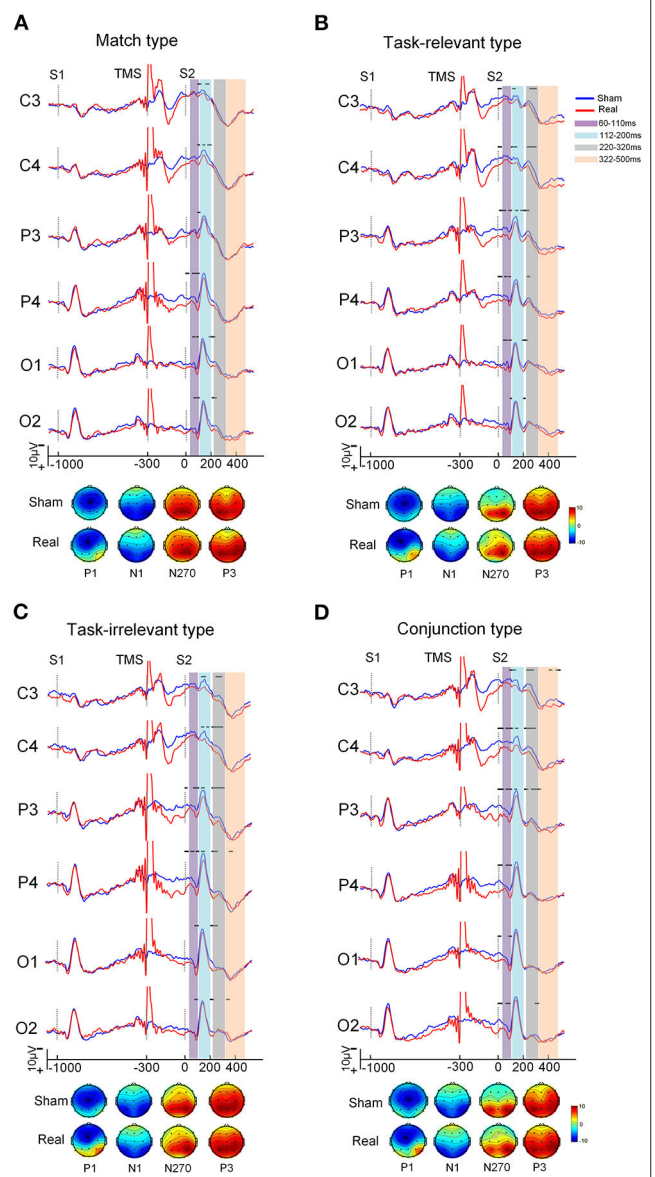

FIGURE 2 | The grand-averaged ERPs and topographies of each type and condition in color task. The ERPs showed difference between the sham and real TMS condition [upper of **(A–D)**]. The amplitude of P1 in real TMS condition was increased than that in sham TMS condition, while the amplitude of N270 was reduced in mismatch types. The topographies [bottom of **(A–D)**] showed that P1 at posterior scalp areas in real TMS condition was more positive than sham one. Vertical lines indicated the TMS, S1, and S2 stimulus onset. Horizontal solid line indicates the time after the onset of S2 with statistical significance between the two conditions. The purple-shaded areas indicate the time window (60–110 ms) in which the P100 was measured, the blue-shaded areas indicate the time window (112–200 ms) in which the N150 was measured, the gray-shaded areas indicate the time window (220–320 ms) in which the N270 was measured, and the orange-shaded areas indicate the time window (322–500 ms) in which the P3000 was measured.

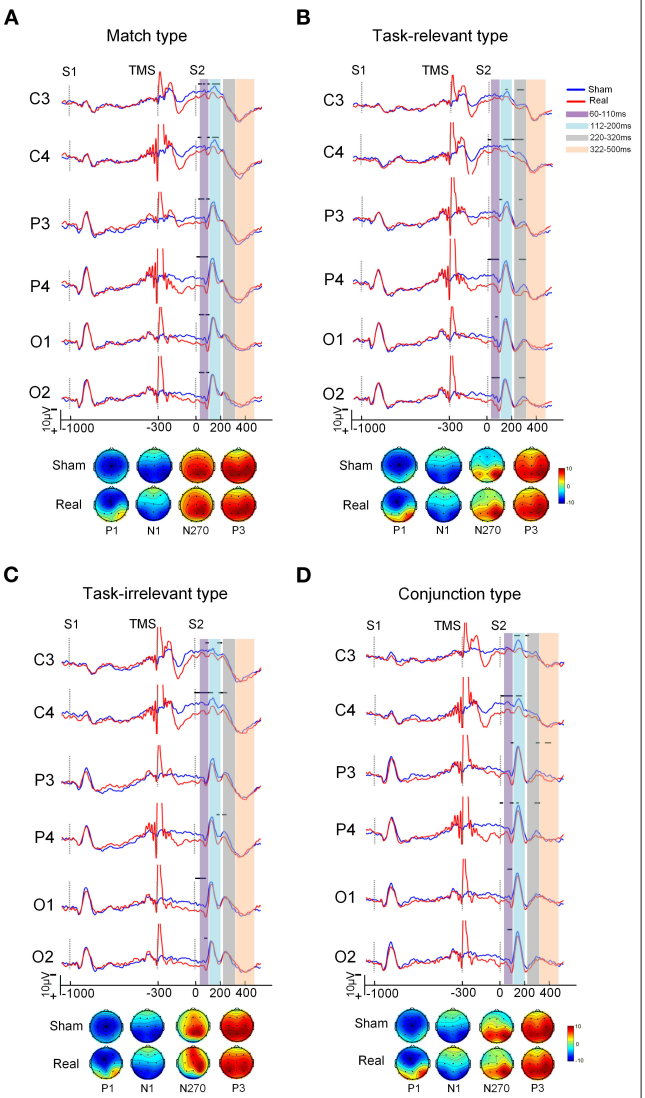

FIGURE 3 | The grand-averaged ERPs and topographies of each type and condition in shape task. The ERPs showed difference between the sham and real TMS condition [upper of **(A–D)**]. The amplitude of P1 in real TMS condition was increased than that in sham TMS condition, while the amplitude of N270 was reduced in mismatch types. The topographies [bottom of **(A–D)**] showed that P1 at posterior scalp areas in real TMS condition was more positive than in sham one. Vertical lines indicated the TMS, S1, and S2 stimulus onset. Horizontal solid line indicates the time after the onset of S2 with statistical significance between the two conditions. The purple-shaded areas indicate the time window (60–110 ms) in which the P100 was measured, the blue-shaded areas indicate the time window (112–200 ms) in which the N150 was measured, the gray-shaded areas indicate the time window (220–320 ms) in which the N270 was measured, and the orange-shaded areas indicate the time window (322–500 ms) in which the P3000 was measured.

sham TMS condition, the information outflows in the frontal and midline region were significantly reduced in the real TMS condition (**Figure 4D**).

DISCUSSION

The results of this study demonstrate that, in addition to the frontal and parietal cortices, the pSTS makes a significant

TABLE 2 | Mean amplitude with all stimulus types for each task and condition in different time windows.

Time window	Condition	Area	Ac				As			
			Match	Task-relevant	Task-irrelevant	Conjunction	Match	Task-relevant	Task-irrelevant	Conjunction
60–110 ms	Sham TMS	P3	−3.5 ± 7.7	−4.2 ± 6.8	−3.0 ± 7.4	−2.5 ± 5.5	−4.2 ± 6.0	−3.6 ± 8.4	−5.2 ± 5.5	−3.2 ± 8.7
		P4	−3.4 ± 8.0	−3.3 ± 7.1	−2.3 ± 8.2	−2.3 ± 6.6	−4.3 ± 6.8	−2.5 ± 5.5	−5.1 ± 7.0	−3.4 ± 7.6
		O1	−0.6 ± 8.5	−1.2 ± 7.0	−0.9 ± 9.3	−0.3 ± 7.6	−2.0 ± 8.3	−0.4 ± 6.2	−2.5 ± 8.2	−1.5 ± 9.8
		O2	−1.3 ± 9.1	−0.9 ± 6.9	−1.1 ± 9.4	−1.0 ± 8.1	−5.2 ± 7.0	−0.8 ± 7.0	−3.5 ± 9.1	−1.8 ± 9.6
	Real TMS	P3	−2.8 ± 6.5	−0.6 ± 7.7*	−1.4 ± 6.5	−1.1 ± 6.7	−2.0 ± 4.4	−0.9 ± 5.7	−2.9 ± 5.1	−0.1 ± 4.2*
		P4	−0.1 ± 7.5*	1.9 ± 8.0*	2.1 ± 8.0*	2.3 ± 6.5*	1.2 ± 6.3*	3.0 ± 6.6*	1.9 ± 6.7*	2.8 ± 5.6*
		O1	0.2 ± 8.2	1.2 ± 7.9	2.8 ± 9.0*	1.0 ± 8.8	0.3 ± 7.1	3.5 ± 7.3*	0.1 ± 7.7	1.0 ± 8.3
		O2	0.7 ± 8.3	0.8 ± 6.6	1.0 ± 8.1	0.9 ± 8.1	0.4 ± 8.4	3.9 ± 7.9*	1.9 ± 5.7*	1.4 ± 7.4
112–200 ms	Sham TMS	P3	−7.4 ± 6.4	−6..2 ± 5.1	−7.6 ± 6.6	−5.4 ± 7.2	−8.1 ± 5.8	−7.0 ± 7.8	−6.9 ± 5.1	−6.4 ± 7.4
		P4	−8.5 ± 6.4	−6.4 ± 7.8	−8.8 ± 6.8	−5.4 ± 6.5^{ac}	−9.2 ± 6.8	−7.8 ± 7.0	−8.0 ± 6.5	−5.1 ± 8.8^{abc}
		O1	−10.7 ± 7.8	−9.5 ± 7.1	−10.9 ± 7.8	−8.1 ± 6.3^{ac}	−11.2 ± 7.2	−9.7 ± 6.9	−9.8 ± 6.7	−10.6 ± 8.4
		O2	−10.4 ± 8.7	−9.1 ± 9.8	−11.2 ± 9.1	−9.1 ± 9.5	−10.6 ± 5.1	−10.6 ± 7.7	−9.5 ± 9.3	−10.6 ± 9.9
	Real TMS	P3	−6.8 ± 5.7	−5.1 ± 5.3	−5.6 ± 5.9	−4.4 ± 5.2^{a}	−6.0 ± 5.0	−5.1 ± 4.7	−4.8 ± 6.5	−6.5 ± 5.9
		P4	−6.8 ± 8.0	−5.6 ± 5.3	−6.2 ± 8.4	−3.2 ± 6.8^{ac}	−8.3 ± 8.7	−4.7 ± 6.0	−8.4 ± 4.5^{b}	−5.1 ± 7.7^{ac}
		O1	−10.1 ± 8.0	−8.5 ± 6.7	−9.9 ± 7.1	−8.6 ± 6.9	−9.4 ± 7.5	−8.6 ± 5.9	−8.1 ± 6.5	−8.6 ± 5.9
		O2	−8.9 ± 9.5	−7.2 ± 7.9	−9.3 ± 8.9	−8.3 ± 8.1	−7.7 ± 7.6	−8.8 ± 7.2	−8.0 ± 8.4	−8.8 ± 6.4
220–320 ms	Sham TMS	C3	6.1 ± 7.8	−1.6 ± 3.9^{a}	1.0 ± 4.7^{ab}	−1.0 ± 5.4^{a}	6.2 ± 7.6	0.5 ± 5.0^{a}	1.9 ± 5.1^{a}	1.0 ± 5.5^{a}
		C4	6.7 ± 7.8	0.1 ± 6.1^{a}	2.0 ± 6.3^{a}	0.5 ± 4.4^{a}	6.2 ± 8.0	0.5 ± 5.6^{a}	2.0 ± 5.4^{a}	1.5 ± 4.1^{a}
		P3	8.2 ± 9.7	0.4 ± 4.2^{a}	4.7 ± 5.0^{b}	3.3 ± 5.6^{ab}	7.0 ± 9.4	2.9 ± 5.4	2.5 ± 6.8^{a}	3.8 ± 4.2
		P4	8.1 ± 9.3	2.1 ± 5.0^{a}	6.1 ± 5.3^{b}	3.2 ± 5.8^{c}	8.0 ± 9.8	3.9 ± 8.5^{a}	2.8 ± 7.2^{a}	4.9 ± 5.2
	Real TMS	C3	6.5 ± 6.2	1.5 ± 5.1*^{a}	4.0 ± 5.8*^{ab}	2.4 ± 5.6*^{a}	6.0 ± 6.4	3.1 ± 4.8*^{a}	3.8 ± 5.4*^{a}	3.6 ± 5.1*^{a}
		C4	8.5 ± 5.9	5.7 ± 5.7*^{a}	5.4 ± 7.2*^{a}	3.5 ± 5.2*^{a}	7.5 ± 6.7	4.5 ± 5.8*^{a}	5.0 ± 5.1*^{a}	4.3 ± 5.2*^{a}
		P3	8.7 ± 7.0	3.7 ± 5.3*^{a}	7.1 ± 3.0*^{b}	5.8 ± 5.2*^{b}	6.6 ± 7.5	6.1 ± 5.3^{a}	3.7 ± 6.9*	6.6 ± 5.0*
		P4	9.9 ± 7.1	5.2 ± 5.4*^{a}	9.1 ± 4.8*^{b}	6.0 ± 5.7*^{ac}	8.4 ± 7.9	7.5 ± 7.0*^{a}	5.6 ± 6.0*	7.4 ± 5.0*
322–500 ms	Sham TMS	C3	−7.4 ± 4.9	7.8 ± 3.4	8.6 ± 4.4	8.1 ± 3.5	9.1 ± 5.5	8.6 ± 5.6	9.9 ± 4.9	7.7 ± 4.7
		C4	9.7 ± 4.6	8.9 ± 4.1	9.2 ± 3.6	8.7 ± 4.8	9.7 ± 5.3	9.7 ± 5.9	9.9 ± 5.2	8.7 ± 5.1
		P3	9.7 ± 5.4	9.7 ± 4.2	10.4 ± 5.4	9.5 ± 3.9	9.9 ± 5.4	10.0 ± 7.0	10.2 ± 5.6	8.7 ± 6.1
		P4	8.4 ± 6.3	9.3 ± 5.9	10.0 ± 4.8	8.9 ± 5.3	9.5 ± 6.3	10.7 ± 8.0	9.8 ± 5.1	8.3 ± 6.7
	Real TMS	C3	7.1 ± 4.4	8.3 ± 6.1	9.6 ± 5.0	7.4 ± 5.0	8.5 ± 3.9	7.2 ± 4.1	10.3 ± 3.9	6.8 ± 4.8
		C4	9.0 ± 4.6	10.2 ± 6.1	10.2 ± 3.9	9.5 ± 5.7	9.8 ± 4.0	9.6 ± 5.3	10.9 ± 4.0	8.9 ± 5.5
		P3	9.7 ± 5.0	10.7 ± 5.7	10.6 ± 6.0	9.6 ± 5.2	9.6 ± 4.9	8.9 ± 4.6	12.0 ± 5.6	8.4 ± 5.3
		P4	9.4 ± 5.0	10.6 ± 5.2	9.9 ± 5.3	9.4 ± 5.5	9.3 ± 5.2	9.9 ± 6.0	11.2 ± 5.6	8.8 ± 6.1

Data represent mean ± SD.

**Significantly different from sham condition in the same type and areas of each task ($P < 0.05$).*

[a] Significantly different from match type in the same condition and areas of each task ($P < 0.05$).

[b] Significantly different from the relevant type in the same condition and areas of each task ($P < 0.05$).

[c] Significantly different from the irrelevant type in the same condition and areas of each task ($P < 0.05$).

Ac, attention to color; As, attention to shape; TMS, transcranial magnetic stimulation.

contribution to attentional processing in humans. Furthermore, we found that single-pulse TMS of the pSTS facilitates feature discrimination, possibly by promoting the formation of a dynamic attention network.

Behavior Performance

In this experiment, the better performance in RTs was not at the cost of correct response in the real TMS condition. Previous studies have reported that behavioral changes were facilitated or inhibited by TMS, depending on the stimulated region or parameters (Shapiro et al., 2001; Cappa et al., 2002).

The results of this study demonstrate that single-pulse TMS of the pSTS facilitated task performance, as the temporal areas have been claimed to be associated with perceptual discrimination (Lambert and Wootton, 2017). The positive outcome on behavior may result from the excitatory effects of TMS on neurons in the pSTS stimulated area (Pascual-Leone et al., 2000). However, Sack and Linden showed that TMS-induced changes in behavior may be ascribed not only to the stimulated area but also to the network of areas connected to the stimulated site (Sack and Linden, 2003).

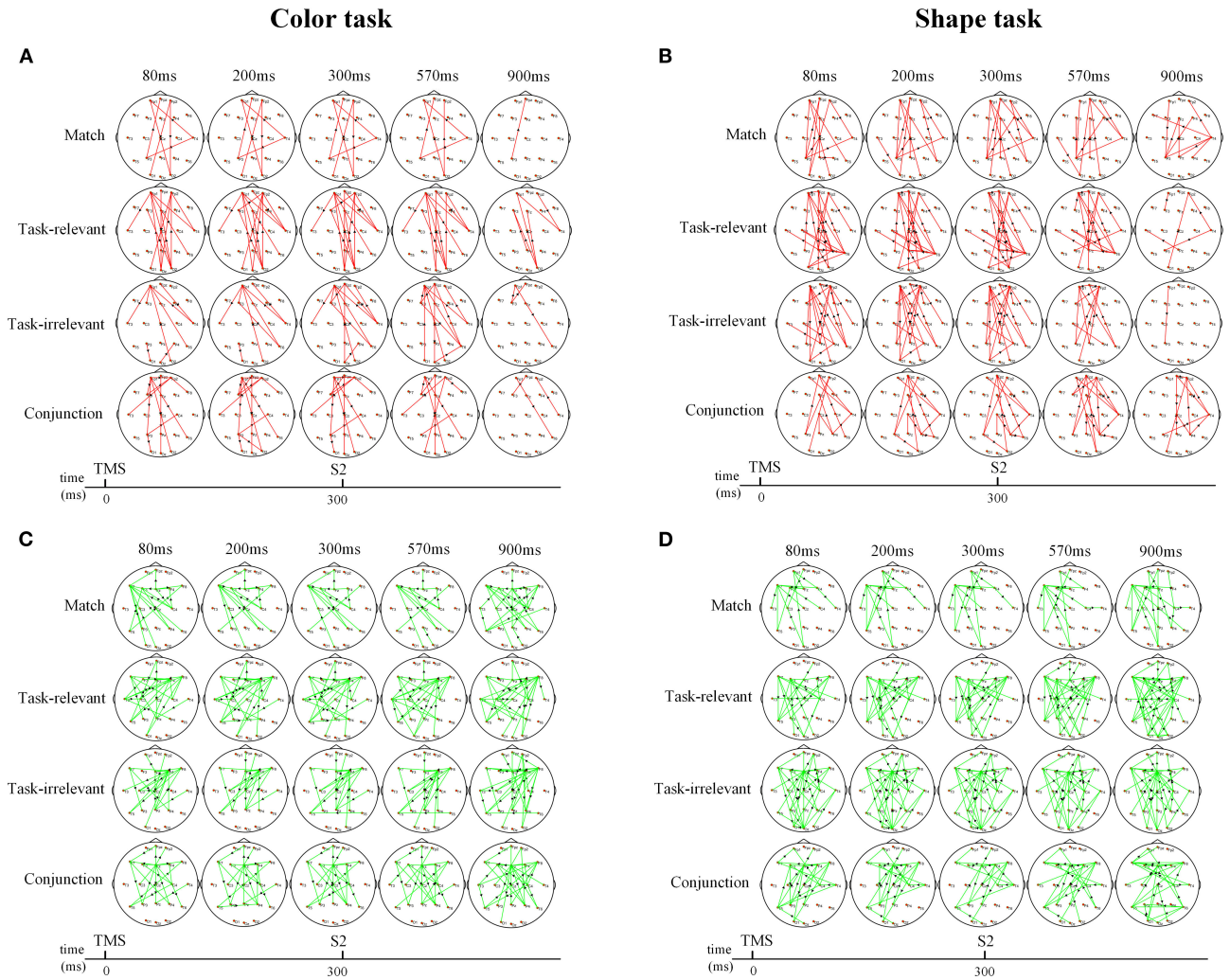

FIGURE 4 | The time-varying network changes between sham and real TMS in color and shape task. **(A,B)** Increased network connections between sham and real TMS condition. **(A)** The connections between the bilateral frontal pole and temporo-occipital region were significantly increased in the color task**. (B)** The connections between the bilateral frontal pole and temporo-occipital, parietal region were significantly increased in the shape task. **(C,D)** Decreased network connections between sham and real TMS condition. **(C)** The connections between the bilateral frontal region and temporo-parietal region, central region were significantly reduced in the color task. **(D)** The information outflows in the frontal and mid-line region were significantly reduced. Red and green lines indicate the increased connection and decreased connection, respectively. Arrows indicate the direction of information flow.

Thus, the greater performance may also probably result from the facilitation of intracortical networks associated with feature discrimination. The overall correct response rates for each type and task in two conditions were over 98%. It is noteworthy that although the RT showed the difference between sham and real TMS, the correct rate did not show a significant difference, which may be due to a ceiling effect.

P1 and N1

P1 and N1 are the earliest components linked to visual perceptual processing. N1 is related to the orientation of attention to relevant stimuli, and the P1 component reflects the early sensory processing of stimulus present to a specific location (Luck et al., 1990; Lin et al., 2019). Previous studies have shown that P1 amplitude enhancement is linked to the attention-related sensory gain control in the early stage of visual processing (Hillyard et al., 1998; Finnigan et al., 2011). The temporal cortex was claimed to be involved in the early stages of feature processing and perceptual discrimination (Lambert and Wootton, 2017). In the present study, P1 amplitude was more positive in the real TMS condition, suggesting the pSTS-TMS facilitates the early stage of feature discrimination, which may result from enhanced sensory gain control through the activation of pSTS.

N270

N270 amplitude was reduced in the real TMS condition with all mismatch types, while no difference was observed with the match type. This is similar to the previous finding that the

pSTS showed greater activation in response to incongruent vs. congruent stimulus pairs and bimodal vs. unimodal conflict (Hocking and Price, 2008). The pSTS plays an important role in conflicts related to face and orientation processing (Eacott et al., 1993), and the results show that pSTS is also involved in color and shape discrimination, as TMS of this brain area facilitated the processing of mismatched stimulus features to a greater extent than that of matched features.

It is well-established that the capacity for processing information is limited. Smaller N270 amplitude reflects reduced neuronal activities and less effort in conflict processing (Wang et al., 2001; Scannella et al., 2016). More attention is required to process conflicting or mismatched information as compared to matched information. The efficient and effective allocation of attentional resources during conflict processing is important for achieving optimal performance (Kramer et al., 1983). In the present study, TMS of the pSTS decreased the amplitude of N270 and shortened RTs, suggesting that less cognitive effort was needed for correct responses with pSTS-TMS. Thus, TMS of the pSTS preferentially enhanced the processing of conflicting over matched information by increasing the availability and efficiency of neurons involved in feature discrimination.

P3

The P3 component is elicited by cognitive tasks that engage attention after an unexpected stimulus (Polich, 1998). P3 amplitude reflects the mental representations of sensory stimuli at the final stage of information processing. In the present study, there was no difference in P3 amplitude between the sham and real TMS conditions, indicating that the pSTS is not involved in the late stage of visual information processing.

Time-Varying Network

In the present study, a comparison of the time-varying network changes between the sham and real TMS conditions revealed that TMS of the right pSTS enhanced information flow in the frontal pole, especially with mismatched stimuli.

Previous studies revealed that the frontal pole was associated with cognitive functions such as selective attention (Burgess et al., 2007) and multitasking (Gilbert et al., 2006). The frontal pole cortex is presumed to play a role in tracking and evaluating competing stimuli and is coactivated with the default mode network (Euston et al., 2012). A previous study also showed that patients with right frontal–temporal lobe brain damage have difficulty in performing visual searches due to deficits in feature-based control (Kumada and Hayashi, 2006).

The results of this study suggest that the frontal pole plays a key role in feature discrimination. TMS of the right pSTS activated bilateral frontal poles and enhanced their connectivity with other brain regions, especially the temporo-parietal and occipital regions. This confirms that TMS can modulate neuronal activity beyond the site of stimulation, impacting a distributed network of brain areas (Ferreri et al., 2011). The information flow and formation of a dynamic attention network with the frontal pole as the core were essential for color and shape discrimination, with the pSTS facilitating attentional processing by increasing information flow in the frontal pole.

LIMITATIONS

The present study had some limitations. First, the experimental conditions were relatively limited in terms of target regions (e.g., vertex or motor cortex) and stimulation conditions. Second, as is common in TMS–EEG studies, the TMS pulse caused a high-amplitude artifact in the EEG signals that lasted 200 ms, which was beyond the scope of our analysis. However, we performed ICA and discarded the signal for 60 ms after each pulse to remove these artifacts. Nonetheless, future studies should address these limitations.

CONCLUSION

In the present study, we investigated the effects of single-pulse TMS over the right pSTS in the dual-feature delayed matching task. We found that the right pSTS activation improved task performance, especially with mismatched types, and reduced N270 amplitude. The results of the time-varying network analysis revealed that pSTS-TMS altered the dynamic attention network by increasing the connectivity between bilateral frontal poles and the temporo-parieto-occipital regions. These findings provide evidence that the pSTS plays an important role in the feature discrimination aspect of attentional processing.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of the Xuanwu Hospital, Capital Medical University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

YW conceived and designed the study. QZ acquired the data and drafted the manuscript. PS, XW, and HL contributed to data analysis and interpretation. All authors contributed to the article and approved the submitted version.

FUNDING

The study was supported in part by the National Key Research and Development Program of China (Grant Nos. 2019YFC0121202, 2019YFC0121203, and 2018YFC1314500); the National Natural Science Foundation of China (Grant Nos. 81501119, 82001388, and 82001374); the China Postdoctoral Science Foundation (Grant No. 2019M660720); and the Beijing Postdoctoral Research Foundation (Grant No. 2020-ZZ-014).

REFERENCES

Barker, A. T., and Jalinous, R. (1985). Non-invasive magnetic stimulation of human motor cortex. *Lancet* 1, 1106–1107. doi: 10.1016/S0140-6736(85)92413-4

Beauchamp, M. S. (2011). "Biological motion and multisensory integration: the role of the superior temporal sulcus," in *The Science of Social Vision*, ed R. B. Adams (New York, NY: Oxford University Press), 409–420. doi: 10.1093/acprof:oso/9780195333176.003.0024

Beauchamp, M. S., Yasar, N. E., Frye, R. E., and Ro, T. (2008). Touch, sound and vision in human superior temporal sulcus. *Neuroimage* 41, 1011–1020. doi: 10.1016/j.neuroimage.2008.03.015

Bergmann, T. O., Karabanov, A., Hartwigsen, G., Thielscher, A., and Siebner, H. R. (2016). Combining non-invasive transcranial brain stimulation with neuroimaging and electrophysiology: current approaches and future perspectives. *Neuroimage* 140, 4–19. doi: 10.1016/j.neuroimage.2016.02.012

Bogadhi, A. R., Bollimunta, A., Leopold, D. A., and Krauzlis, R. J. (2018). Brain regions modulated during covert visual attention in the macaque. *Sci Rep*. 8:15237. doi: 10.1038/s41598-018-33567-9

Bogadhi, A. R., Bollimunta, A., Leopold, D. A., and Krauzlis, R. J. (2019). Spatial attention deficits are causally linked to an area in macaque temporal cortex. *Curr. Biol*. 29, 726–736.e4. doi: 10.1016/j.cub.2019.01.028

Burgess, P. W., Dumontheil, I., and Gilbert, S. J. (2007). The gateway hypothesis of rostral prefrontal cortex (area 10) function. *Trends Cogn Sci*. 11, 290–298. doi: 10.1016/j.tics.2007.05.004

Cappa, S. F., Sandrini, M., Rossini, P. M., Sosta, K., and Miniussi, C. (2002). The role of the left frontal lobe in action naming: rTMS evidence. *Neurology* 59, 720–723. doi: 10.1212/WNL.59.5.720

Carter, C. S., Macdonald, A. M., Botvinick, M., Ross, L. L., Stenger, V. A., Noll, D., et al. (2000). Parsing executive processes: strategic vs. evaluative functions of the anterior cingulate cortex. *Proc. Natl. Acad. Sci. U.S.A*. 97, 1944–1948. doi: 10.1073/pnas.97.4.1944

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

Eacott, M. J., Heywood, C. A., Gross, C. G., and Cowey, A. (1993). Visual-discrimination impairments following lesions of the superior temporal sulcus are not specific for facial stimuli. *Neuropsychologia* 31, 609–619. doi: 10.1016/0028-3932(93)90055-5

Euston, D. R., Gruber, A. J., and McNaughton, B. L. (2012). The role of medial prefrontal cortex in memory and decision making. *Neuron* 76, 1057–1070. doi: 10.1016/j.neuron.2012.12.002

Ferreri, F., Pasqualetti, P., Maatta, S., Ponzo, D., Ferrarelli, F., Tononi, G., et al. (2011). Human brain connectivity during single and paired pulse transcranial magnetic stimulation. *Neuroimage* 54, 90–102. doi: 10.1016/j.neuroimage.2010.07.056

Finnigan, S., O'Connell, R. G., Cummins, T. D., Broughton, M., and Robertson, I. H. (2011). ERP measures indicate both attention and working memory encoding decrements in aging. *Psychophysiology* 48, 601–611. doi: 10.1111/j.1469-8986.2010.01128.x

Gilbert, S. J., Spengler, S., Simons, J. S., Steele, J. D., Lawrie, S. M., Frith, C. D., et al. (2006). Functional specialization within rostral prefrontal cortex (area 10): a meta-analysis. *J. Cogn. Neurosci*. 18, 932–948. doi: 10.1162/jocn.2006.18.6.932

Hallett, M. (2007). Transcranial magnetic stimulation: a primer. *Neuron* 55, 187–199. doi: 10.1016/j.neuron.2007.06.026

Hein, G., and Knight, R. T. (2008). Superior temporal sulcus-it's my area: or is it? *J. Cogn. Neurosci*. 20, 2125–2136. doi: 10.1162/jocn.2008.20148

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

Hocking, J., and Price, C. J. (2008). The role of the posterior superior temporal sulcus in audiovisual processing. *Cereb Cortex*. 18, 2439–2449. doi: 10.1093/cercor/bhn007

Hyvärinen, A., and Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural Netw*. 13, 411–430. doi: 10.1016/S0893-6080(00)00026-5

Ilmoniemi, R. J., and Kicic, D. (2010). Methodology for combined TMS and EEG. *Brain Topogr*. 22, 233–248. doi: 10.1007/s10548-009-0123-4

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

Kramer, A. F., Wickens, C. D., and Donchin, E. (1983). An analysis of the processing requirements of a complex perceptual-motor task. *Hum. Fact*. 25, 597–621. doi: 10.1177/001872088302500601

Kumada, T., and Hayashi, M. (2006). Deficits in feature-based control of attention in a patient with a right fronto-temporal lesion. *Cogn. Neuropsychol*. 23, 401–423. doi: 10.1080/02643290542000085

Lambert, A. J., and Wootton, A. (2017). The time-course of activation in the dorsal and ventral visual streams during landmark cueing and perceptual discrimination tasks. *Neuropsychologia* 103, 1–11. doi: 10.1016/j.neuropsychologia.2017.07.002

Lin, Y., Cui, S., Du, J., Li, G., He, Y., Zhang, P., et al. (2019). N1 and P1 components associate with visuospatial-executive and language functions in normosmic parkinson's disease: an event-related potential study. *Front. Aging Neurosci*. 11:18. doi: 10.3389/fnagi.2019.00018

Luck, S. J. (2014). *An Introduction to the Event-related Potential Technique*. Cambridge, MA: The MIT Press.

Luck, S. J., Heinze, H. J., Mangun, G. R., and Hillyard, S. A. (1990). Visual event-related potentials index focused attention within bilateral stimulus arrays. II. Functional dissociation of P1 and N1 components. *Electroencephalogr. Clin. Neurophysiol*. 75, 528–542. doi: 10.1016/0013-4694(90)90139-B

Majerus, S., Peters, F., Bouffier, M., Cowan, N., and Phillips, C. (2018). The dorsal attention network reflects both encoding load and top-down control during working memory. *J. Cogn. Neurosci*. 30, 144–159. doi: 10.1162/jocn_a_01195

Mao, W., and Wang, Y. (2008). The active inhibition for the processing of visual irrelevant conflict information. *Int. J. Psychophysiol*. 67, 47–53. doi: 10.1016/j.ijpsycho.2007.10.003

Moore, T., and Zirnsak, M. (2017). Neural mechanisms of selective visual attention. *Annu. Rev. Psychol*. 68, 47–72. doi: 10.1146/annurev-psych-122414-033400

Pascual-Leone, A., Walsh, V., and Rothwell, J. (2000). Transcranial magnetic stimulation in cognitive neuroscience - virtual lesion, chronometry, and functional connectivity. *Curr. Opin. Neurobiol*. 10, 232–237. doi: 10.1016/S0959-4388(00)00081-7

Pellicciari, M. C., Veniero, D., and Miniussi, C. (2017). Characterizing the cortical oscillatory response to tMS Pulse. *Front. Cell Neurosci*. 11:38. doi: 10.3389/fncel.2017.00038

Polich, J. (1998). P300 clinical utility and control of variability. *J. Clin. Neurophysiol*. 15, 14–33. doi: 10.1097/00004691-199801000-00004

Rogasch, N. C., and Fitzgerald, P. B. (2013). Assessing cortical network properties using TMS-EEG. *Hum. Brain Mapp*. 34, 1652–1669. doi: 10.1002/hbm.22016

Sack, A. T., and Linden, D. E. J. (2003). Combining transcranial magnetic stimulation and functional imaging in cognitive brain research: possibilities and limitations. *Brain Res. Rev*. 43, 41–56. doi: 10.1016/S0165-0173(03)00191-7

Scannella, S., Pariente, J., De Boissezon, X., Castel-Lacanal, E., Chauveau, N., Causse, M., et al. (2016). N270 sensitivity to conflict strength and working memory: a combined ERP and sLORETA study. *Behav. Brain Res*. 297, 231–240. doi: 10.1016/j.bbr.2015.10.014

Shapiro, K. A., Pascual-Leone, A., Mottaghy, F. M., Gangitano, M., and Caramazza, A. (2001). Grammatical distinctions in the left frontal cortex. *J. Cogn. Neurosci*. 13, 713–720. doi: 10.1162/08989290152541386

Siebner, H. R., Bergmann, T. O., Bestmann, S., Massimini, M., Johansen-Berg, H., Mochizuki, H., et al. (2009). Consensus paper: combining transcranial stimulation with neuroimaging. *Brain Stimul*. 2, 58–80. doi: 10.1016/j.brs.2008.11.002

Stemmann, H., and Freiwald, W. A. (2019). Evidence for an attentional priority map in inferotemporal cortex. *Proc. Natl. Acad. Sci. U.S.A*. 116, 23797–23805. doi: 10.1073/pnas.1821866116

Taylor, M. J. (2002). Non-spatial attentional effects on P1. *Clin. Neurophysiol*. 113, 1903–1908. doi: 10.1016/S1388-2457(02)00309-7

Wang, H., Wang, Y., Kong, J., Cui, L., and Tian, S. (2001). Enhancement of conflict processing activity in human brain under task relevant condition. *Neurosci. Lett*. 298, 155–158. doi: 10.1016/S0304-3940(00)01757-2

Wang, Y., Cui, L., Wang, H., Tian, S., and Zhang, X. (2004). The sequential processing of visual feature conjunction mismatches in the human brain. *Psychophysiology* 41, 21–29. doi: 10.1111/j.1469-8986.2003.00134.x

Wang, Y., Tian, S., Wang, H., Cui, L., Zhang, Y., and Zhang, X. (2003). Event-related potentials evoked by multi-feature conflict under different attentive conditions. *Exp. Brain Res.* 148, 451–457. doi: 10.1007/s00221-002-1319-y

Wang, Y., Wang, H., Cui, L., Tian, S., and Zhang, Y. (2002). The N270 component of the event-related potential reflects supramodal conflict processing in humans. *Neurosci. Lett.* 332, 25–28. doi: 10.1016/S0304-3940(02)00906-0

Yang, J., and Wang, Y. (2002). Event-related potentials elicited by stimulus spatial discrepancy in humans. *Neurosci. Lett.* 326, 73–76. doi: 10.1016/S0304-3940(02)00204-5

Zhang, L., Liang, Y., Li, F., Sun, H., Peng, W., Du, P., et al. (2017). Time-varying networks of inter-ictal discharge g reveal epileptogenic zone. *Front. Comput. Neurosci.* 11:77. doi: 10.3389/fncom.2017.00077

Zhang, R., Hu, Z., Roberson, D., Zhang, L., Li, H., and Liu, Q. (2013). Neural processes underlying the"same"-"different" judgment of two simultaneously presented objects–an EEG study. *PLoS One* 8:e81737. doi: 10.1371/journal.pone.0081737

Zhang, X., Wang, Y., Li, S., and Wang, L. (2003). Event-related potential N270, a negative component to identification of conflicting information following memory retrieval. *Clin. Neurophysiol.* 114, 2461–2468. doi: 10.1016/S1388-2457(03)00251-7

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 in Human Neuroscience

ORIGINAL RESEARCH
published: 16 August 2021
doi: 10.3389/fnhum.2021.659040

Driving With Distraction: Measuring Brain Activity and Oculomotor Behavior Using fMRI and Eye-Tracking

*Nicole H. Yuen[1,2], Fred Tam[2], Nathan W. Churchill[3], Tom A. Schweizer[3,4] and Simon J. Graham[1,2]**

[1] Department of Medical Biophysics, Faculty of Medicine, University of Toronto, Toronto, ON, Canada, [2] Physical Sciences Platform, Sunnybrook Research Institute, Toronto, ON, Canada, [3] Keenan Research Centre for Biomedical Science, St. Michael's Hospital, Toronto, ON, Canada, [4] Division of Neurosurgery, St. Michael's Hospital, Toronto, ON, Canada

Introduction: Driving motor vehicles is a complex task that depends heavily on how visual stimuli are received and subsequently processed by the brain. The potential impact of distraction on driving performance is well known and poses a safety concern – especially for individuals with cognitive impairments who may be clinically unfit to drive. The present study is the first to combine functional magnetic resonance imaging (fMRI) and eye-tracking during simulated driving with distraction, providing oculomotor metrics to enhance scientific understanding of the brain activity that supports driving performance.

Materials and Methods: As initial work, twelve healthy young, right-handed participants performed turns ranging in complexity, including simple right and left turns without oncoming traffic, and left turns with oncoming traffic. Distraction was introduced as an auditory task during straight driving, and during left turns with oncoming traffic. Eye-tracking data were recorded during fMRI to characterize fixations, saccades, pupil diameter and blink rate.

Results: Brain activation maps for right turns, left turns without oncoming traffic, left turns with oncoming traffic, and the distraction conditions were largely consistent with previous literature reporting the neural correlates of simulated driving. When the effects of distraction were evaluated for left turns with oncoming traffic, increased activation was observed in areas involved in executive function (e.g., middle and inferior frontal gyri) as well as decreased activation in the posterior brain (e.g., middle and superior occipital gyri). Whereas driving performance remained mostly unchanged (e.g., turn speed, time to turn, collisions), the oculomotor measures showed that distraction resulted in more consistent gaze at oncoming traffic in a small area of the visual scene; less time spent gazing at off-road targets (e.g., speedometer, rear-view mirror); more time spent performing saccadic eye movements; and decreased blink rate.

Conclusion: Oculomotor behavior modulated with driving task complexity and distraction in a manner consistent with the brain activation features revealed by fMRI.

OPEN ACCESS

Edited by:
Kuniyoshi L. Sakai, The University of Tokyo, Japan

Reviewed by:
Nina Maria Hanning, Ludwig Maximilian University of Munich, Germany
Stephen D. Mayhew, University of Birmingham, United Kingdom

***Correspondence:**
Simon J. Graham
simon.graham@utoronto.ca

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

Received: *26 January 2021*
Accepted: *29 June 2021*
Published: *16 August 2021*

Citation:
Yuen NH, Tam F, Churchill NW, Schweizer TA and Graham SJ (2021) Driving With Distraction: Measuring Brain Activity and Oculomotor Behavior Using fMRI and Eye-Tracking. Front. Hum. Neurosci. 15:659040. doi: 10.3389/fnhum.2021.659040

The results suggest that eye-tracking technology should be included in future fMRI studies of simulated driving behavior in targeted populations, such as the elderly and individuals with cognitive complaints – ultimately toward developing better technology to assess and enhance fitness to drive.

Keywords: driving simulation, distraction, neural correlates of driving, fMRI, eye-tracking

INTRODUCTION

Distracted driving poses a danger to drivers, passengers, pedestrians and cyclists and is a growing threat to road safety across the world. For example, the United States (US) National Highway Traffic Safety Administration reported 37,461 fatalities caused by motor vehicle crashes (MVCs) in 2016, with over 2 million people in the United States injured in MVCs each year (National Center for Statistics and Analysis, 2018). In many of these cases, the MVC occurred because drivers were engaged in dangerous multi-tasking (including activities such as texting, eating, or talking to a passenger). Such behavior can distract the driver from performing driving-related tasks safely, such as maintaining proper lane position (Hamish Jamson and Merat, 2005). The rapid development and widespread use of smart cellphone technology in particular has increased concerns over distracted driving - as texting while driving takes a hand away from the wheel, decreases visual attention to the road (Fitch et al., 2015), and impairs driving response and reaction time (Strayer and Drew, 2004; Drews et al., 2009). Studies have shown that distracted driving with cell phone usage results in failure to stop completely at stop signs, delayed braking responses and more rear-end collisions (Strayer and Drew, 2004; Kramer et al., 2006; Thompson et al., 2012). Cell phone usage alone has been attributed to approximately 15 – 25% of fatal distracted driving crashes (National Center for Statistics and Analysis, 2013; Ortiz et al., 2016).

From the perspective of psychological science, driving is a complex, demanding task that requires numerous cognitive abilities, including attention, memory, decision-making, and alertness to adapt to a rapidly changing environment. A decline in driving performance can arise due to processes that negatively affect such abilities, such as distraction (Lee et al., 2003; Cantin et al., 2009; Rizzo, 2011). Distraction, introduced as a secondary task that requires executive functions (as opposed to distraction by sudden, attention capturing events, such as children running into the street) has the potential to divert attention and resources from the primary driving task, thereby increasing risk of MVCs from degraded driving ability (Thompson et al., 2012).

Moreover, cognitive abilities are relevant for processing and responding to the continuous stream of sensory input that is received by the driver, requiring them to make informed and correct decisions, especially during complex road or traffic scenarios. Examples of these scenarios include driving at busy intersections, making left turns with oncoming traffic, or merging into traffic and making lane changes (Thompson et al., 2012; Noyce et al., 2017). Declines in cognitive abilities also pose a safety risk and potentially make the driver more prone to the effects of distracted driving (Fraade-Blanar et al., 2018). Elderly and middle-aged drivers have been shown to commit significantly more driving errors and reduced steering wheel control when distracted (Thompson et al., 2012).

In many cases, physicians lack quantitative and objective tools to assess fitness to drive in patients, and must make such decisions based on clinical exam and rapid cognitive assessments that are often insufficiently informative (Marshall and Gilbert, 1999; Withaar et al., 2000). The decision to declare an individual unfit to drive is also complicated by the fact that in the developed world, driving is very often key to functional dependence and to employment.

Neuropsychological tests (NPTs) may also have a role in clinical and scientific settings in the assessment of fitness to drive and the ability to drive safely. These standardized behavioral tests are more detailed than clinical assessments and can be used to assess cognitive functions that underpin driving abilities, determining whether test performance is impaired in relation to population norms (Lundberg et al., 1997; Lezak et al., 2004). Although NPTs are useful to identify cognitive decline, they have not been established as good predictors of driving ability, however. For example, one study investigating a range of NPTs found that the best four-test combination was able to identify at-risk drivers with 95% specificity, but only 80% sensitivity (Bowers et al., 2013). Further research is needed to develop complex cognitive tasks or screening tools to predict driving performance better. As an initial part of this work, there is a strong need to understand the underlying brain activity associated with normal and unimpaired driving performance. From this information, it may be possible in the future to develop objective, robust behavioral methods to assess fitness to drive.

Many researchers have followed this line of thinking to study the task-related brain activity associated with various types of driving behavior, especially using the method of functional magnetic resonance imaging (fMRI). By necessity, these studies involve simulated rather than actual driving behavior, using various forms of virtual reality or video game technology to give test participants the simulated experience of driving while they lie in the magnet bore of an MRI system. The costs associated with fMRI and other practical considerations (e.g., access and availability) make it very unlikely that this imaging method can provide direct utility as a clinical tool for assessing fitness to drive – however, fMRI is very useful for providing scientific insight. Simulated driving studies using fMRI have revealed a broad network of task-related brain activity, engaging areas in all major lobes of the brain (Calhoun et al., 2002; Graydon et al., 2004; Callan et al., 2009; Crundall and Underwood, 2011; Calhoun and Pearlson, 2012; Choi

et al., 2017). In simple straight driving, increased activations in the visual-association, parietal and occipital regions were observed, related to visuomotor integration; additionally, the precentral gyrus, superior and inferior parietal lobules, and cerebellum were seen to activate and contribute to motor control (Calhoun and Pearlson, 2012). In a study investigating driving while concurrently performing arithmetic tasks, activations of the motor cortex, parietal and occipital lobes were seen to decrease in comparison to the driving only task; whereas the temporal and inferior frontal regions, associated with auditory processing and additional task performance, showed an increase in activation when concurrently performing a secondary task (Choi et al., 2017). Similarly, activation in the motor cortex, parietal and occipital lobes and superior temporal gyrus was reported when driving with an auditory task (Uchiyama et al., 2012). Furthermore, an auditory distraction task was reported to shift activation from the posterior and visual areas to the frontal regions, which are responsible for planning, decision-making and cognition (Schweizer et al., 2013).

Despite the knowledge gained from these studies, the interpretation of the brain activity associated with simulated driving behavior remains limited, as many brain areas are active when performing simulated driving tasks and it is difficult to link the contribution of specific areas to the various behavioral subcomponents of driving behavior. One method to overcome this difficulty involves augmenting the behavioral recording that is undertaken during fMRI experiments of simulated driving. In particular, vision and visual attention are extremely crucial to safe driving. Attention has been described to be involved in guiding eye movements, including point of gaze spatiotemporal characteristics, eye blink rate and pupil dilation (Schneider, 1995; Moore and Fallah, 2001). Such parameters can be readily measured using modern eye-tracking devices (Land, 2006). The introduction of a secondary task to driving has been shown to decrease the percentage of gaze points in peripheral regions (e.g., speedometer, rearview mirror), and to cause spatial gaze concentration due to the increased mental workload (Recarte and Nunes, 2003; Tsai et al., 2007). In addition, eye blink rate and duration have been shown to have associations with cognitive load in a range of behavior including word-naming tasks (Ohira, 1996), visuospatial memory tasks (Van Orden et al., 2001) and driving in various environments (Brookings et al., 1996; Tsai et al., 2007; Haak et al., 2009; McIntire et al., 2014; Faure et al., 2016). It has been suggested that the blink rate decreases to minimize the loss of incoming information, or to prevent disruption to cognitive processes involved in the mental task (Holland and Tarlow, 1972; McIntire et al., 2014; Maffei and Angrilli, 2018). One study that investigated changes in driving behavior in the presence of a secondary cognitive task found that when the driving environment was more demanding, blink rate decreased (Faure et al., 2016). Pupil dilation has also been observed to modulate cognitive processing. Greater dilation has been correlated with greater cognitive load across various mental tasks (Recarte and Nunes, 2003; Tsai et al., 2007; Palinko et al., 2010; Gable et al., 2015; Niezgoda et al., 2015) and individuals with better task performance have been shown to undergo a larger percent change in pupil size (Tsai et al., 2007; Palinko et al., 2010).

Because eye-tracking technology can be used to measure cognitive load, combining eye-tracking with fMRI is likely to support and inform the interpretations of the ensuing activation maps associated with simulated driving behavior by providing additional information about the participant behaviors during each task. At present, no studies have been conducted using both fMRI and eye-tracking simultaneously in the context of simulated driving.

To fill this gap in the scientific literature, the present study adopts simultaneous eye-tracking and fMRI. It is hypothesized that complex driving conditions will engage greater brain activity in areas involved in visual and motor processing, and that the addition of a distracting cognitive task will engage medial and lateral frontal areas, while decreasing the extent of occipital activation. It is also hypothesized that an increase in driving task demand and distraction during simulated driving will result in decreased gaze distribution and blink rate, and increased pupil diameter, and that these effects will correlate with similar changes in fMRI signals.

MATERIALS AND METHODS

Ethics Statement

This study was carried out in accordance with the recommendations of the Research Ethics Board at Sunnybrook Health Sciences Centre, Canada. All participants were given a full explanation of the experimental procedures and provided written informed consent prior to participating in the study.

Participants

Nineteen healthy adults between the ages of 20 and 30 were recruited for this study. Seven participants were excluded due to unusable eye-tracker data quality (arising from factors such as blue eye color, and obstruction of the view between the eye and the eye-tracking video camera due to long eyelashes). Twelve participants remained for subsequent analysis (4 females and 8 males, mean age = 23.4 years, SD = 1.1 years) with mean driving experience of 5.3 years (SD = 2.1 years, range = 1.5 – 9.0 years). All participants had a valid driver's license in the Canadian province of Ontario, no history of psychological or neurological illness, and were able to be imaged in a 3 T MRI system (i.e., they were free from MRI exclusion criteria, such as claustrophobia or ferromagnetic implants). All participants were right-handed with normal or corrected vision. Ten of the participants were right-eye dominant (4 females, 6 males), whereas two were left-eye dominant (2 males).

Driving Simulation

Simulated driving tasks were administered using STISIM Drive Software (Systems Technology, Inc., Hawthorne, CA, United States), controlled by custom fMRI-compatible driving simulator hardware (with foot pedals and steering wheel) (Kan et al., 2012). The steering wheel included response buttons embedded on the wheel that enabled participants to answer questions while driving, similar to modern steering wheels in real life that include buttons to adjust volume or answer cell

phone calls. The simulation environment was shown on a screen using an fMRI-compatible projector (Avotec SV-6011 LCD Projection System, Stuart, FL). The environment was observed by the participants through a mirror mounted on the head coil, and auditory stimuli were delivered through fMRI-compatible headphones (Avotec SS-3100, Stuart, FL).

Prior to the fMRI experiment, participants underwent training in an fMRI simulator for approximately one hour to familiarize them with the driving hardware and software. The training scenario included the same tasks to be performed during fMRI, in pseudo-randomized order. Participants were told to follow traffic laws and maintain the posted speed limit of 60 km/h during the session.

Driving Tasks

When in the magnet, participants were again instructed to adhere to traffic lights and road rules while maintaining the posted speed of 60 km/h. The environment of the driving scenario was mainly rural, with minimal scenery, trees and buildings. Driving tasks included straight driving ("Straight," "S," **Figure 1A**) and turns at intersections with and without oncoming traffic, with or without having to perform an auditory task simultaneously. At task onset, participants were presented with an auditory recording of a male voice to mimic instructions from a modern navigation device (e.g., "At the intersection, turn right"). The recording occurred one hundred virtual meters in advance of the intersection (a time duration of approximately 6 s when driving at the speed limit). For all turning tasks, the intersection included a set of traffic lights with the appropriate light in the green "go" condition. Each participant performed eight trials of right-hand turns ("Right Turn," "R," not shown) and seven trials of left-hand turns ("Left Turn," "L," **Figure 1B**) without any traffic or distraction. To increase complexity of the driving task, oncoming traffic was introduced to left turns (without distraction), which required participants to decide when it was safe to turn ("Left Turn + Traffic," "LT"; seven trials, **Figure 1C**). To simulate distraction, an auditory task was presented to the participant as a general knowledge true or false question (e.g., "a hammer is lighter than a feather") during straight driving ("Straight + Audio," "SA"; six trials) and in left turns with oncoming traffic ("Left Turn + Traffic + Audio," "LTA"; seven trials), immediately after the turn instruction was presented. Questions were answered by pressing buttons corresponding to true and false on the steering wheel. During LTA and SA conditions, participants were not told to prioritize either the driving or the auditory task; and no error feedback was provided throughout all the task conditions, other than what each participant detected using their senses. The simulated driving scenarios were administered in pseudo-randomized order and were split into two runs to enable participants to remain vigilant without becoming fatigued. Each trial duration was approximately 20 s long, and each run had a duration of approximately 12 min (depending on the driving speed of the participant), for a total driving time of approximately 25 min. Straight driving after each turn task served as the baseline condition and was the same for all participants. Driving task trials were separated by at least 10 s of baseline straight driving to minimize overlapping of fMRI and eye movement response signals between the various driving task conditions.

Functional MRI Protocol

Once each participant completed the fMRI simulator session, they were asked to complete a screening form to ensure that they were able to undergo fMRI. The form was reviewed by the MRI technologist and after clearance was granted to enter the magnet room, the participant was positioned on the patient table of the MRI system. To increase participant comfort and to minimize body and head motion, pillows were placed under the knees for leg support, and pads were placed under the arms and elbows and beside the head. The foot pedals (accelerator and brake) were placed appropriately such that when they were fully depressed, there was no bulk body movement that could translate into unwanted head motion. The steering wheel was then placed over the torso. The fMRI-compatible headset was fitted on the participant over noise-reducing ear plugs, and the sound system was tested to ensure that the participant was able to hear instructions clearly during the driving scenarios. Because both hands were used to steer, the peripheral pulse monitor was placed on the toe to record heart beat information, rather than the standard placement on the finger. Respiration was monitored, and an emergency call squeeze bulb was placed on the participant so that they could discontinue the study for any reason, if they wished.

Two fMRI runs with simulated driving were performed to enable the participant to maintain vigilance throughout task performance and to reduce fatigue. Anatomical imaging was acquired at high spatial resolution in between the runs to allow the participant to rest their eyes. The fMRI session took approximately one hour to complete, involving 30 min of setup and calibration and 30 minutes of imaging. The driving scenarios were triggered synchronously with the start of fMRI data collection. Images were acquired using a 3.0 Tesla MRI system (Prisma, Siemens, Erlangen, Germany). Functional MRI during simulated driving was performed using $T2^*$-weighted echo planar imaging (EPI; repetition time (TR) = 1.75 s, echo time (TE) = 30 ms, flip angle (FA) = 40°, field of view (FOV) = 256 × 256 mm^2, 60 slices, voxel size = 2.5 × 2.5 × 2.5 mm^3, 377 time points). Anatomical imaging was performed using T1-weighted, three-dimensional magnetization prepared rapid acquisition gradient echo imaging (3D MPRAGE; TR = 1.8 s, TE = 2.21 ms, FA = 10°, FOV = 256 × 256 mm^2, 176 slices, voxel size = 1.0 × 1.0 × 1.0 mm^3).

Eye-Tracking Setup

An fMRI-compatible high-speed eye-tracker was used to record movements of the right eye for consistency of methodology, at 1000 Hz using a monocular system with a 50 mm lens (EyeLink 1000 Plus, SR Research Ltd., Mississauga, Canada) and infrared illumination (910 nm wavelength). The eye was tracked using corneal reflection, and the pupil was detected using a centroid fitting algorithm, using the standard hardware and software of the system. Metrics that were measured included brief pauses (ocular fixations) as well as rapid shifts (saccades) in the point of

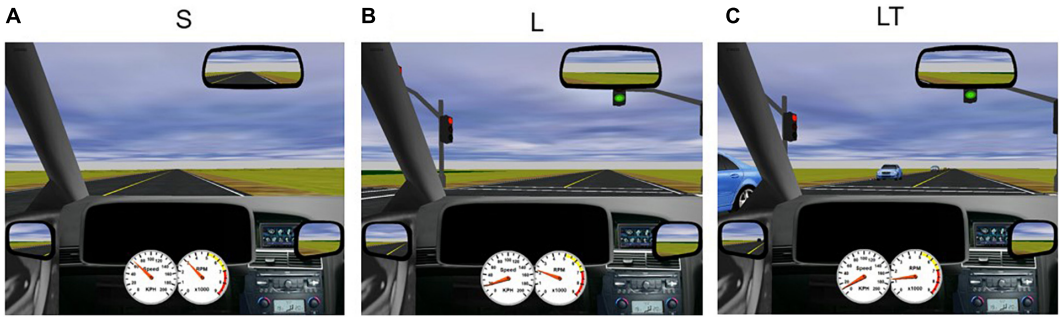

FIGURE 1 | Screenshots of STISIM driving scenarios. **(A)** S, straight driving; **(B)** L, left turn at an intersection without traffic; and **(C)** LT, left turn with traffic.

gaze, fixations and blinks. Each sample included the time latency (measured from the time the tracker software was started), eye position on the display screen (measured as coordinates on the calibrated display screen), and pupil size (measured in arbitrary units).

If vision correction was required, the appropriate prescription MRI-compatible lens was selected and prepared. When the lens was placed in MRI-compatible frames, it created reflection onto the eye-tracker camera, preventing the eye to be tracked. To avoid this, the lens was gently placed on the participant's face, supported with tape and cotton pads (five participants). To calibrate the eye-tracking system, the participant was instructed to fixate on each of thirteen target locations on the display screen, appearing consecutively after a fixation on the target was detected. Following calibration, a validation procedure was undertaken to measure gaze position accuracy on the screen using the prior calibration parameters. Calibration and validation were performed before each simulated driving run. Eye-tracking measurements were started manually, then subsequent initialization of the fMRI acquisition sent a synchronous trigger to the eye-tracker to denote the time that the fMRI and driving scenario were started.

Data Extraction and Analysis

Driving Simulator Data

Driving simulator data were automatically generated and saved at the end of each driving scenario as an STISIM Drive data file. This output was parsed to extract the time series data of driving metrics that were measured. Driving metrics extracted included length of run, longitudinal position, lane position (deviation of lateral lane position of the vehicle referenced in relation to the center of the driver vehicle with respect to the roadway center dividing line), driving speed, vehicle heading angle, and number of collisions. These metrics, which have been used to characterize impaired driving performance in various patient populations with neurological impairment (e.g., Mansur et al., 2018; Hird et al., 2017), were converted into a matrix of size N × M (with N variables and M time points) and read into MATLAB (The MathWorks Inc., Natick, MA, United States). Metrics that were calculated in MATLAB included event onset time (longitudinal position of event onset converted into time), time to turn (time between event onset and when the vehicle heading angle was at 25°, considered to be start of the turn position), turn speed (mean speed between turn start time and turn end time, after the vehicle heading angle returned to 0°), and question response time (time a response button was pressed minus the time the question was administered). Metrics of interest were concatenated for both driving runs. Statistical analyses of behavioral data were completed using Statistical Product and Service Solutions (SPSS) software (IBM SPSS Statistics Version 25, IBM, Armonk, NY, United States). Repeated Measures ANOVA tests were used to determine if there were significant differences between the time to turn and turn speed across R, L, LT, and LTA conditions, with subsequent Bonferroni-corrected *post hoc* contrasts to determine which factors were significantly different. Paired samples *t*-tests were used to assess significance of driving speed and lane position between S and SA conditions, and question response time and accuracy between SA and LTA conditions.

MRI Data

The fMRI time series data were preprocessed using Analysis of Functional NeuroImages (AFNI) freeware (Cox, 1996) to remove large spikes in the time series, to correct for the physiological effects of the cardiac and respiratory cycles, to suppress artifacts from head motion correction (and ensuring no excessive head motion), and to correct for the slight differences in time that each image slice was collected in each multi-slice acquisition of the brain volume (i.e., "slice-time correction"). The data were spatially smoothed using a 5 mm full width at half maximum Gaussian kernel, then normalized by the mean of each voxel. From the driving simulator data, the times of each event onset for all trials of all tasks (L, R, LT, LTA, S and SA conditions as defined in section "Driving Tasks") were extracted for both driving runs and saved as individual text files for each task. These event onset times, reflecting the time the auditory instructions were presented, were used to enter a stimulus timing file into a general linear model (GLM) using straight driving (the S condition) as the control, to estimate task-related brain activity at each voxel location. The stimulus-timing file, which included the full 20 s duration of each driving task, was convolved with a canonical hemodynamic response function with six regressors (head motion parameters in six degrees of freedom). The maps were then averaged in Talairach and Tournoux (TT) atlas space (Talairach and Tournoux, 1988) and thresholded using a false

discovery rate (Genovese et al., 2002) of q = 0.05 to correct for multiple statistical comparisons. Maps for all 12 participants were entered into a within-participants, random effects analysis of variance (ANOVA) to create final group activation maps, which were overlaid on an anatomical average in Talairach atlas space for interpretation. Activation maps were generated for the following task contrasts of interest: R - S, L - S, (L – S) – (R – S), LT - S, LT – L, LT – R, (LT – L) – (LT – R), SA - S, LTA - S, LTA - SA and LTA - LT.

Eye-Tracker Data

The eye-tracker output data were converted from the native Eyelink Data File (.EDF) to ASCII text format and were then read into MATLAB. These data consisted of the raw pupil size time series, blink start and end times, saccade start and end times, fixation start and end times and the associated gaze locations on the display screen. The first data pre-processing step involved removing invalid pupil size samples (non-positive values, created by loss of eye target, eyelid occlusion or blinks). The remaining pupil size samples were subjected to additional criteria for removal, including dilation speed outliers and edge artifacts, trend-line deviation outliers, and samples that were temporally isolated (Kret and Sjak-Shie, 2019). Dilation speed outliers are samples that have a disproportionately large absolute pupil size change relative to adjacent samples, likely due to artifacts. These were detected based on the median absolute deviation calculated from the dilation speed. Trend-line deviation outliers are samples that deviate greatly from the signal trend line, which was generated through interpolation and data smoothing. Temporally isolated samples are likely the result of noise or system error (e.g., pupil detection when eyes are shut or not properly tracked) (Kret and Sjak-Shie, 2019).

A temporal threshold of 100 ms was subsequently used in the eye-tracking data to distinguish between fixations and saccades. This threshold is suitable for tasks that involve visual processing of complex geometric stimuli, whereas highly automatic processing (e.g., of the human face) typically involves longer fixations (Manor and Gordon, 2003). Fixations were grouped by tasks, then further grouped by areas of interest (road, non-road) within each task. Percentage of time spent in fixation (i.e., percent fixation duration) was measured as the total duration of fixations from the time of trial onset to the time the participant turned, divided by this time duration and multiplied by 100%. The spatial distribution of percent fixation duration was determined for each participant by the additive superposition of Gaussians, centered at each fixation location and weighted by the fixation duration. The distribution maps for each of the L, LT and LTA conditions were then averaged over all participants, and depicted as colored contour lines overlaid on the driving simulation screen. A Gaussian function was also fit to the distributions for each participant, resulting in estimated values for the percent fixation duration peak and width for each task condition. The sampling distributions of the peaks and widths were tested for normality using the Kolmogorov-Smirnov test using SPSS, then tested for differences using paired samples t-tests and 95% Confidence Interval (CI) testing by bootstrapping.

Saccades that did not meet the minimal saccadic duration of 10 ms were removed, and the percent saccade duration was measured as the total duration of saccades over the length of the trial (Nij Bijvank et al., 2018). The saccade amplitude, peak velocity, and number of performed saccades were also compared for each task.

The average percentage duration of each eye behavior, including saccades, blinks and fixations, was also plotted on a stacked bar graph for each left turn condition from the time of trial onset to the time of turning. Fixations were categorized as "road" (including motor vehicles), and non-road (including the horizon and interior locations of the driven car, such as the rear-view mirror and speedometer). In separate analyses, the effect of blink rate and pupil size was investigated over the trial duration, measured as the time of trial onset to offset. Blink rate was calculated for each task as the number of recorded blinks for each participant divided by the total trial duration (reported as blinks per minute). Similarly, the mean pupil dilation size was calculated as the temporally averaged pupil size for each task condition.

Relationships Between BOLD Signal and Eye-Tracking Metrics

To investigate the relationships between the BOLD signal and eye-tracking metrics, the mean BOLD signal intensity was first calculated in key brain regions that have been associated with eye movements, attention and cognitive processing (inferior frontal, inferior occipital, medial frontal, middle frontal, middle occipital, orbitofrontal, superior frontal, superior occipital, superior temporal and thalamic regions). The brain regions of interest were masked and the time series were averaged over all voxels in a given mask region. The mean BOLD signal intensity was then calculated for each left turn task (L, LT, and LTA). The differences in BOLD signal from the LT condition were then calculated as a change score for each region of interest (i.e., mean BOLD signal intensity of L – mean BOLD signal intensity of LT, and mean BOLD signal intensity of LTA – mean BOLD signal intensity of LT). Change scores for blink rate and pupil dilation were calculated in the analogous manner. Scatter plots were then generated to compare the change scores for the BOLD signal intensity and each eye-tracking metric to investigate the relationships based on the Pearson correlation value.

RESULTS

Driving Behavior

Descriptive statistics for the behavioral metrics of simulated driving performance, as obtained from the STISIM software and further calculated in MATLAB, are shown in **Table 1** for the S, SA, R, L, LT, and LTA conditions. Participants followed the rules of the road and were able to complete all the driving conditions "safely": a total of 2 collisions was observed across all participants for each of R, LT and LTA conditions, whereas only 1 collision was observed for the L condition (The metric was not relevant for the S and SA conditions). Moreover, a number of the behavioral metrics were not significantly different across the driving conditions. This included the time to turn

TABLE 1 | Descriptive statistics of behavioral metrics of simulated driving performance: turn speed, time to turn and number of collisions in the turning tasks (R, L, LT, and LTA), driving speed and lane position in straight driving tasks (S, SA), and response time and response accuracy in audio tasks (SA, LTA).

Condition	Turn speed				Time to turn				Collisions
	Mean (km/hr)	SD (km/hr)	95% CI		Mean (s)	SD (s)	95% CI		total count
			LB	UB			LB	UB	
R	30.5	5.9	26.8	34.3	10.4	1.3	9.6	11.3	2
L	37.9	4.8	34.9	40.9	10.2	1.2	9.4	11.0	1
LT	35.6	3.4	33.4	37.7	9.9	2.5	8.3	11.5	2
LTA	34.3	3.0	32.4	36.3	10.5	1.7	9.4	11.6	2

	Driving speed				Lane position			
	Mean (km/hr)	SD (km/hr)	95% CI		Mean (m)	SD (m)	95% CI	
			LB	UB			LB	UB
S	62.5	2.0	61.2	63.8	2.2	0.4	1.9	2.5
SA	60.4	0.8	59.9	60.9	2.3	0.3	2.1	2.5

	Response time				Response accuracy			
	Mean (s)	SD (s)	95% CI		Mean (%)	SD (%)	95% CI	
			LB	UB			LB	UB
SA	4.1	0.7	3.6	4.6	83.3	20.1	70.6	96.1
LTA	4.1	1.3	3.3	5.0	77.4	14.2	68.3	86.4

SD, standard deviation; CI, Confidence Interval; LB, Lower Bound; UB, Upper Bound. R, right turns; L, left turns; LT, left turns with oncoming traffic; LTA, left turns with oncoming traffic and auditory distraction; S, straight driving; SA, straight driving with auditory distraction.

(Repeated Measures ANOVA; $p > 0.05$), and lane position for the S and SA conditions (paired samples *t*-test; $p > 0.05$). However, a paired samples t-test showed that driving speed (reported as the mean (SD), where SD is the standard deviation) for the SA condition (60.4 (0.8) km/hr) was significantly slower than that of S condition (62.5 (2.0) km/hr). In addition, there was a significant effect of turning condition on turn speed (Repeated Measures ANOVA; $p < 0.05$; **Figure 2B**). Bonferroni-corrected *post hoc* contrasts showed no significant differences in the turn speed between L and LT conditions, and between R and LTA conditions. The turn speed was significantly slower for the R condition (30.5 (5.9) km/hr) than for the L condition (37.9 (4.8) km/hr), and similar contrast was found between the R and LT condition (35.6 (3.4) km/hr) (corrected $p < 0.05$). In addition, the turn speed for the LTA condition (34.3 (3.1) km/hr) was significantly slower than for the L condition (corrected $p < 0.05$).

Regarding the auditory task performance during distracted simulated driving, paired samples t-tests did not reveal significant differences between the SA and the LTA conditions, neither for question response time, nor response accuracy ($p > 0.05$). Notably, the mean (SD) of the response time was 4.1 (1.0) s over both conditions, whereas the time to turn in the LTA task was 10.5 (1.7) s. Because the time to turn was not significantly elevated for the LTA condition in comparison to the other turn conditions and the turn speed was slower for the LTA condition than for the L condition, this suggests that participants were dual-tasking rather than performing the two tasks sequentially during the LTA condition. If the tasks were performed sequentially, then a time to turn of approximately 10.5 + 4.1 = 14.6 s would have been expected for the LTA task. As this was not observed, it can be inferred that the tasks were performed in parallel.

Brain Activity

Regular Driving

Brain activations are reported relative to the S control condition. The R (**Figure 2A**) and L (**Figure 2B**) turning conditions without traffic showed similar activation maps ($p < 0.001$), and contrast maps between the two conditions were not significantly different. Both R and L conditions showed significant positive bilateral activation in the occipital (lingual gyrus, middle occipital), inferior parietal, primary motor, primary somatosensory and premotor cortices. Activation was also observed in the precuneus, cerebellum and basal ganglia.

The addition of oncoming traffic to left turns (LT, **Figure 2C**) produced activation maps ($p < 0.001$) that were very similar to those for the R and L conditions. Although negative activations were slightly more extensive in the frontal (left inferior, left and right medial, left middle, left superior) cortex and the anterior cingulate cortex, the contrast maps between the regular driving conditions (LT – L, LT – R) did not show statistically significant differences ($p > 0.01$).

Distracted Driving

When evaluating the differences in brain activation between distracted straight driving and straight driving conditions, the

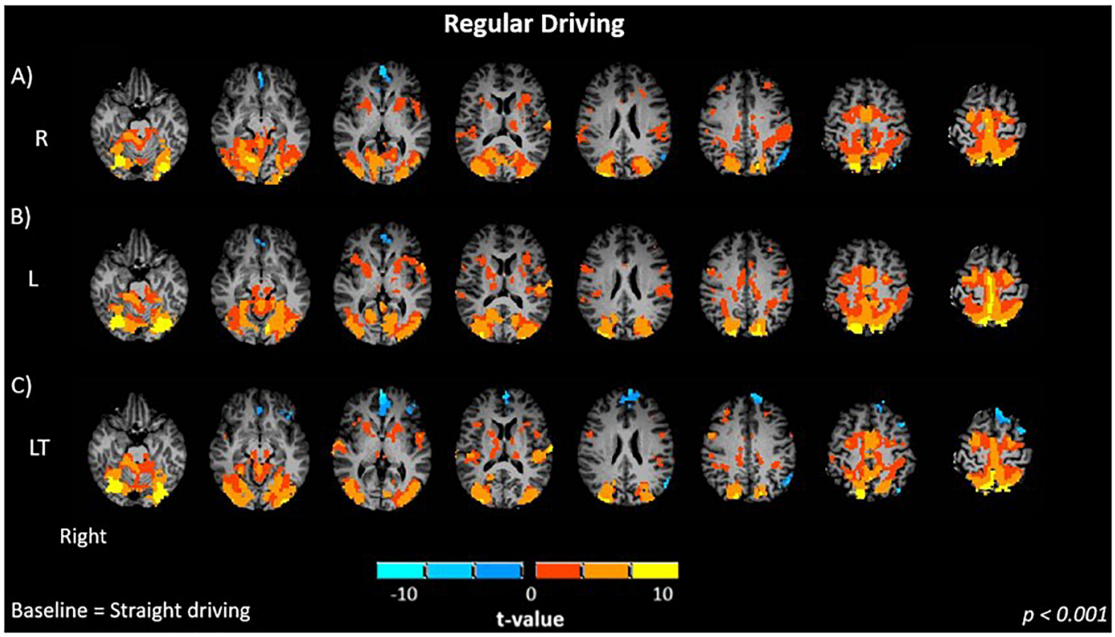

FIGURE 2 | Brain activations of participants when performing regular driving tasks: **(A)** R, right turns; **(B)** L, left turns; **(C)** LT, left turns with oncoming traffic. Activations are shown as *t* contrasts in comparison with straight driving, and are reported with a statistical significance threshold of $p < 0.001$. "Right" indicates the right cerebral hemisphere.

(SA – S) contrast showed significant positive activations in the temporal (left and right superior, right middle) and frontal (left and right inferior) cortices and the anterior cingulate cortex ($p < 0.001$, **Figure 3A**). No significant activation was observed in the occipital, parietal and somatosensory areas. The (LTA – S) contrast was evaluated in analogous fashion. In addition to activations in occipital, parietal, motor, somatosensory and cerebellar regions observed for the L and LT conditions, the LTA condition showed greater positive activations in relation to the S condition in frontal (right middle, left and right inferior) and temporal (left and right middle and superior) cortices, anterior cingulate cortex and insula ($p < 0.001$, **Figure 3B**).

Next, the difference maps between LTA – SA and LTA – LT are shown in **Figure 4**. These two contrasts are of interest as they serve to illustrate the differing effects of distraction response for different driving demands. Compared to SA, LTA showed significantly greater activation in the left and right superior temporal gyrus ($p < 0.001$, **Figure 4A**). Significant increases in activation were also seen in the occipital (middle, superior); the primary motor, somatosensory and premotor cortices; as well as the precuneus and cerebellum. Significant decreases in activation were seen in the left inferior frontal gyrus. Compared to LT, LTA showed significantly greater activation in the frontal (left and right inferior, left superior, left medial, left middle), temporal (left and right middle, left and right superior) cortices and precuneus ($p < 0.001$, **Figure 4B**). Significant decreases were observed in the right occipital (middle, superior), right inferior parietal, and cerebellar areas.

The activations shown in **Figures 2–4** are subsequently listed in **Table 2**, which provides the *t*-value of peak activation, the location of each peak in Talairach and Tourneaux (TT) atlas coordinates and the associated brain region in the TT atlas.

Oculomotor Measures

Contour lines representing the spatial distribution of percent fixation duration for the L, LT and LTA conditions are shown in **Figure 5**, overlaid on the driving simulation environment. The distribution map for the L condition showed the widest span of fixation across the screen, including peripheral areas (mirror and speedometer) and the left side of the screen. The LT condition showed less fixation on the left side of the screen, as participants directed their gaze more to the oncoming traffic. The addition of auditory distraction in the LTA condition further narrowed the percent fixation duration onto the road immediately ahead.

Statistical analyses were subsequently conducted to test for the effects of driving condition on the spatial distribution of percent fixation duration. The sampling distributions of the peak amplitudes and widths of the percent fixation duration were normal for the L, LT, and LTA conditions (Kolmogorov-Smirnov Test, $p < 0.05$). No significant differences were found in the percent fixation duration distribution widths between these conditions, and no significant differences were found in the peak amplitudes of the fixation distribution between LT and L. However, the 95% CI testing of the mean difference of estimated peak amplitudes between "LTA – L" and "LTA – LT" rejected the null hypothesis. The peak amplitude for the LTA condition (56.8 (7.8)%) was found to be significantly greater than the values for both the LT condition (27.4 (4.5)%) and the L condition (25.4 (3.4)%) (95% CI Testing, Paired Samples *T*-Test, $p < 0.05$). The interpretation of this effect is that during the LTA condition,

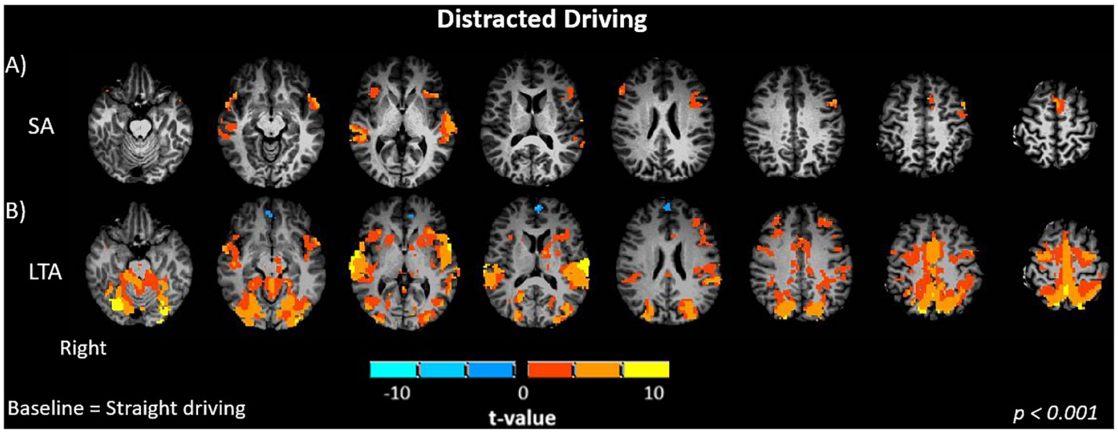

FIGURE 3 | Brain activations of participants when performing driving tasks while distracted. **(A)** SA, straight driving with audio task; **(B)** LTA, left turns with oncoming traffic and audio. Activations are shown as t contrasts in comparison with straight driving, and are reported with a statistical significance threshold of $p < 0.001$. "Right" indicates the right cerebral hemisphere.

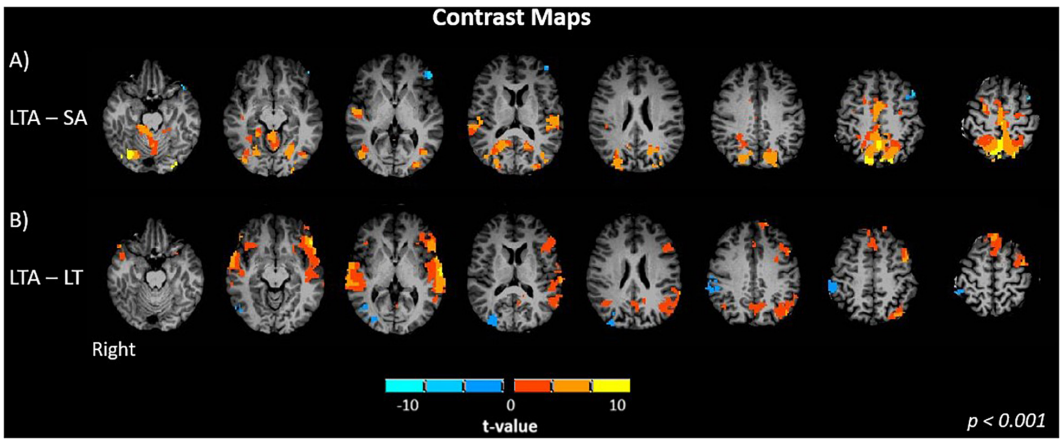

FIGURE 4 | Brain activation maps of contrast between task factors: **(A)** LTA – SA, Left turns with oncoming traffic and audio vs. straight driving with audio; and **(B)** LTA – L, left turns with oncoming traffic and audio vs. left turns with oncoming traffic without audio. Activations are shown with a statistical significance threshold of $p < 0.001$. "Right" indicates the right cerebral hemisphere.

the spatial distribution of percent fixation duration for each participant was more consistent across the group than during the other two conditions.

Next, the average percentage of time spent performing each different eye behavior (fixations, saccades, blinks) was plotted in a stacked bar graph for each left turn task condition (**Figure 6**). No significant differences were found for the time spent fixating on the road (including motor vehicles) between LTA and LT conditions (Repeated Measures ANOVA; $p > 0.05$), whereas a significant elevation was found for LT compared to L (mean difference = 7.2%, Repeated Measures ANOVA; $p < 0.05$). Time spent fixating on non-road locations showed decreases with increasing task complexity, with mean differences of −7.3% and −3.3% for LTA compared to L and LTA compared to LT, respectively (Repeated Measures ANOVA; $p < 0.05$). The percentage of time performing saccades was significantly increased for LTA compared to LT and compared to L (mean difference = 5.4%, 5.0% respectively, Repeated Measures ANOVA; $p < 0.05$). No significant differences were observed between the peak velocity and number of saccades performed across the three tasks, however the mean saccade amplitude for the LTA condition (2.6 (0.28) units) was significantly less than that for the L condition (3.3 (0.35) units) (Repeated Measures ANOVA; $p < 0.05$).

In a separate analysis of eye-blink behavior (i.e., not reported in **Figure 6**), the mean blink rate over the total trial duration for the LTA condition [34.8 blinks/min (3.8)] was significantly lower than that for the L condition (37.0 blinks/min (5.3), Repeated Measures ANOVA; $p < 0.05$). No significant differences were found for the mean pupil dilation size for L, LT and LTA, although an increasing trend was observed across the conditions.

Relationships Between Eye-Tracking Metrics and BOLD Signals

No significant relationships were found between the mean BOLD signal intensity and pupil dilation size. However,

TABLE 2 | Peak activation (*t*-value), spatial coordinates (Left, Posterior, Inferior) and brain region of statistically significant brain activity with respect to the Talairach and Tournoux (TT) brain atlas.

Peak activation intensity (*t*-value)	Atlas coordinates			TT Atlas region
	L	P	I	
R - S				
16.45	9	59	64	Left Precuneus, BA 7
14.20	11	34	71	Left Postcentral Gyrus, BA 3
9.89	−26	61	−14	Right Declive
7.48	39	64	−14	Left Fusiform Gyrus, BA 19
7.06	29	66	−14	Left Declive
7.03	−1	1	49	Right Medial Frontal Gyrus, BA 6
6.39	−11	69	6	Right Cuneus, BA 18
6.04	−11	76	4	Right Lingual Gyrus, BA 18
5.93	−21	6	61	Right Middle Frontal Gyrus, BA 6
5.03	9	69	11	Left Cuneus, BA 30
5.02	9	94	-9	Left Lingual Gyrus, BA 17
L - S				
12.67	−14	79	41	Right Precuneus, BA 7
10.10	26	74	−16	Left Declive
9.72	−16	81	36	Right Cuneus, BA 19
9.51	−24	66	−14	Right Declive
8.14	−1	14	54	Right Medial Frontal Gyrus, BA 6
7.86	1	26	56	Left Medial Frontal Gyrus, BA 6
6.21	26	49	−11	Left Culmen, BA 37
6.11	29	84	6	Left Middle Occipital Gyrus, BA 18
5.94	−19	59	16	Right Posterior Cingulate, BA 30
5.85	−24	49	−11	Right Culmen, BA 37
5.75	−1	41	−9	Right Cerebellar Lingual
LT - S				
11.38	−14	79	41	Right Precuneus, BA 7
9.51	−26	66	−16	Right Declive
8.60	26	76	−16	Left Declive
6.62	−4	1	46	Right Cingulate Gyrus, BA 24
6.34	−1	11	49	Right Medial Frontal Gyrus
6.08	1	76	−1	Left Lingual Gyrus, BA 18
6.06	6	44	49	Left Precuneus, BA 7
5.82	16	36	−39	Left Cerebellar Tonsil
5.66	−24	46	−11	Right Culmen, BA 37
5.38	44	26	11	Left Transverse Temporal Gyrus, BA 41
5.07	−11	46	−46	Right Cerebellar Tonsil
5.02	−19	59	19	Right Posterior Cingulate
SA - S				
4.48	1	−9	51	Left Superior Frontal Gyrus, BA 6
4.33	−49	31	1	Right Middle Temporal Gyrus, BA 21
3.71	−49	39	6	Right Superior Temporal Gyrus, BA 22
LTA - S				
9.39	−29	66	−16	Right Declive
8.28	−1	56	49	Right Precuneus, BA 7
7.13	16	79	36	Left Precuneus, BA 7
7.05	29	66	−14	Left Declive
6.92	−51	21	11	Right Transverse Temporal Gyrus, BA 41
6.47	−1	6	44	Right Cingulate Gyrus, BA 24
6.32	−1	11	51	Right Medial Frontal Gyrus, BA 6
5.36	1	-11	39	Left Cingulate Gyrus, BA 24, 32
LTA - SA				
10.18	−1	56	49	Right Precuneus, BA 7
7.67	−29	69	−14	Right Declive
6.60	−19	56	19	Right Posterior Cingulate
6.46	−21	26	−21	Right Culmen
5.91	−1	11	51	Right Medial Frontal Gyrus, BA 6
5.53	46	24	11	Left Transverse Temporal Gyrus, BA 41
5.42	11	44	−44	Left Cerebellar Tonsil
LTA - LT				
9.23	51	−19	−1	Left Inferior Frontal Gyrus, BA 47
6.61	64	29	6	Left Superior Temporal Gyrus, BA 42
4.12	44	−4	44	Left Middle Frontal Gyrus, BA 6
3.83	54	51	9	Left Middle Temporal Gyrus, BA 22
3.67	14	−46	41	Left Superior Frontal Gyrus, BA 8
3.25	−46	36	4	Right Superior Temporal Gyrus, BA 22
1.71	−36	56	24	Right Middle Temporal Gyrus, BA 39

BA, Brodmann Area; S, straight driving; SA, straight driving with auditory distraction; L, left turns; R, right turns; LT, left turns with oncoming traffic; LTA, left turns with oncoming traffic and auditory distraction. Atlas coordinates L, left; P, posterior; I, inferior. Not all regions of significant activation are reported, with a cutoff value of t = 5 (unless there were less than 3 regions greater than this cutoff). Each peak within a given activation zone is only reported once.

scatter plots showing the change scores for the mean BOLD signal intensity and blink rate (L – LT, LTA – LT) showed negative, statistically significant Pearson correlations in the insula [mean slope of −0.13 (0.05)], middle frontal [−0.01 (0.04)] and superior temporal [−0.03 (0.07)] regions, as shown in **Figure 7**. A positive, statistically significant Pearson correlation was found in the middle occipital gyrus [0.02 (0.02)].

DISCUSSION

The present study investigated simulated driving using the novel combination of fMRI and eye-tracking. As the study was intended to investigate whether it was possible to replicate the fMRI findings of Schweizer et al. (2013) using the same task design, comparisons to this prior work are made throughout. In addition, given that driving requires considerable visual processing demands, the study extends beyond Schweizer et al. (2013) to include eye-tracking for an initial evaluation of oculomotor measures for each task, and their relationship with simulated driving task-related behavior and brain activity. The results of the study are in good agreement with the stated hypotheses, and are discussed below for the behavioral metrics of the driving simulator performance, brain activity and oculomotor measures. The inter-relationships between these measures are then discussed, followed by the limitations of the study.

Behavioral Metrics of Simulated Driving

The behavioral metrics of simulated driving performance (**Table 1**) indicated that the R, L, LT, and LTA conditions were performed well by participants for the most part, while they obeyed the "rules of the road." Over all participants and fMRI runs, only two collisions were observed for both the LT and LTA conditions, whereas only one collision was observed for the L condition. Only one participant was involved in multiple collisions (one in each of the R, LT, and LTA conditions). The left turn conditions were performed similarly by all participants on the time to turn metric, and the turn speed was similar in the L and LT conditions. In addition, the S and SA conditions were performed without differences in lane position. These results suggest that each of the simulated driving conditions, while containing variations in the complexity of task, were well within the skill level for efficient execution by the participants without large changes in the behavioral metrics studied. It is important to emphasize that these results are specific to young healthy drivers, and analogous results for other populations of interest, such as elderly drivers or individuals with cognitive impairments, are likely to be different.

Two noteworthy differences in simulated driving performance were found between certain conditions. First, the turn speed was significantly slower for the R condition than for the L, LT and LTA conditions. This may have resulted from the smaller turning radius of the road when turning right, resulting in slower turn speeds to drive safely and maintain lane position. However, the R and L conditions showed similar brain activation maps (see below), suggesting that both turns require similar behavioral elements. Second, there were slight but significant reductions in driving speed associated with distracted driving. The effect was slightly more pronounced when comparing the SA and S conditions, whereas the slowing in the LTA condition was only significant in comparison to the L condition, not the LT condition. These results suggest that the primary task (driving) was slightly impacted by the additional cognitive demand associated with responding to the secondary task (auditory questions) in both cases. From the auditory task performance, there was no evidence that participants were making trade-offs between how well to perform the driving tasks in comparison to the auditory task – as the question response time and response accuracy was not statistically significant between the LTA and SA conditions. This may be due to the age of the participants, or the auditory tasks that were administered (general knowledge true or false questions), as they were not very demanding. This effect on driving speed could be investigated in more detail in the future by conducting a behavioral study with more participants, and by manipulating the strength of the auditory distraction. Certain behavioral data related to the LTA condition were also noteworthy: the question response time was substantially shorter than the time to turn, and the time to turn was not significantly different from that of the LT condition. As described in the Results, these findings suggest that participants engaged in driving-auditory dual-tasking in the LTA condition rather than performing the two tasks sequentially.

The results observed in this study were similar to those previously reported by Schweizer et al. (2013), however some slight discrepancies were also observed. In both studies, no significant differences were found between the lane position in straight driving and distracted straight driving. The average response accuracy to the audio distraction task was also similar, and the observed turn speed of right turns was slower than all left turns (L, LT and LTA). In this study, the turn speed of the L condition was faster than both LT and LTA, whereas Schweizer et al. (2013) reported the turn speed of the L condition was slower than both LT and LTA. Additionally, no significant differences were previously observed in the average speed during straight driving and distracted straight driving. It is reasonable to speculate that these discrepancies were observed due to the different cohorts investigated in the two studies, the relatively low sample size and the variability in human behavior. None of the participants volunteered in both studies. Although the two studies were conducted using slightly different versions of fMRI-compatible driving simulator, one-hour simulator training periods were part of the experimental design in both cases to control for systematic bias. Irrespective of the discrepancies, the simulated driving behavior in the present study helps to inform the interpretation of the task-related brain activation maps, as discussed below.

Brain Activity

Previous simulated driving studies have shown a network of brain regions activated during straight driving, including the occipital and parietal lobes, cerebellum, and areas involved in motor control and somatosensory integration (Calhoun et al., 2002; Uchiyama et al., 2003; Graydon et al., 2004; Spiers and Maguire, 2007; Just et al., 2008; Callan et al., 2009; Crundall and Underwood, 2011; Calhoun and Pearlson, 2012; Li et al., 2012; Schweizer et al., 2013; Hyung-Sik et al., 2014; Choi et al., 2017; Ware et al., 2020). Straight driving involves visual vigilance and small adjustments in driving control to maintain speed and lane position. To isolate the visual, motor and cognitive functions activated in more demanding driving tasks (e.g., performing turns), straight driving was chosen as the baseline condition. As such, all the subsequent activation maps of the task conditions must be thought of as referenced to this baseline behavior, and

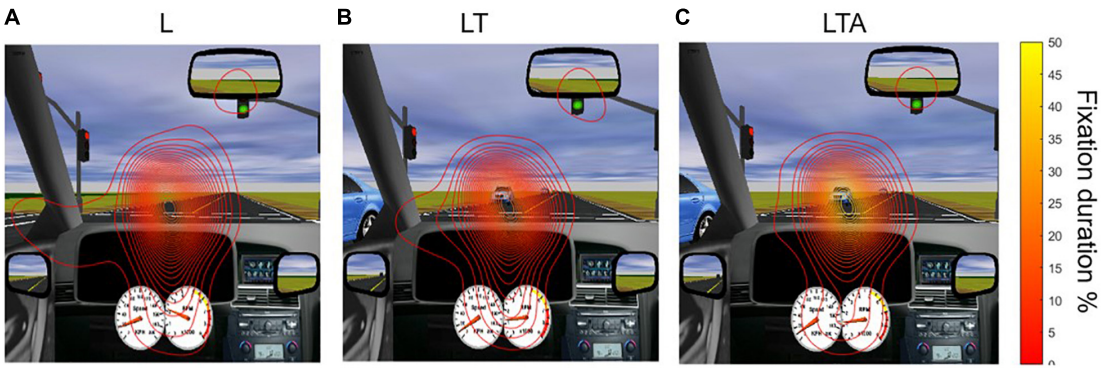

FIGURE 5 | Distribution of gaze in the various tasks, averaged over all participants. Maps of percent fixation duration were computed as described in the text, and are shown as contour lines overlaid on the driving simulation screen. **(A)** L, left turn condition; **(B)** LT, left turn with traffic condition; and **(C)** LTA, left turns with traffic and auditory distraction.

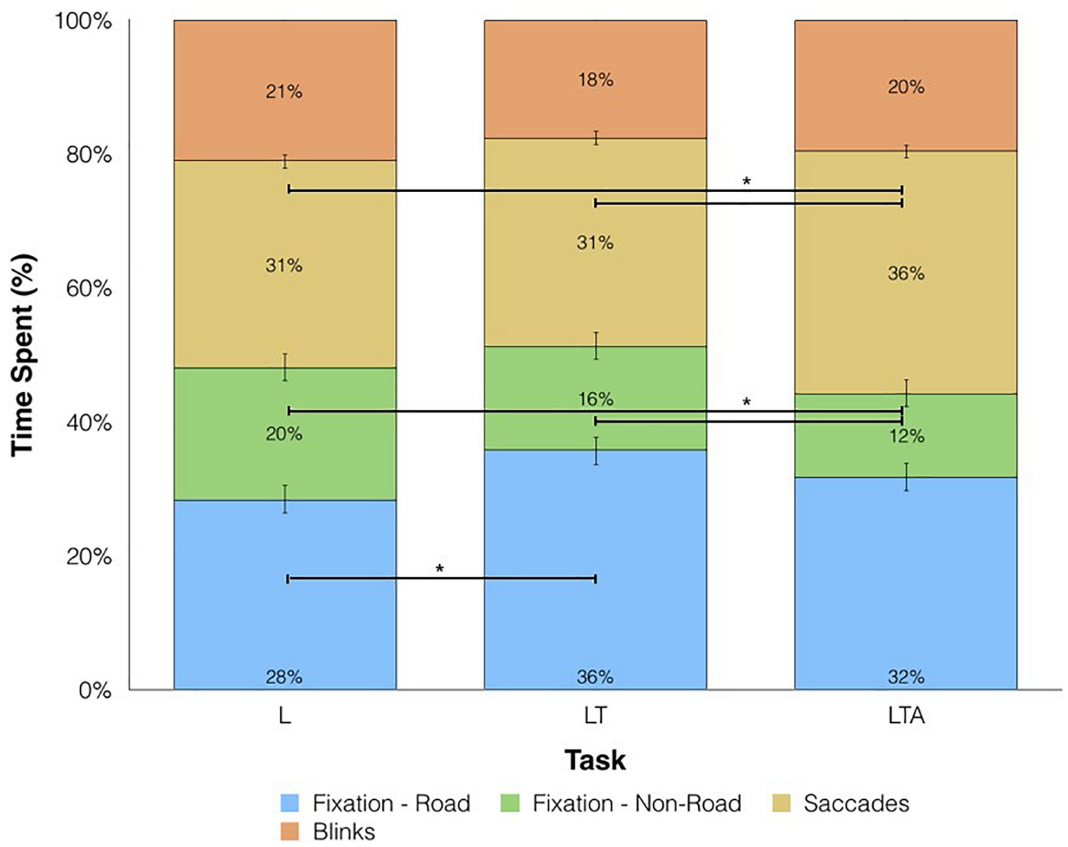

FIGURE 6 | Percentage of time spent on eye movements for L, left turns; LT, left turns with oncoming traffic; and LTA, left turns with oncoming traffic and auditory distraction. Fixations on the road are shown in blue, fixations away from the road are shown in green, saccades are shown in yellow and blinks are shown in orange. Horizontal bars and asterisks represent statistically significant differences at $p < 0.05$.

represent an augmentation of this supposedly least-taxing state of driving and associated brain activity.

In this study, occipital, parietal, cerebellar and motor areas were recruited in both the R and L conditions, and also the LT condition (**Figure 2**). The brain regions that were found to be activated (occipital, parietal, cerebellar and motor areas) were consistent with previous studies investigating simple driving as noted above (Calhoun et al., 2002; Uchiyama et al., 2003; Graydon et al., 2004; Spiers and Maguire, 2007; Callan et al., 2009; Crundall and Underwood, 2011; Calhoun and Pearlson, 2012; Li et al., 2012; Hyung-Sik et al., 2014; Choi et al., 2017; Ware et al., 2020). Some additional aspects are noteworthy. First, the driving

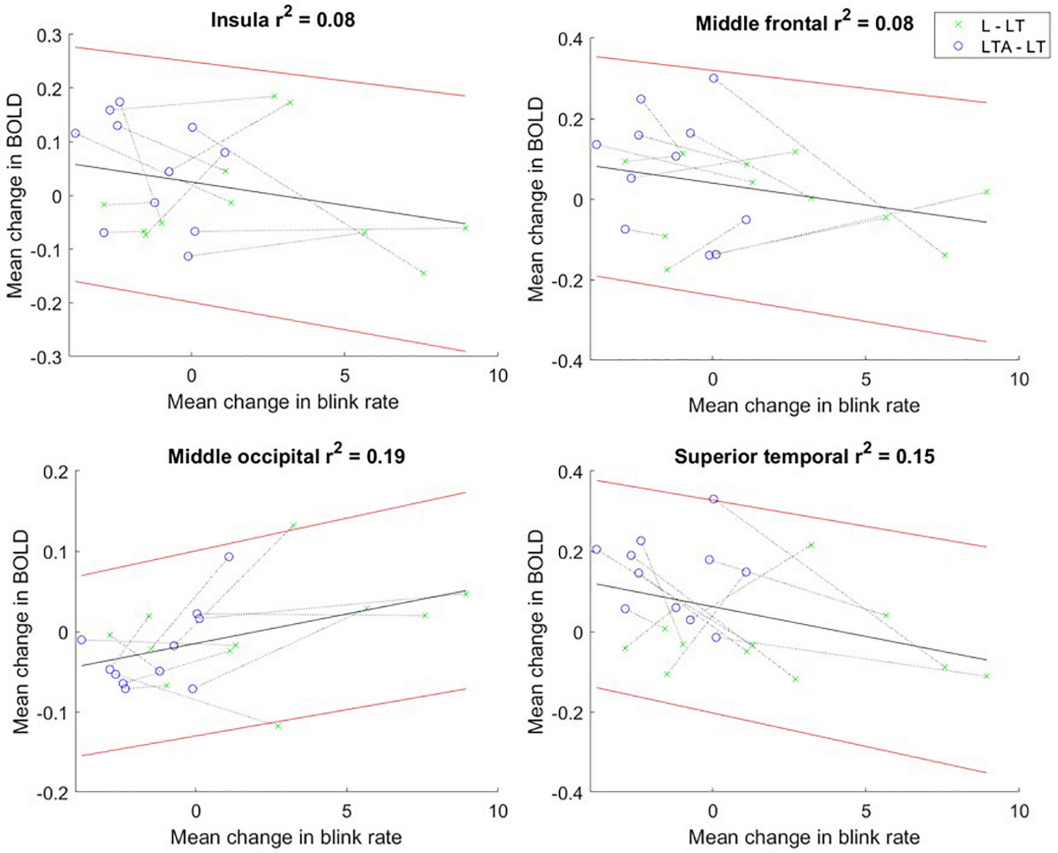

FIGURE 7 | Mean change in BOLD fMRI signal vs. mean change in blink rate for "L – LT" (green) and "LTA – LT" (blue) task contrasts for the insula, and regions of interest in the middle frontal, middle occipital and superior temporal gyri. The dotted line connects the data points for each participant. The linear regression line is shown in black, with the 95% confidence interval for the regression line shown in red. BOLD, blood oxygenation level dependent; L, left turns; LT, left turns with oncoming traffic; LTA, left turns with oncoming traffic and auditory distraction.

conditions showed decreased activity in the medial frontal cortex and anterior cingulate cortex (ACC). Both regions are part of the default mode network (DMN), which decreases in activity when performing tasks (Raichle, 2015). Thus, it is speculated that the more engaging R, L, and LT conditions produced more suppression of the DMN than occurred during straight driving. Second, minimal activity was observed in the rest of the prefrontal cortex, suggesting that the R, L, and LT conditions were performed by the participants as a consequence of extensive training and practice in driving maneuvers, leading to efficient mental processing controlled primarily by the posterior brain, largely in the absence of executive control. Third, no significant differences were found in the brain activation between R, L and LT conditions. The lack of difference in brain activation between R and L conditions is expected because the task requirements are largely the same, except that the direction of simulated movement is different. Additionally, the lack of difference in brain activation between the LT and L or R conditions is consistent with the behavioral metrics, which showed that each condition produced a similar time of turning. This suggests that the inclusion of traffic to left turns was not very challenging for this cohort of young but experienced drivers.

The present results are similar to but slightly different from those reported by Schweizer et al. (2013), in which the R condition produced minimal significant activations, including the somatosensory association, parietal and visual cortices, whereas the L condition showed greater activation in the posterior regions (occipital, parietal, motor, cerebellar), and the LT condition showed a progressive increase in the extent of activation in the posterior network and cingulate cortex. Two methodological elements are noteworthy in discussing the findings of this prior study in relation to the present work. First, the differences in brain activation between conditions were assessed visually and qualitatively in the prior study, whereas the present study used a quantitative assessment of statistical contrasts. Thus, it is possible that effects interpreted between conditions in the previous study were not statistically significant. Second, as previously discussed in section 4.1 above, the two studies consist of different samples of young healthy adults. Although the two studies were conducted using slightly different versions of 3T Siemens MRI system, systematic differences arising from this effect are anticipated to be minor when comparing task-related brain activity at the group level.

Shifting to the effects of distraction on brain activity, brain activations were observed in the middle and superior temporal gyrus (MTG, STG) when the SA – S contrast was evaluated (**Figure 3A**). These regions have previously been associated with cognitive processes including language and semantic processing, spatial awareness, and attentional processing from the dorsal and ventral streams (Chao et al., 1999; Tranel et al., 1997; Onitsuka et al., 2004). Similar effects were observed for the LTA – S contrast (**Figure 3B**). Furthermore, the LTA – SA contrast showed activation of the STG, occipital, primary motor and premotor cortices, whereas significant decreases in activations were observed in the left inferior frontal gyrus (IFG) (**Figure 4A**). The activations in the occipital, primary motor and premotor cortices are consistent with the increased visuomotor demands associated with performing left turns with oncoming traffic in the LTA condition in relation to the straight driving required in the SA condition. Greater activation in the STG suggests a higher demand for cognitive processing in the LTA condition in relation to the SA condition. In addition, the LTA condition showed less activation of the left IFG compared to the SA condition. The left IFG has been linked to language processing and working memory, and has been found to be activated during semantic tasks (Petersen et al., 1989; Démonet et al., 1992; Moss et al., 2005; Liakakis et al., 2011). Two possible mechanisms may be speculated for the observed decrease in activation, although future work with a larger sample size will be required to determine which is correct. From **Table 1**, there was a trend for lower response accuracy for the auditory task in the LTA condition compared to the SA condition, but without statistical significance. Considering this trend as a "false positive," then the difference in IFG activity (and inferred resources for language processing) should disappear with larger sample size. Alternatively, if the trend is thought of as a "false negative," then both the difference in response accuracy and IFG activity should persist in a study with larger sample size. As the task instructions were not specific about prioritizing the auditory task, it is conceivable that participants performed as they would in the real world, to trade off auditory task performance to ensure safe driving behavior in the more cognitively demanding task (LTA compared to SA). The observed trend in response accuracy is consistent with this interpretation, although no more definite statements can be made at present.

Distraction poses a great risk to drivers especially when they perform left turns. Of the MVCs that occur while drivers turn or cross an intersection, 61% involve turning in the left direction (U.S. Department of Transportation National Highway Traffic Safety Administration, 2008). Compared to L or LT conditions, the LTA condition showed significantly greater bilateral activation in the middle frontal gyrus (MFG), IFG, SFG, MTG, STG and ACC (**Figure 4B**). In particular, the increase in frontal activity is characteristic of dual-tasking (Dux et al., 2006) and is consistent with the hypotheses of the study. The right MFG and IFG have been associated with the ventral attention network, activated in response to unexpected stimuli and to attend to relevant task information and distractors (Friedman-Hill et al., 2003; Hahn et al., 2006; Buschman and Miller, 2007; Fan et al., 2008; Neufang et al., 2011; Weissman and Prado, 2012; Japee et al., 2015). The IFG has also been suggested to be play a role in sustained attention, attentional control, working memory and motor response inhibition (Swick et al., 2008; Hampshire et al., 2010; Tops and Boksem, 2011). The activation of this region suggests the involvement of attentional control and reorientation when a secondary task was introduced. The left SFG is known to play a role in working memory and executive processing, suggesting that recruitment of this region is necessary for performance of the auditory task (Boisgueheneuc et al., 2006). Recruitment of the MTG and STG suggests involvement of spatial awareness and attentional processing to undertake executive functions and additional task processing, consistent with previous findings (Schweizer et al., 2013). The ACC has been linked to cognitive modulation and error processing (Bush et al., 2000; Paus, 2001; Fan et al., 2008; Orr and Hester, 2012). Greater activation in the ACC suggests that the secondary auditory task increased the overall cognitive load, and thus the need to monitor and correct for errors.

In addition to the regions of increased activation in the LTA condition compared to the LT or L conditions, significantly decreased activation was observed in the right middle occipital gyrus (MOG), superior occipital gyrus (SOG), right inferior parietal lobe (IPL), cerebellum and precuneus. This observation is also in line with the study hypotheses. The right MOG and SOG are considered to be part of the visual dorsal stream, associated with spatial and visuomotor processing (Dumoulin et al., 2000; Wandell et al., 2007; Matthys et al., 2009; Renier et al., 2010), and the right IPL is involved in sustained visual and somatosensory attention (Wilkins et al., 1987; Pardo et al., 1991; Rueckert and Grafman, 1996, 1998). These results are in agreement with those of previous studies that investigated brain activity in simulated driving with a relatively simple secondary behavioral task, showing a shift in activation from the posterior, visual areas to the frontal areas responsible for planning, decision-making and cognition (Just et al., 2008; Schweizer et al., 2013; Choi et al., 2017).

Oculomotor Measures

Investigating the same driving tasks as in the present work, Schweizer et al. (2013) stated that "the distracted brain sacrificed areas in the posterior brain important for visual attention and alertness to recruit enough brain resources to perform a secondary cognitive task." However, the inclusion of oculomotor measures in the present study, as subsequently described below, suggest an additional, alternative interpretation. Given the task instructions, participants modified their gaze behavior to maintain the ability to make left turns in the presence of oncoming traffic. As the task was performed in a simulator under very simple conditions, there were no significant task penalties associated with not gazing at off-road areas such as the dash-board and rear-view mirror. Instead, in the presence of distraction there were performance advantages to directing gaze much more consistently at the oncoming traffic to simplify the visual processing demands (and the resulting brain activity). Additional experimental methodology involving augmented simulated driving tasks will be required to evaluate these two interpretations further.

There is a correlation between distraction and increased risk of crashes and near-crashes (2006). The implication of interest here is that one of the effects of distraction, contributing to its role in dangerous driving, is an associated modification in visual processing that can potentially stop drivers from reacting to possible threats. Supporting the modulation of visual processing in the present study, and as hypothesized, it was found that the less demanding left turn task (i.e., L condition) resulted in a wide distribution of gaze points across the scene as quantified using the percent fixation duration map, including peripheral areas (mirror and speedometer) and the left side of the road (**Figure 5**); as the task became more demanding, (i.e., LT condition), the distribution remained similar, although there was an increase in the fraction of time spent fixating on the road at oncoming traffic, as appropriate for the specific task (**Figure 6**). With the introduction of distraction (i.e., the LTA condition), the fixation distribution narrowed: the fixation duration onto the immediate road and oncoming vehicles was increased further, as participants concentrated on the essential features of the visual scene to maintain driving performance. Although differences in the widths of the fixation distribution were not significantly different (LTA – LT, LTA – L), the effect was nevertheless detected as an increase in the mean peak amplitude of the fixation distribution for these contrasts. This suggests that as a group, the participants were gazing more consistently at the oncoming traffic during the LTA condition. Although not specifically studied in these driving tasks, it would be useful in future studies to investigate whether these changes in gaze behavior result in decreased ability to react appropriately to other dynamic changes in the visual scene that would constitute a threat to safe driving - such as cross-traffic failing to obey traffic lights – requiring aversive maneuvers. Furthermore, the current methodology and study findings are not definitive regarding how to quantify attention (and modulations thereof) across the driving task conditions, and this should be addressed in future studies.

Irrespective of the two interpretations discussed above, the oculomotor results are consistent with the brain activation maps showing a decrease in activation in the visual cortex for the LTA – LT contrast (**Figure 4B**). **Figures 5**, **6** confirm that this reduction is accompanied by a narrowed field of view, involving an increasing trend in the percentage of on-road fixations and a significant decrease in the percentage of off-road fixations when comparing LTA to LT. Furthermore, compared to the LT and L conditions, the LTA condition showed less percentage time of overall fixations (i.e., road and non-road) with significantly greater percentage of time spent performing saccades. This observation is consistent with recent eye-tracking research showing that the percentage of time spent in fixation can decrease when visual processing tasks are performed in the presence of auditory distraction tasks (Wu et al., 2019). Control of saccades has been associated with the frontal eye fields (Brodmann Area 8), the dorsolateral prefrontal cortex (DLPFC) and the cingulate eye field within the ACC (Künzle et al., 1976; Vernet et al., 2014). In **Figure 4B**, increased activation was observed in these regions when comparing the LTA condition to the LT condition, consistent with increased saccades in the former. Together with the decrease in fixation on the periphery in the LTA condition, these results further support the interpretations that for the simulated driving tasks investigated, distraction resulted in less visual processing resources (a) made available or (b) required to support driving performance.

No significant differences were found in blink rate for the LT – L contrast. This is consistent with the brain activation maps and behavioral results, which did not show significant differences between the two tasks. As hypothesized, however, distraction resulted in a statistically significant decrease in blink rate for the LTA – L contrast. This finding is consistent with consensus in the literature that blink rate decreases with increased cognitive load (Brookings et al., 1996; Ohira, 1996; Van Orden et al., 2001; Haak et al., 2009; McIntire et al., 2014; Faure et al., 2016; Maffei and Angrilli, 2018). However, one study found an increase in blink rate when a straight driving task was combined with an auditory arithmetic task - suggesting that other factors may influence the effect, such as disinhibition of blink rate effects under certain task conditions and ranges of workload (Tsai et al., 2007). The experimental design of the present study cannot refute this claim.

A trend of increasing pupil diameter was observed across the left turn conditions (L, LT, LTA) suggesting that greater cognitive processing may have occurred across this progression of tasks. Although this effect was small and not statistically significant, the direction of the effect is consistent with the brain activation data, and with the study hypotheses. Pupil dilation in cognitive processing has been linked to activation of the MFG and ACC, measuring attention and effort (Siegle et al., 2003; Alnaes et al., 2014; Murphy et al., 2014; Zekveld et al., 2014; Joshi et al., 2016). This is consistent with the increase in activation of the MFG and ACC observed in the brain activation maps for the LTA - LT contrast (**Figure 4B**).

Relationships Between Eye-Tracking Measures and Bold Signals

The change score relationships for BOLD signals and eye-tracking metrics were invested as r^2 values, representing the proportion of variance of the change in BOLD signal that can be attributed to the change in each eye-tracking metric. As hypothesized, significant Pearson correlations were found between the BOLD signal change score and blink rate change score for several regions. Negative correlations were found for the insula, the STG and the MFG, whereas a positive correlation was found for the MOG. The insula and STG have been described to play a role in the suppression of eye blinking (Lerner et al., 2009; Berman et al., 2012). The MFG, which has been associated with the reorientation of attention, has also been correlated with blink inhibition (Andersson et al., 2009; Mazzone et al., 2010; Neufang et al., 2011; Japee et al., 2015). The increased activation in the insula, STG and MFG is consistent with the suppression of the blink mechanism as the cognitive load is increased. The MOG has been related to visual attention, spatial and visuomotor processing, suggesting a decrease in one, two, or all of these components across the driving tasks in step with the associated decrease in blink rate (Mangun et al., 1998; LaBerge, 2000; Hahn et al., 2006; Wandell et al., 2007; Renier et al., 2010). However, it must be noted that the observed r^2 values were modest for all four brain regions. In addition to the natural variations

in human performance and brain function, the BOLD signal is influenced by numerous confounding factors, including age, sex and caffeine levels (D'Esposito et al., 2003; Chang et al., 2018). This can result in high inter-subject variability, reducing the power of the fMRI data. More work will be required to determine the main factor underlying the change in BOLD signal related to oculomotor behavior in the present context, involving a larger sample size and/or different analysis tools. Despite this, the observed relationships are important to note nevertheless. They support the interpretation that an increase in cognitive load across L, LT and LTA tasks increases certain brain activity to suppress blink rate; and also results in decreasing activity related to aspects of visual processing – suggesting a strategy of using neural resources efficiently to maintain task performance.

Limitations

The results of the present study must be considered in relation to a number of limitations in the experimental design and execution. Technical issues associated with fMRI-compatible eye-tracking caused the sample size to be somewhat lower than originally intended. One disadvantage of eye-tracking is that participant factors (e.g., glasses, contacts, eye color) can negatively impact how the pupil or the eye are tracked from the video recordings (Jacob and Karn, 2003; Nyström et al., 2013; Kok and Jarodzka, 2017). Additionally, cognitive processes cannot be inferred directly and conclusively from eye-movements. Visual attention can also be guided by peripheral vision, which cannot be captured through eye-tracking (Kok and Jarodzka, 2017). Just as it is possible to attend to something without directly fixating on it, it is also possible to fixate on something without attending to it - an effect known as attentional blindness (Mack, 2003). As a result, it can be challenging to develop a complete understanding of the relationships between eye movements and cognition. Thus, the eye-tracking results in the present study must be viewed as supporting (rather than defining) the interpretation of brain/behavior relationships across the various driving conditions.

Although the results largely replicate (while also extending from) a previous fMRI study with a very similar experimental design (Schweizer et al., 2013), additional work involving a larger sample size would be beneficial to better characterize human variability and potentially to detect important new findings with potentially smaller effect sizes in the driving conditions studied. The current study also examines the effect of distraction in a young adult population, which is like to differ when conducting analogous work involving the elderly. When distracted, older drivers have been found to drive slower with increased steering variability, and to commit more driving errors (Thompson et al., 2012). Although MVCs more frequently involve younger drivers (possibly due to lack of sufficient driving skills and incomplete maturation), older drivers have been attributed to an increased risk of MVCs in highly complex situations (e.g., intersections) due to age-associated changes in attention and cognitive decline (Stinchcombe and Gagnon, 2013). In addition, deficits in visual functioning, perception and information processing have been observed in the elderly (Kline et al., 1992; Stinchcombe and Gagnon, 2013). However, the present study of young adults is useful, as distracted driving is a prominent issue in this demographic with the rise in cell phone technology and social media (Klauer et al., 2014). This study also provides a normative data set to be used as a baseline for future studies.

The eye-tracking data obtained in this study were useful, but did not exhibit sufficient detection power to reveal statistically significant effects in all cases. In this study cohort, an increasing trend was seen in the pupil dilation size with the increase in cognitive load, although the effect was not significant. If this work were to be extended to an older adult or clinical population with cognitive deficits, this eye-tracking metric will remain important to investigate, as eye-tracking has been suggested for assessment of cognitive impairment and diagnosis of neurodegenerative conditions. Patients with dementia and Alzheimer's disease have been shown to demonstrate altered pupillary responses and oculomotor impairments, which can be associated to brain function and neural mechanisms to assess cognitive dysfunction (Pavisic et al., 2017; Marandi and Gazerani, 2019; Tao et al., 2020). Aging has also been associated with an increased latency of saccades and increased blink frequency in addition to neurocognitive decline (Marandi and Gazerani, 2019). Thus, the inclusion of eye-tracking methodology to future simulated driving studies may provide critical insight to the understanding of brain activity and behavior.

Another limitation of this study is that the simulation used did not replicate potential anxiety experienced in real driving due to the lack of stressing factors (e.g., other drivers, no real crash risk or danger). A subsequent experiment may expand on this area by incorporating more factors into the simulation such as vehicles to maintain a safe following distance, multiple lanes on the road with other vehicles, increased traffic or construction; as well as narratives, rewards and penalties related to driving performance.

Lastly, the simulated driving apparatus used inside the MRI system inevitably produced an experience that was different from real world driving, as the participants viewed the simulation through a projector screen while lying on a patient table, driving with minimal head and body motion. Usage of driving simulator hardware is suggested to provide a sufficient level of realism for evaluating demanding driving scenarios that would be dangerous to assess in on-road testing (Kan et al., 2012). However, the setup may create activation maps for simulated driving that imperfectly reflect the mental processes that occur when individuals drive motor vehicles in the real world (Kan et al., 2012). To mitigate this possibility, each participant underwent a training session with the driving simulator prior to fMRI. The training session was intended to familiarize participants with the simulator controls and thus ensure that their driving skills encompassed use of the experimental apparatus.

CONCLUSION

The present study uses a unique multi-measurement approach, combining fMRI with eye-tracking to measure brain and oculomotor behavior during simulated driving. It was observed that simple driving primarily involved the visual and motor systems, whereas the introduction of auditory distraction

shifted resources to the frontal and temporal systems for greater cognitive processing. Neural resources were modulated as task complexity increased, with re-allocation of resources when there were competing task demands in driving with auditory distraction. The inclusion of eye-tracking data furthered understanding of the effect of distraction on driving behavior and brain activity. Decreased brain activation in the visual system for distracted driving was supported by a more peaked mean distribution of gaze points as quantified by the percent fixation duration metric, concentrated on the road at the expense of the periphery, as well as a decrease in blink rate. Providing a baseline for comparison, these findings may be applied to future research involving an elderly or clinical population to determine the effect of aging or brain damage on the ability to multitask while driving. Because distraction promotes dangerous driving behavior and has been shown to display significant changes in brain activation and eye movements, assessing fitness to drive should consider including varying driving demands (e.g., with a secondary task introduced as distraction) to mimic real world scenarios where distraction is prevalent. Automotive companies may also wish to consider implementing strategies to mitigate visual tunneling when distracted, or to minimize the effect of distracting activities and communication devices.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Research Ethics Board at Sunnybrook Health Sciences Centre, Canada. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

TS and SG did the conceptualization and supervision. NY, FT, and SG did the investigation. SG did the resources, project administration, and funding acquisition. NY and SG did the data curation and writing-original draft preparation. NY, FT, NC, TS, and SG wrote-reviewed and edited and did the visualization. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by the Canadian Institutes of Health Research (grant number PGT 162241) and the Ontario Graduate Scholarship (OGS, provided by the University of Toronto to NY).

REFERENCES

Alnaes, D., Sneve, M. H., Espeseth, T., Endestad, T., van de Pavert, S. H. P., and Laeng, B. (2014). Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. *J. Vis.* 14:1. doi: 10.1167/14.4.1

Andersson, M., Ystad, M., Lundervold, A., and Lundervold, A. J. (2009). Correlations between measures of executive attention and cortical thickness of left posterior middle frontal gyrus - a dichotic listening study. *Behav. Brain Funct.* 5:41. doi: 10.1186/1744-9081-5-41

Berman, B. D., Horovitz, S. G., Morel, B., and Hallett, M. (2012). Neural correlates of blink suppression and the buildup of a natural bodily urge. *Neuroimage* 59, 1441–1450. doi: 10.1016/j.neuroimage.2011.08.050

Boisgueheneuc, F. D., Levy, R., Volle, E., Seassau, M., Duffau, H., Kinkingnehun, S., et al. (2006). Functions of the left superior frontal gyrus in humans: a lesion study. *Brain* 129, 3315–3328. doi: 10.1093/brain/awl244

Bowers, A. R., Anastasio, R. J., Sheldon, B. S., O'Connor, M. G., Hollis, A. M., Howe, P. D., et al. (2013). Can we improve clinical prediction of at-risk older drivers? *Accident Anal. Prevent.* 59, 537–547. doi: 10.1016/j.aap.2013.06.037

Brookings, J. B., Wilson, G. F., and Swain, C. R. (1996). Psychophysiological responses to changes in workload during simulated air traffic control. *Biol. Psychol.* 42, 361–377. doi: 10.1016/0301-0511(95)05167-8

Buschman, T. J., and Miller, E. K. (2007). Top-Down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* 315, 1860–1862. doi: 10.1126/science.1138071

Bush, G., Luu, P., and Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn. Sci.* 4, 215–222. doi: 10.1016/s1364-6613(00)01483-2

Calhoun, V. D., and Pearlson, G. D. (2012). A selective review of simulated driving studies: combining naturalistic and hybrid paradigms, analysis approaches, and future directions. *NeuroImage* 59, 25–35. doi: 10.1016/j.neuroimage.2011.06.037

Calhoun, V. D., Pekar, J. J., McGinty, V. B., Adali, T., Watson, T. D., and Pearlson, G. D. (2002). Different activation dynamics in multiple neural systems during simulated driving. *Hum. Brain Mapp.* 16, 158–167. doi: 10.1002/hbm.10032

Callan, A. M., Osu, R., Yamagishi, Y., Callan, D. E., and Inoue, N. (2009). Neural correlates of resolving uncertainty in driver's decision making. *Hum. Brain Mapp.* 30, 2804–2812. doi: 10.1002/hbm.20710

Cantin, V., Lavallière, M., Simoneau, M., and Teasdale, N. (2009). Mental workload when driving in a simulator: effects of age and driving complexity. *Accident Anal. Prevent.* 41, 763–771. doi: 10.1016/j.aap.2009.03.019

Chang, D., Song, D., Zhang, J., Shang, Y., Ge, Q., and Wang, Z. (2018). Caffeine caused a widespread increase of resting brain entropy. *Sci. Rep.* 8:2700.

Chao, L. L., Haxby, J. V., and Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nat. Neurosci.* 2, 913–919. doi: 10.1038/13217

Choi, M.-H., Kim, H.-S., Yoon, H.-J., Lee, J.-C., Baek, J.-H., Choi, J.-S., et al. (2017). Increase in brain activation due to sub-tasks during driving: fMRI study using new MR-compatible driving simulator. *J. Physiol. Anthropol.* 36:11.

Cox, R. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comp. Biomed. Res.* 29, 162–173. doi: 10.1006/cbmr.1996.0014

Crundall, D., and Underwood, G. (2011). "Visual attention while driving," in *Handbook of Traffic Psychology*, ed. B. E. Porter (Amsterdam: Elsevier), 137–148. doi: 10.1016/b978-0-12-381984-0.10011-6

Démonet, J. F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. L., Wise, R., et al. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain* 115(Pt 6), 1753–1768. doi: 10.1093/brain/115.6.1753

D'Esposito, M., Deouell, L. Y., and Gazzaley, A. (2003). Alterations in the BOLD fMRI signal with ageing and disease: a challenge for neuroimaging. *Nat. Rev. Neurosci.* 4, 863–872. doi: 10.1038/nrn1246

Drews, F. A., Yazdani, H., Godfrey, C. N., Cooper, J. M., and Strayer, D. L. (2009). Text messaging during simulated driving. *Hum. Factors* 51, 762–770. doi: 10.1177/0018720809353319

Dumoulin, S. O., Bittar, R. G., Kabani, N. J., Baker, C. L., Le Goualher, G., Bruce Pike, G., et al. (2000). A new anatomical landmark for reliable identification of human area V5/MT: a quantitative analysis of sulcal patterning. *Cereb. Cortex* 10, 454–463. doi: 10.1093/cercor/10.5.454

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

Fan, J., Hof, P. R., Guise, K. G., Fossella, J. A., and Posner, M. I. (2008). The functional integration of the anterior cingulate cortex during conflict processing. *Cereb. Cortex* 18, 796–805. doi: 10.1093/cercor/bhm125

Faure, V., Lobjois, R., and Benguigui, N. (2016). The effects of driving environment complexity and dual tasking on drivers' mental workload and eye blink behavior. *Transportation Res. Part F: Traffic Psychol. Behav.* 40, 78–90. doi: 10.1016/j.trf.2016.04.007

Fitch, G. M., Bartholomew, P. R., Hanowski, R. J., and Perez, M. A. (2015). Drivers' visual behaviour when using handheld and hands-free cell phones. *J. Safety Res.* 54, 105–108.e29. doi: 10.1016/j.jsr.2015.06.008

Fraade-Blanar, L. A., Ebel, B. E., Larson, E. B., Sears, J. M., Thompson, H. J., Chan, K. C. G., et al. (2018). Cognitive decline and older driver crash risk. *J. Am. Geriatr. Soc.* 66, 1075–1081. doi: 10.1111/jgs.15378

Friedman-Hill, S. R., Robertson, L. C., Desimone, R., and Ungerleider, L. G. (2003). Posterior parietal cortex and the filtering of distractors. *Proc. Natl. Acad. Sci. U S A.* 100, 4263–4268. doi: 10.1073/pnas.0730772100

Gable, T. M., Kun, A. L., Walker, B. N., and Winton, R. J. (2015). "Comparing heart rate and pupil size as objective measures of workload in the driving context: initial look," in *Proceedings of the Adjunct Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '15*, (Nottingham: ACM Press), 20–25.

Genovese, C. R., Lazar, N. A., and Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage* 15, 870–878. doi: 10.1006/nimg.2001.1037

Graydon, F. X., Young, R., Benton, M. D., Genik, R. J., Posse, S., Hsieh, L., et al. (2004). Visual event detection during simulated driving: identifying the neural correlates with functional neuroimaging. *Transportation Res. Part F: Traffic Psychol. Behav.* 7, 271–286. doi: 10.1016/j.trf.2004.09.006

Haak, M., Bos, S., Panic, S., and Rothkrantz, L. J. M. (2009). "Detecting stress using eye blinks and brain activity from EEG signals," in *Proceeding of the 1st Driver Car Interaction and Interface (DCII 2008)*, Netherland 35–60.

Hahn, B., Ross, T. J., and Stein, E. A. (2006). Neuroanatomical dissociation between bottom–up and top–down processes of visuospatial selective attention. *NeuroImage* 32, 842–853. doi: 10.1016/j.neuroimage.2006.04.177

Hamish Jamson, A., and Merat, N. (2005). Surrogate in-vehicle information systems and driver behaviour: effects of visual and cognitive load in simulated rural driving. *Transportation Res. Part F: Traffic Psychol. Behav.* 8, 79–96. doi: 10.1016/j.trf.2005.04.002

Hampshire, A., Chamberlain, S. R., Monti, M. M., Duncan, J., and Owen, A. M. (2010). The role of the right inferior frontal gyrus: inhibition and attentional control. *Neuroimage* 50, 1313–1319. doi: 10.1016/j.neuroimage.2009.12.109

Hird, M. A., Vesely, K. A., Fischer, C. E., Graham, S. J., Naglie, G., and Schweizer, T. A. (2017). Investigating simulated driving errors in amnestic single- and multiple-domain mild cognitive impairment. *J. Alzheimer's Dis.* 56, 447–452. doi: 10.3233/jad-160995

Holland, M. K., and Tarlow, G. (1972). Blinking and mental load. *Phych. Rep.* 31, 119–127.

Hyung-Sik, K., Mi-Hyun, C., Hee-Jeong, Y., Hyun-Joo, K., Ul-Ho, J., Sung-Jun, P., et al. (2014). Cerebral activation and lateralization due to the cognition of a various driving speed difference: an fMRI study. *Biomed. Mater. Eng.* 24, 1133–1139. doi: 10.3233/bme-130913

Jacob, R. J. K., and Karn, K. S. (2003). "Eye tracking in human-computer interaction and usability research: ready to deliver the promises," in *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, eds H. Deubel and J. R. I. Hyönä (Amsterdam: Elsevier Science).

Japee, S., Holiday, K., Satyshur, M. D., Mukai, I., and Ungerleider, L. G. (2015). A role of right middle frontal gyrus in reorienting of attention: a case study. *Front. Syst. Neurosci.* 9:23. doi: 10.3389/fnsys.2015.00023

Joshi, S., Li, Y., Kalwani, R. M., and Gold, J. I. (2016). Relationships between pupil diameter and neuronal activity in the locus coeruleus. colliculi, and cingulate cortex. *Neuron* 89, 221–234. doi: 10.1016/j.neuron.2015.11.028

Just, M. A., Keller, T. A., and Cynkar, J. (2008). A decrease in brain activation associated with driving when listening to someone speak. *Brain Res.* 1205, 70–80. doi: 10.1016/j.brainres.2007.12.075

Kan, K., Schweizer, T. A., Tam, F., and Graham, S. J. (2012). Methodology for functional MRI of simulated driving: functional MRI of simulated driving. *Med. Phys.* 40:012301. doi: 10.1118/1.4769107

Klauer, S. G., Guo, F., Simons-Morton, B. G., Ouimet, M. C., Lee, S. E., and Dingus, T. A. (2014). Distracted driving and risk of road crashes among novice and experienced drivers. *New Engl. J. Med.* 370, 54–59. doi: 10.1056/nejmsa1204142

Kline, D., Kline, T., Fozard, J., Kosnik, W., Schieber, F., and Sekuler, R. (1992). Vision, aging, and driving: the problems of older drivers. *J. Gerontol.* 47, 27–34.

Kok, E. M., and Jarodzka, H. (2017). Before your very eyes: the value and limitations of eye tracking in medical education. *Med. Educ.* 51, 114–122. doi: 10.1111/medu.13066

Kramer, A. F., Wiegmann, D. A., and Kirlik, A. (2006). *AttentionFrom Theory to Practice*. Oxford: Oxford University Press.

Kret, M. E., and Sjak-Shie, E. E. (2019). Preprocessing pupil size data: guidelines and code. *Behav. Res. Methods* 51, 1336–1342. doi: 10.3758/s13428-018-1075-y

Künzle, H., Akert, K., and Wurtz, R. H. (1976). Projection of area 8 (frontal eye field) to superior colliculus in the monkey. an autoradiographic study. *Brain Res.* 117, 487–492. doi: 10.1016/0006-8993(76)90754-x

LaBerge, D. (2000). "Networks of attention," in *The New Cognitive Neurosciences*, ed. Michael S. Gazzaniga (Cambridge, MA: MIT Press), 711–724.

Land, M. F. (2006). Eye movements and the control of actions in everyday life. *Prog. Retinal Eye Res.* 25, 296–324. doi: 10.1016/j.preteyeres.2006.01.002

Lee, H. C., Cameron, D., and Lee, A. H. (2003). Assessing the driving performance of older adult drivers: on-road versus simulated driving. *Accident Anal. Prevent.* 35, 797–803. doi: 10.1016/s0001-4575(02)00083-0

Lerner, A., Bagic, A., Hanakawa, T., Boudreau, E. A., Pagan, F., Mari, Z., et al. (2009). Involvement of insula and cingulate cortices in control and suppression of natural urges. *Cereb. Cortex* 19, 218–223. doi: 10.1093/cercor/bhn074

Lezak, M. D., Howieson, D. B., Loring, D. W., and Fischer, J. S. (2004). *Neuropsychological Assessment*. New York, NY: Oxford University Press.

Li, Y.-O., Eichele, T., Calhoun, V. D., and Adali, T. (2012). Group study of simulated driving fMRI data by multiset canonical correlation analysis. *J. Signal Process. Systems* 68, 31–48. doi: 10.1007/s11265-010-0572-8

Liakakis, G., Nickel, J., and Seitz, R. J. (2011). Diversity of the inferior frontal gyrus—a meta-analysis of neuroimaging studies. *Behav. Brain Res.* 225, 341–347. doi: 10.1016/j.bbr.2011.06.022

Lundberg, C., Johansson, K., Ball, K., Bjerre, B., Blomqvist, C., Braekhus, A., et al. (1997). Dementia and driving: an attempt at consensus. *Alzheimer Dis. Assoc. Disord.* 11, 28–37. doi: 10.1097/00002093-199703000-00006

Mack, A. (2003). Inattentional blindness: looking without seeing. *Curr. Direct. Psychol. Sci.* 12, 180–184. doi: 10.1111/1467-8721.01256

Maffei, A., and Angrilli, A. (2018). Spontaneous eye blink rate: an index of dopaminergic component of sustained attention and fatigue. *Int. J. Psychophysiol.* 123, 58–63. doi: 10.1016/j.ijpsycho.2017.11.009

Mangun, G. R., Buonocore, M. H., Girelli, M., and Jha, A. P. (1998). ERP and fMRI measures of visual spatial selective attention. *Hum. Brain Mapp.* 6, 383–389. doi: 10.1002/(sici)1097-0193(1998)6:5/6<383::aid-hbm10>3.0.co;2-z

Manor, B. R., and Gordon, E. (2003). Defining the temporal threshold for ocular fixation in free-viewing visuocognitive tasks. *J. Neurosci. Methods* 128, 85–93. doi: 10.1016/s0165-0270(03)00151-1

Mansur, A., Hird, M. A., Desimone, A., Pshonyak, I., Schweizer, T. A., and Das, S. (2018). Driving habits and behaviors of patients with brain tumors: a self-report, cognitive and driving simulation study. *Sci. Rep.* 15:4635.

Marandi, R. Z., and Gazerani, P. (2019). Aging and eye tracking: in the quest for objective biomarkers. *Future Neurol.* 14:FNL33.

Marshall, S. C., and Gilbert, N. (1999). Saskatchewan physicians' attitudes and knowledge regarding assessment of medical fitness to drive. *CMAJ* 160, 1701–1704.

Matthys, K., Smits, M., Geest, J. N. V., der, Lugt, A. V., der, et al. (2009). Mirror-Induced visual illusion of hand movements: a functional magnetic resonance

imaging study. *Arch. Phys. Med. Rehabil.* 90, 675–681. doi: 10.1016/j.apmr.2008.09.571

Mazzone, L., Yu, S., Blair, C., Gunter, B. C., Wang, Z., Marsh, R., et al. (2010). An fMRI study of frontostriatal circuits during the inhibition of eye blinking in persons with tourette syndrome. *Am. J. Psychiatry* 167, 341–349. doi: 10.1176/appi.ajp.2009.08121831

McIntire, L. K., McKinley, R. A., Goodyear, C., and McIntire, J. P. (2014). Detection of vigilance performance using eye blinks. *Appl. Ergon.* 45, 354–362. doi: 10.1016/j.apergo.2013.04.020

Moore, T., and Fallah, M. (2001). Control of eye movements and spatial attention. *Proc. Natl. Acad. Sci. U S A.* 98, 1273–1276.

Moss, H. E., Abdallah, S., Fletcher, P., Bright, P., Pilgrim, L., Acres, K., et al. (2005). Selecting among competing alternatives: selection and retrieval in the left inferior frontal gyrus. *Cereb. Cortex* 15, 1723–1735. doi: 10.1093/cercor/bhi049

Murphy, P. R., O'Connell, R. G., O'Sullivan, M., Robertson, I. H., and Balsters, J. H. (2014). Pupil diameter covaries with bold activity in human locus coeruleus: pupil diameter and locus coeruleus activity. *Hum. Brain Mapp.* 35, 4140–4154. doi: 10.1002/hbm.22466

National Center for Statistics and Analysis (2013). *Distracted Driving 2011. Traffic Safety Facts Research Note.* Report No. DOT HS 811 737, Washington, DC: National Highway Traffic Safety Administration.

National Center for Statistics and Analysis (2018). *Distracted Driving 2016. Traffic Safety Facts Research Note.* Report No. DOT HS 812 517, Washington, DC: National Highway Traffic Safety Administration.

Neufang, S., Akhrif, A., Riedl, V., Förstl, H., Kurz, A., Zimmer, C., et al. (2011). Disconnection of frontal and parietal areas contributes to impaired attention in very early Alzheimer's disease. *J. Alzheimer's Dis. JAD* 25, 309–321. doi: 10.3233/jad-2011-102154

Niezgoda, M., Tarnowski, A., Kruszewski, M., and Kamiński, T. (2015). Towards testing auditory–vocal interfaces and detecting distraction while driving: a comparison of eye-movement measures in the assessment of cognitive workload. *Transportation Res. Part F: Traffic Psychol. Behav.* 32, 23–34. doi: 10.1016/j.trf.2015.04.012

Nij Bijvank, J. A., Petzold, A., Balk, L. J., Tan, H. S., Uitdehaag, B. M. J., Theodorou, M., et al. (2018). A standardized protocol for quantification of saccadic eye movements: DEMoNS. *PLoS One* 13:e0200695. doi: 10.1371/journal.pone.0200695

Noyce, D. A., Chitturi, M. V., Nassereddine, H., Santiago-Chaparro, K. R., and Bill, A. R. (2017). *Neural Correlates of Older Driver Performance*. Washington, DC: University Transportation Center.

Nyström, M., Andersson, R., Holmqvist, K., and van de Weijer, J. (2013). The influence of calibration method and eye physiology on eyetracking data quality. *Behav. Res. Methods* 45, 272–288. doi: 10.3758/s13428-012-0247-4

Ohira, H. (1996). Eyeblink activity in a word-naming task as a function of semantic priming and cognitive load. *Percept. Mot. Skills* 82, 835–842. doi: 10.2466/pms.1996.82.3.835

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

Orr, C., and Hester, R. (2012). Error-related anterior cingulate cortex activity and the prediction of conscious error awareness. *Front. Hum. Neurosci.* 6:177. doi: 10.3389/fnhum.2012.00177

Ortiz, N., Ramnarayan, M., and Mizenko, K. (2016). The effect of distraction on road user behavior: an observational pilot study across intersections in Washington, DC. *J. Transport Health* 3:S67.

Palinko, O., Kun, A. L., Shyrokov, A., and Heeman, P. (2010). "Estimating cognitive load using remote eye tracking in a driving simulator," in *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, (New York, NY: ACM), 141–144.

Pardo, J. V., Fox, P. T., and Raichle, M. E. (1991). Localization of a human system for sustained attention by positron emission tomography. *Nature* 349, 61–64. doi: 10.1038/349061a0

Paus, T. (2001). Primate anterior cingulate cortex: where motor control, drive and cognition interface. *Nat. Rev. Neurosci.* 2, 417–424. doi: 10.1038/35077500

Pavisic, I. M., Firth, N. C., Parsons, S., Rego, D. M., Shakespeare, T. J., Yong, K. X. X., et al. (2017). Eyetracking metrics in young onset Alzheimer's disease: a window into cognitive visual functions. *Front. Neurol.* 8:377. doi: 10.3389/fneur.2017.00377

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., and Raichle, M. E. (1989). Positron emission tomographic studies of the processing of singe words. *J. Cogn. Neurosci.* 1, 153–170. doi: 10.1162/jocn.1989.1.2.153

Raichle, M. E. (2015). The brain's default mode network. *Annu. Rev. Neurosci.* 38, 433–447.

Recarte, M. A., and Nunes, L. M. (2003). Mental workload while driving: effects on visual search, discrimination, and decision making. *J. Exp. Psychol. Appl.* 9, 119–137. doi: 10.1037/1076-898x.9.2.119

Renier, L. A., Anurova, I., De Volder, A. G., Carlson, S., VanMeter, J., and Rauschecker, J. P. (2010). Preserved functional specialization for spatial processing in the middle occipital gyrus of the early blind. *Neuron* 68, 138–148. doi: 10.1016/j.neuron.2010.09.021

Rizzo, M. (2011). Impaired driving from medical conditions: a 70-year-old man trying to decide if he should continue driving. *JAMA* 305:1018. doi: 10.1001/jama.2011.252

Rueckert, L., and Grafman, J. (1996). Sustained attention deficits in patients with right frontal lesions. *Neuropsychologia* 34, 953–963. doi: 10.1016/0028-3932(96)00016-4

Rueckert, L., and Grafman, J. (1998). Sustained attention deficits in patients with lesions of posterior cortex. *Neuropsychologia* 36, 653–660. doi: 10.1016/s0028-3932(97)00150-4

Schneider, W. X. (1995). VAM: a neuro-cognitive model for visual attention control of segmentation, object recognition, and space-based motor action. *Visual Cogn.* 2, 331–376. doi: 10.1080/13506289508401737

Schweizer, T. A., Kan, K., Hung, Y., Tam, F., Naglie, G., and Graham, S. J. (2013). Brain activity during driving with distraction: an immersive fMRI study. *Front. Hum. Neurosci.* 7:53. doi: 10.3389/fnhum.2013.00053

Siegle, G. J., Steinhauer, S. R., Stenger, V. A., Konecky, R., and Carter, C. S. (2003). Use of concurrent pupil dilation assessment to inform interpretation and analysis of fMRI data. *NeuroImage* 20, 114–124. doi: 10.1016/s1053-8119(03)00298-2

Spiers, H. J., and Maguire, E. A. (2007). Neural substrates of driving behaviour. *NeuroImage* 36, 245–255. doi: 10.1016/j.neuroimage.2007.02.032

Stinchcombe, A., and Gagnon, S. (2013). Aging and driving in a complex world: exploring age differences in attentional demand while driving. *Transportation Res. Part F: Traffic Psychol. Behav.* 17, 125–133. doi: 10.1016/j.trf.2012.11.002

Strayer, D. L., and Drew, F. A. (2004). Profiles in driver distraction: effects of cell phone conversations on younger and older drivers. *Hum. Factors: J. Hum. Factors Ergonom. Soc.* 46, 640–649. doi: 10.1518/hfes.46.4.640.56806

Swick, D., Ashley, V., and Turken, A. U. (2008). Left inferior frontal gyrus is critical for response inhibition. *BMC Neurosci.* 9:102. doi: 10.1186/1471-2202-9-102

Talairach, J., and Tournoux, P. (1988). *Co-planar Stereotaxic Atlas of the Human Brain: 3-Dimensional Proportional System: an Approach to Cerebral Imaging*. New York, NY: Thieme Medical Publishers, Inc.

Tao, L., Wang, Q., Liu, D., Wang, J., Zhu, Z., and Feng, L. (2020). Eye tracking metrics to screen and assess cognitive impairment in patients with neurological disorders. *Neurol. Sci.* 41, 1697–1704. doi: 10.1007/s10072-020-04310-y

Thompson, K. R., Johnson, A. M., Emerson, J. L., Dawson, J. D., Boer, E. R., and Rizzo, M. (2012). Distracted driving in elderly and middle-aged drivers. *Accident Anal. Prevent.* 45, 711–717. doi: 10.1016/j.aap.2011.09.040

Tops, M., and Boksem, M. A. S. (2011). A potential role of the inferior frontal gyrus and anterior insula in cognitive control, brain rhythms, and event-related potentials. *Front. Psychol.* 2:330. doi: 10.3389/fpsyg.2011.00330

Tranel, D., Damasio, H., and Damasio, A. R. (1997). A neural basis for the retrieval of conceptual knowledge. *Neuropsychologia* 35, 1319–1327. doi: 10.1016/s0028-3932(97)00085-7

Traffic safety Association. (2013). *Distracted Driving 2011. Traffic Safety Facts*. Fraser, MI: Traffic safety Association.

Tsai, Y.-F., Viirre, E., Strychacz, C., Chase, B., and Jung, T.-P. (2007). Task performance and eye activity: predicting behavior relating to cognitive workload. *Aviat. Space Environ. Med.* 78, B176–B185.

Uchiyama, Y., Ebe, K., Kozato, A., Okada, T., and Sadato, N. (2003). The neural substrates of driving at a safe distance: a functional MRI study. *Neurosci. Lett.* 352, 199–202. doi: 10.1016/j.neulet.2003.08.072

Uchiyama, Y., Toyoda, H., Sakai, H., Shin, D., Ebe, K., and Sadato, N. (2012). Suppression of brain activity related to a car-following task with an auditory task: an fMRI study. *Transportation Res. Part F: Traffic Psychol. Behav.* 15, 25–37. doi: 10.1016/j.trf.2011.11.002

U.S. Department of Transportation National Highway Traffic Safety Administration (2008). *National Motor Vehicle Crash Causation Survey: Report to Congress.*

Van Orden, K. F., Limbert, W., Makeig, S., and Jung, T.-P. (2001). Eye activity correlates of workload during a visuospatial memory task. *Hum. Fact.* 43, 111–121. doi: 10.1518/001872001775992570

Vernet, M., Quentin, R., Chanes, L., Mitsumasu, A., and Valero-Cabré, A. (2014). Frontal eye field, where art thou? anatomy, function, and non-invasive manipulation of frontal regions involved in eye movements and associated cognitive operations. *Front. Integr. Neurosci.* 8:66. doi: 10.3389/fnint.2014.00066

Wandell, B. A., Dumoulin, S. O., and Brewer, A. A. (2007). Visual field maps in human cortex. *Neuron* 56, 366–383. doi: 10.1016/j.neuron.2007.10.012

Ware, M., Feng, J., and Nam, C. S. (2020). "Neuroergonomics behind the wheel: neural correlates of driving," in *Neuroergonomics: Principles and Practice*, ed. C. S. Nam (Cham: Springer International Publishing), 353–388. doi: 10.1007/978-3-030-34784-0_18

Weissman, D. H., and Prado, J. (2012). Heightened activity in a key region of the ventral attention network is linked to reduced activity in a key region of the dorsal attention network during unexpected shifts of covert visual spatial attention. *NeuroImage* 61, 798–804. doi: 10.1016/j.neuroimage.2012.03.032

Wilkins, A. J., Shallice, T., and McCarthy, R. (1987). Frontal lesions and sustained attention. *Neuropsychologia* 25, 359–365. doi: 10.1016/0028-3932(87)90024-8

Withaar, F. K., Brouwer, W. H., and Van Zomeren, A. H. (2000). Fitness to drive in older drivers with cognitive impairment. *J. Int. Neuropsychol. Soc.* 6, 480–490. doi: 10.1017/s1355617700644065

Wu, D., Huang, H., Liu, N., and Miao, D. (2019). Information processing under high and low distractions using eye tracking. *Cogn. Process.* 20, 11–18. doi: 10.1007/s10339-018-0876-3

Zekveld, A. A., Heslenfeld, D. J., Johnsrude, I. S., Versfeld, N. J., and Kramer, S. E. (2014). The eye as a window to the listening brain: neural correlates of pupil size as a measure of cognitive listening load. *NeuroImage* 101, 76–86. doi: 10.1016/j.neuroimage.2014.06.069

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.

frontiers in Behavioral Neuroscience

ORIGINAL RESEARCH
published: 10 September 2021
doi: 10.3389/fnbeh.2021.743064

Effects of Aging on the Neural Mechanisms Underlying the Recollection of Memories Encoded by Social Interactions With Persons in the Same and Different Age Groups

Eri Tsuruha[1,2] *and Takashi Tsukiura*[1]*

[1] Department of Cognitive and Behavioral Sciences, Graduate School of Human and Environmental Studies, Kyoto University, Kyoto, Japan, [2] Graduate School of Advanced Integrated Studies in Human Survivability, Kyoto University, Kyoto, Japan

OPEN ACCESS

Edited by:
Kuniyoshi L. Sakai,
The University of Tokyo, Japan

Reviewed by:
Yuji Naya,
Peking University, China
Edgardo O. Alvarez,
Laboratorio de Neuropsicofarmacología Experimental – CONICET Mendoza, Argentina

***Correspondence:**
Takashi Tsukiura
tsukiura.takashi.6c@kyoto-u.ac.jp

Specialty section:
This article was submitted to Learning and Memory, a section of the journal Frontiers in Behavioral Neuroscience

Received: *17 July 2021*
Accepted: *23 August 2021*
Published: *10 September 2021*

Citation:
Tsuruha E and Tsukiura T (2021) Effects of Aging on the Neural Mechanisms Underlying the Recollection of Memories Encoded by Social Interactions With Persons in the Same and Different Age Groups. Front. Behav. Neurosci. 15:743064. doi: 10.3389/fnbeh.2021.743064

Memories related to ingroup members are remembered more accurately than those related to outgroup members. However, little is known about the age-dependent differences in neural mechanisms underlying the retrieval of memories shared with ingroup or outgroup members that are categorized by age-group membership. The present functional magnetic resonance imaging (fMRI) study investigated this issue. Healthy young and older adults participated in a 2-day experiment. On the first day outside fMRI, participants were presented with words by unfamiliar persons in movie clips and exchanged each word with persons belonging to the same age group (SAG) or different age group (DAG). On the second day during fMRI, participants were randomly presented with learned and new words one by one, and they judged whether each word had been encoded with either SAG or DAG members or neither. fMRI results demonstrated that an age-dependent decrease in successful retrieval activation of memories presented by DAG was identified in the anterior temporal lobe (ATL) and hippocampus, whereas with memories presented by SAG, an age-dependent decrease in activation was not found in any regions. In addition, an age-dependent decrease in functional connectivity was significant between the hippocampus/ATL and posterior superior temporal sulcus (pSTS) during the successful retrieval of memories encoded with the DAG people. The "other"-related mechanisms including the hippocampus, ATL, and pSTS with memories learned with the outgroup members could decrease in older adults, whereas with memories learned with the ingroup members, the "self"-related mechanisms could be relatively preserved in older adults.

Keywords: fMRI, source memory, aging, generation, social interaction, hippocampus, anterior temporal lobe, superior temporal sulcus

INTRODUCTION

Memories for persons who belong to the same age group (SAG) as ingroup members are remembered more accurately than memories for persons who belong to the different age group (DAG) as outgroup members (for review, see Rhodes and Anastasi, 2012). This enhancing effect on memory is known as the own-age bias, which is a type of intergroup bias (for review, see Molenberghs and Louis, 2018). Previous studies have consistently demonstrated that faces, words, and behaviors related to ingroup members are perceived and remembered more accurately than those related to outgroup members (for review, see Molenberghs and Louis, 2018). In addition, there is functional neuroimaging evidence that recollection-related processes, such as association memory or source memory, are significantly disturbed by aging, and age-dependent disturbances in recollection are caused by decreased activation in the hippocampus (for review, see Shing et al., 2010; Koen and Yonelinas, 2014). However, little is known about how neural mechanisms underlying the recollection of memories related to the SAG people as an ingroup member and the DAG people as an outgroup member are changed by aging. To address this issue, the present functional magnetic resonance imaging (fMRI) study investigated age-related differences in task-related activation and functional connectivity during the recollection of memories encoded by social interactions with persons belonging to the SAG people as an ingroup member and the DAG people as an outgroup member.

Neural mechanisms during the processing of information related to ingroup members are shared with neural mechanisms during the processing of information related to "self" (for review, see Northoff et al., 2006), whereas neural mechanisms are common between the processing of information related to outgroup members and of "other"-related information (Golby et al., 2001; Rilling et al., 2008; Greven and Ramsey, 2017). For example, people belonging to the ingroup are regarded as the social self, which involves the dorsomedial prefrontal cortex (dmPFC; Volz et al., 2009) or precuneus (Scheepers et al., 2013). In addition, functional neuroimaging studies have shown that functional connectivity between the memory-related hippocampus and the self-related cortical midline structures (CMS), which include the dmPFC and precuneus, contributes to the self-reference effect known as the memory enhancement of self-related information (for review, see Murray et al., 2015). The importance of hippocampus-CMS interactions in the self-reference effect on memory was also reported in a social context, in which memories for unfamiliar faces encoded with the anticipation of future friendship with self were significantly enhanced compared to face memories encoded with the anticipation of friendship with others (Yamawaki et al., 2017). Regarding the processing of information related to outgroup members, one fMRI study demonstrated that the posterior superior temporal sulcus (pSTS), reflecting social cognition of others (for review, see Patel et al., 2019), showed significant activation in prosocial behavior toward outgroup members compared to those toward ingroup members (Do and Telzer, 2019). In addition, activation in the right hippocampus and the right pSTS, which reflects "theory of mind" related to the process of inferring mental states in others, was significantly identified in the competition with other persons compared to that with machines (Polosan et al., 2011). Significant activation in the right pSTS has been found when perceiving social information such as faces (Dasgupta et al., 2017) or when watching movies of faces with rich social information (Pitcher et al., 2011). Thus, memories of ingroup members could involve interacting mechanisms between the self-related CMS and the memory-related hippocampus, whereas interactions between the right pSTS related to social cognition for others and the memory-related hippocampus could contribute to memories of outgroup members.

Episodic memory is disrupted by aging (for review, see Tromp et al., 2015; Nyberg, 2017), and age-dependent decreases in the recollection of episodic details have been linked with disrupted function of the hippocampus in older adults (for review, see Shing et al., 2010; Koen and Yonelinas, 2014). For example, one fMRI study reported that recollection-related activation in the hippocampus age-dependently decreased, whereas familiarity-related activation in the entorhinal and perirhinal cortex was relatively preserved in older adults (Daselaar et al., 2006). An age-dependent decrease in hippocampal activation was also identified in the recollection of autobiographical memory (St. Jacques et al., 2012). In addition, there is functional neuroimaging evidence that an age-dependent decrease in activation and functional connectivity related to memories for face-name associations has been found in the anterior temporal lobe (ATL) as well as the hippocampus. For example, an age-related decrease in correlations between hippocampal and ATL activation was observed in the retrieval of social knowledge (name or job title) associated with faces (Tsukiura et al., 2011). Another study using transcranial direct current stimulation (tDCS) demonstrated that the retrieval of face-name associations was significantly improved by tDCS stimulation over the left ATL in both young and older adults, whereas tDCS stimulation over the right ATL had a beneficial effect on the retrieval of face-name associations only in young adults (Ross et al., 2011). In neuropsychological studies of brain-damaged patients, patients with ATL atrophy were significantly impaired in perceiving faces or voices of familiar persons (for review, see Gainotti, 2013; Cosseddu et al., 2018). Taken together with previous findings, an age-related decrease in the functional network between the self-related CMS and the memory-related hippocampus-ATL regions could be critical in age-related decreases in memory for faces belonging to the SAG people as an ingroup member, whereas the aging effects on memory for faces belonging to the DAG people as an outgroup member could be modulated by age-related decreases in the functional network between the other-related pSTS and the memory-related hippocampus-ATL regions. However, little is known about the neural mechanisms underlying age-related differences in the retrieval of memories shared with ingroup and outgroup members that are categorized by age-group memberships.

To address this issue, using the event-related fMRI technique, we scanned healthy young and older adults during the retrieval of memories shared with the SAG people as ingroup members and with the DAG people as outgroup members. On the

basis of previous studies, we made two predictions. First, activation in the dmPFC, precuneus, hippocampus, and right ATL would significantly decrease by aging during the recollection of memories associated with the SAG people as ingroup members, whereas age-dependent decreased activation in the right pSTS, hippocampus, and right ATL would be found in the recollection of memories associated with the DAG people as outgroup members. Second, functional connectivity between the CMS regions, including the dmPFC or precuneus, and the hippocampus-ATL regions would be significantly impaired by aging during the recollection of memories encoded with the SAG people as ingroup members, whereas age-dependent decreases in functional connectivity between the right pSTS and the hippocampus-ATL regions would be identified in the recollection of memories encoded with the DAG people as outgroup members.

MATERIALS AND METHODS

Participants

Thirty-six healthy young adults (17 females; mean age: 22.42, SD: 1.66) and 36 healthy older adults (18 females; mean age: 66.22, SD: 3.01) participated in a 2-day experiment and were paid for their participation. All participants were right-handed, native Japanese speakers, with no history of neurological or psychiatric disease. All young participants were recruited from the Kyoto University community, and all older participants were recruited from Kyoto City Silver Human Resource Center. All participants gave informed consent to an experimental protocol approved by the Institutional Review Board (IRB) of the Graduate School of Human and Environmental Studies, the Kyoto University (29-H-8).

All participants performed several neuropsychological tests, including the FLANDERS handedness test (Nicholls et al., 2013; Okubo et al., 2014), the Japanese version of Montreal Cognitive Assessment (MoCA-J; Nasreddine et al., 2005; Fujiwara et al., 2010), and Center for Epidemiologic Studies Depression Scale (CES-D; Radloff, 1977; Shima, 1998). In addition, personality and social traits in each participant were assessed by NEO–FFI (Costa and MarcCrae, 1992; Shimonaka et al., 1999), University of California Los Angeles Loneliness Scale (UCLA-LS; Russell et al., 1978; Masuda et al., 2012), Subjective Well-Being Scale (SWBS; Sell and Nagpal, 1992; Ito et al., 2003), Interdependent Happiness Scale (IHS; Hitokoto and Uchida, 2015), Japanese version of Rosenberg Self-Esteem Scale (RSES-J; Rosenberg, 1965; Mimura and Griffiths, 2007), and Loyola Generativity Scale (LGS; Mcadams and Destaubin, 1992; Tabuchi et al., 2012). However, these personality and social trait scores were not included in all analyses of the present study.

One young participant and one older participant had MoCA-J scores lower than 2 SD below the mean scores in each group, and four young participants and five older participants had CES-D scores that extended far beyond the cutoff value. In addition, one young participant was found to have possible pathological changes (probable arachnoid cyst) in a structural MRI, one older participant did not complete her/his MRI scan due to a piece of internal metal, one older participant had difficulty communicating without hearing aids, one older participant had difficulty perceiving visual stimuli due to glaucoma, one young participant and one older participant showed head movement greater than 3 mm during the scanning, and one older participant showed more no-response trials than 2 SD above the mean number of no-response trials in all participants. Based on the exclusion criteria, behavioral and MRI data from 19 participants (seven young and 12 older adults, some of whom met several exclusion criteria) were excluded from all analyses. Thus, we analyzed data from 29 young adults (15 females; mean age: 22.45, SD: 1.72) and 24 older adults (11 females; mean age: 65.96, SD: 3.09) in the present study.

Neuropsychological test scores were compared between the two groups of 29 young and 24 older adults by two-sample *t*-tests (two-tailed). Significant differences between the two groups were identified in age [$t_{(51)} = 64.78$, $p < 0.01$, $d = 17.88$], education years [$t_{(51)} = 5.87$, $p < 0.01$, $d = 1.62$], and MoCA-J scores [$t_{(51)} = 5.01$, $p < 0.01$, $d = 1.38$]. However, we did not find a significant difference between the two groups in the FLANDERS handedness test [$t_{(51)} = 0.37$, $p = 0.71$, $d = 0.10$] and CES-D scores [$t_{(51)} = 1.01$, $p = 0.32$, $d = 0.28$]. Detailed profiles of young and older participants are summarized in **Table 1**.

Stimuli

A total of 120 Japanese words were selected from a database of two-letter Japanese Kanji words (Itsukushima et al., 1991). They were divided into three lists of 40 words, the two of which were used for target stimuli in two encoding conditions and the remaining one was used for distracter stimuli. These lists were counterbalanced across participants. In addition, each list was subdivided into two lists of 20 words, which were applied to two presenters in the encoding task described below. The attributions of imagery, concreteness, ease of learning, and frequency in each word were statistically equalized among these six lists. In one-way analyses of variance (ANOVA), no significant differences were found in attribution scores for imagery [$F_{(5,119)} = 1.70$, $p = 0.14$, $\eta^2 = 0.07$], concreteness [$F_{(5,119)} = 0.91$, $p = 0.48$, $\eta^2 = 0.04$], ease of learning [$F_{(5,119)} = 0.59$, $p = 0.71$, $\eta^2 = 0.03$], and frequency [$F_{(5,119)} = 0.41$, $p = 0.84$, $\eta^2 = 0.02$].

Target words were visually presented in a movie clip for approximately 18 min (see **Figure 1**) and were encoded through exchanges with four unfamiliar persons, two of whom were

TABLE 1 | Participant characteristics.

	Young (SD)	Old (SD)	Group difference
Age (years)	22.45 (1.72)	65.96 (3.09)	Young < Old**
Sex (male:female)	14:15	13:11	
Education	16.14 (1.43)	13.33 (2.04)	Young > Old**
MoCA-J	28.41 (1.38)	26.04 (2.05)	Young > Old**
FLANDERS	9.83 (0.47)	9.88 (0.45)	n.s.
CES-D	8.69 (4.31)	7.58 (3.56)	n.s.

*SD, standard deviation. FLANDERS, Japanese version of FLANDERS handedness questionnaire; MoCA-J, Japanese version of Montreal Cognitive Assessment; CES-D, Center for Epidemiologic Studies Depression Scale. **p < 0.01.*

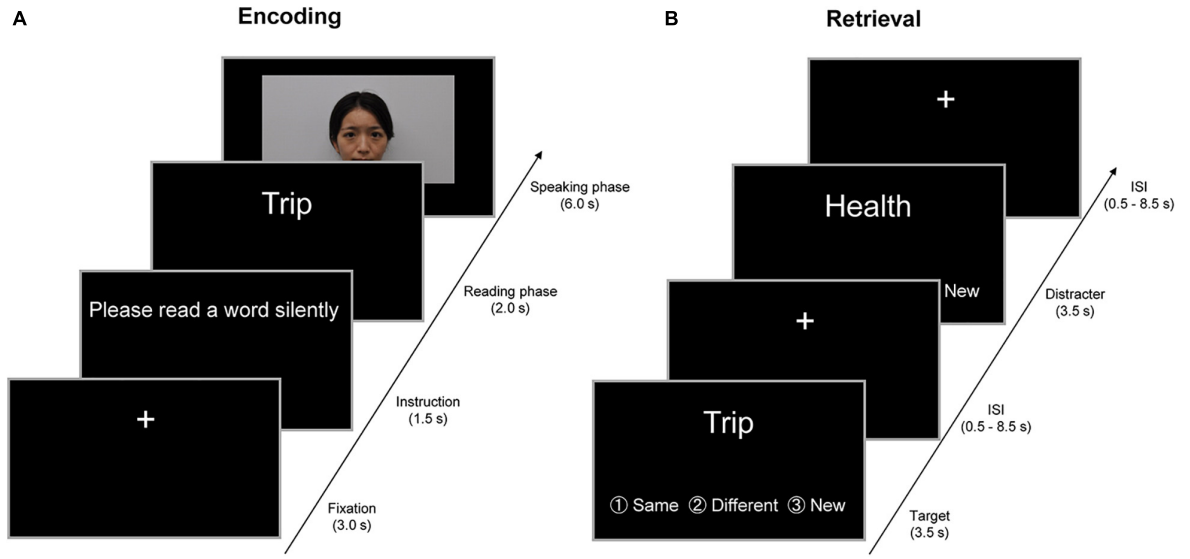

FIGURE 1 | Example of the encoding and retrieval trials in a 2-day experiment. **(A)** Example of the encoding trials on the first day outside fMRI. In a trial of the encoding task, the instruction "Please read a word silently" was visually presented on a PC display for 1.5 s. In the subsequent "reading" phase, participants were visually presented with a target word for 2 s and were required to read the target word silently. In the "speaking" phase for 6 s, participants were randomly presented with a person belonging to the same age group (SAG) or different age group (DAG) in a movie clip. In this phase, participants initially watched the person in the movie and spoke the target word to the person. Subsequently, the person in the movie clip orally returned the same word to participants, and then participants spoke the same word again to the person in the movie clip. In the intertrial interval, a visual fixation was presented for 3 s at the center of the PC display. **(B)** Examples of the retrieval trials on the second day inside fMRI. In the retrieval task, participants were randomly presented with 80 learned and 40 new words one by one on an MRI-compatible PC display, and were required to judge whether each word had been learned from either SAG or DAG members or neither in the encoding task. Three response options (same generation, different generation, and new) were provided for the judgment. Each word was presented for 3.5 s, and then a visual fixation was shown during the interstimulus interval (ISI), which was jittered with variable durations (0.5–8.5 s). All verbal items were presented in Japanese. English is used here for illustration purposes only.

young adults and the other two of whom were older adults. To serve as a young adult in the movie clip, four graduate students (two females and two males) were recruited from the Kyoto University community, and four older adults (two females and two males) who were living outside the city of Kyoto appeared in the movie clip. To avoid sex effects on memory processes, target words were exchanged in the movie clip with persons of the same sex as the participants. The experimenter confirmed that the four persons in the movie clip were not familiar to the participants before the encoding task.

Experimental Procedures

All young and older adults participated in a 2-day experiment in which participants were required to perform the encoding task outside the fMRI scanner on the first day, and then neural activation was measured in the retrieval task inside the scanner approximately 24 h after the encoding task (see **Figure 1**). No reference was made to a subsequent memory test during the encoding task, and hence, the encoding operation was incidental.

In the encoding task on the first day, the participants were visually presented with target words one by one by unfamiliar persons in the movie clip and were required to exchange each word with either a person who belonged to SAG or DAG. In a trial of the encoding task, the initial instruction "Please read a word silently" was visually presented on a PC display for 1.5 s. In the subsequent "reading" phase, the participants were visually presented with a target word for 2 s and were required to read the target word silently. In the "speaking" phase for 6 s, the participants were randomly presented with a SAG or DAG person in the movie clip. In this phase, the participants initially watched the person in the movie and spoke the target word to the person. Subsequently, the person in the movie clip orally returned the same word to the participants, and then participants spoke the same word again to the person in the movie clip. By these experimental operations, we created the virtual exchange of target words with the SAG and DAG people during the encoding task. In the intertrial interval (ITI), a visual fixation was presented for 3 s at the center of the PC display. All experimental procedures during the encoding task were recorded by a video camera.

In the retrieval task on the second day, which was performed in the MRI scanner approximately 24 h after the encoding task, the participants were randomly presented with 120 words one by one, including 80 target words and 40 new words, on an MRI-compatible PC display. The participants were required to judge whether each word was learned from the SAG or DAG person or not learned in the encoding task. Three response options (same generation, different generation, and new) were provided for the judgment, and the participants were instructed to show their response by pressing one of three buttons as soon as possible. Each word was presented for 3.5 s, and then a visual fixation was shown during the interstimulus interval (ISI), which

was jittered with variable durations (0.5–8.5 s). The retrieval task was carried out in one run (approximately 12 min), and retrieval-related activation was measured by the event-related fMRI method. After the retrieval task, the participants evaluated their subjective feelings of attractiveness, trustworthiness, and familiarity of the two SAG and two DAG persons shown in the encoding task on a visual analog scale (VAS) that was 10 cm in length. The results of the VAS evaluations are summarized in the **Supplementary Material**.

MRI Acquisition and Analysis

Data Acquisition

All MRI data were acquired by a MAGNETOM Verio 3.0 T MRI scanner (Siemens, Erlangen, Germany), which is located in Kokoro Research Center, Kyoto University. The stimulus presentation and recording of behavioral responses were controlled by MATLAB® programs[1] on a Windows PC. Experimental stimuli during the encoding task were presented on a Windows PC display, and experimental stimuli during the retrieval task were presented on an MRI-compatible display (Nordic Neuro Lab Inc., Bergen, Norway). During the retrieval task in the scanner, the participants viewed stimuli with a mirror attached to the head coil, and their responses during the retrieval task were recorded by three buttons in a four-button fiber-optic response device for the right hand (Current Designs Inc., Philadelphia, PA, United States). A set of earplugs helped reduce scanner noise, and foam pads were used to minimize head motions during scanning.

During the MRI scanning, three directional T1-weighted anatomical images were initially acquired to localize the subsequent functional and structural images. Second, 5-min resting-state functional images (TR = 2,000 ms, TE = 30 ms, flip angle = 87 degrees, FOV = 19.2 cm × 19.2 cm, matrix size = 64 × 64, slice thickness/gap = 3/1 mm, 33 horizontal slices collected by ascending order) and task-related functional images (TR = 2,000 ms, TE = 25 ms, flip angle = 75 degrees, FOV = 22.4 cm × 22.4 cm, matrix size = 64 × 64, slice thickness/gap = 3.5/0 mm, 39 horizontal slices collected by ascending order) were collected using a pulse sequence of gradient-echo echo-planar imaging (EPI), which is sensitive to blood oxygenation level-dependent (BOLD) contrasts. However, resting-state functional images were not included in the analysis of the present study. Finally, high-resolution T1-weighted structural images were obtained by MPRAGE (TR = 2,250 ms, TE = 3.51 ms, FOV = 25.6 cm × 25.6 cm, matrix size = 256 × 256, slice thickness/gap = 1.0/0 mm, 208 horizontal slices).

Preprocessing of fMRI Data

All MRI data were preprocessed and statistically analyzed by Statistical Parametric Mapping 12 (SPM12, Functional Imaging Laboratory, University College London, London, United Kingdom) implemented in MATLAB® (see text footnote 1). Regarding the preprocessing, after discarding the initial four volumes, functional images in a retrieval run were initially corrected for slice timing, and then parameters of head motion were extracted from functional images. Second, high-resolution T1-weighted structural images for each participant were spatially aligned to functional images in the first scan of these functional images by the coregistration method. Third, structural images spatially aligned to functional images were spatially normalized into the tissue probability map (TPM) template in Montreal Neurological Institute (MNI) space, and parameters estimated by this spatial normalization were applied to all functional images (resampled resolution = 3.5 mm × 3.5 mm × 3.5 mm). Finally, normalized functional images were spatially smoothed by a Gaussian kernel of 8-mm FWHM.

Univariate Analysis of fMRI Data

Trials in which a target word was not spoken twice in the encoding task and those in which no response was shown in the retrieval task were excluded from all statistical analyses. Retrieval trials in which both target words and encoding conditions associated with the words were successfully recognized were categorized into Source Hit, and retrieval trials in which target words were successfully recognized but encoding conditions associated with the words were missed or in which target words were not recognized were regarded as Item-Only Hit and Miss. In addition, the Source Hit and Item-Only Hit + Miss trials were subdivided by encoding conditions, that is, trials encoded with the SAG people and those with the DAG people (Source Hit-SAG, Source Hit-DAG, Item-Only Hit + Miss-SAG, Item-Only Hit + Miss-DAG). Retrieval trials showing a response of "same generation" for distracter words were defined as False Alarm in the SAG response (False Alarm-SAG), retrieval trials showing a response of "different generation" for distracter words were categorized as False Alarm in the DAG response (False Alarm-DAG), and retrieval trials showing a "new" response for distracter words were categorized as Correct Rejection.

Statistical analyses of fMRI data were performed first at the individual level and then at the group level. In the individual-level (fixed-effect) analysis, trial-related activation during the retrieval task was modeled by convolving vectors of onset with a canonical hemodynamic response function (HRF) in the context of a general linear model (GLM). In this model, the timing when each word was presented was defined as the onset with an event duration of 0 s. Six parameters related to head motion were also included as confounding variables in this model. Trial-related activation in this model included six experimental conditions decided by encoding condition and retrieval performance (Source Hit-SAG, Source Hit-DAG, Item-Only Hit + Miss-SAG, Item-Only Hit + Miss-DAG, False Alarm-SAG + False Alarm-DAG, Correct Rejection) and one No-Response condition. Activation related to the Source Hit trial was identified by comparing trial-related activation of Source Hit with baseline activation in each condition of SAG and DAG, and the contrasts yielded a *t*-statistic in each voxel. These contrast images (Source Hit-SAG and Source Hit-DAG) were created for each participant.

In the group-level (random-effect) analysis, using two *t*-contrast images (Source Hit-SAG and Source Hit-DAG) obtained in the individual-level analyses, the Source Hit-related activation in each condition of SAG and DAG was analyzed by

[1] www.mathworks.com

a two-way ANOVA with factors of the encoding condition (SAG and DAG) and subgroup (Young and Old). This ANOVA model was created by a flexible factorial design with a subject factor. Two types of statistical analyses were performed in this ANOVA. First, regions reflecting a significant aging effect on Source Hit-SAG vs. Source Hit-DAG were identified in an *F*-contrast of encoding condition-subgroup interaction masked inclusively by a *t*-contrast of [(Source Hit-SAG vs. Source Hit-DAG in Young) vs. (Source Hit-SAG vs. Source Hit-DAG in Old)] ($p < 0.05$). Second, to identify regions showing a significant aging effect on Source Hit-DAG vs. Source Hit-SAG, an *F*-contrast of encoding condition-subgroup interaction was masked inclusively by a *t*-contrast of [(Source Hit-DAG vs. Source Hit-SAG in Young) vs. (Source Hit-DAG vs. Source Hit-SAG in Old)] ($p < 0.05$).

In these analyses, the height threshold at the voxel level was corrected for multiple comparisons in the whole-brain and hypothesis-driven regions of interest (ROI) (FWE, $p < 0.05$). ROIs in the first analysis were set in the dmPFC, precuneus, hippocampus, and right ATL. In addition, the right pSTS, hippocampus, and right ATL ROIs were applied to the second analysis. ROIs in the hippocampus and precuneus were created bilaterally in the AAL ROI package (Tzourio-Mazoyer et al., 2002). The right ATL ROI covered the entire temporal pole (Brodmann's area: BA 38), extending caudally up to the MNI coordinate $y = -21$ but did not include the superior temporal gyrus. In addition, this ROI included the fusiform and inferior temporal gyrus up to the MNI coordinate $y = -25$ and the fusiform gyrus alone up to the MNI coordinate $y = -39$ but did not include BA37 (Binney et al., 2010). The dmPFC ROI was extracted from the bilateral superior frontal and cingulate gyri in the AAL ROI package, which included the MNI coordinate $y = 1$ or more and $z = 20$ or more (Sugimoto and Tsukiura, 2018). The right pSTS ROI was defined as a sphere with a 5-mm radius at the center of a peak voxel in the pSTS (MNI coordinates: $x = 62$, $y = -32$, $z = 0$), which was identified in a previous study (Dodell-Feder et al., 2011).

Functional Connectivity Analysis of fMRI Data

Functional connectivity with the right hippocampus and right ATL regions, which showed significant activation in the univariate analyses, was investigated by a generalized form of context-dependent psychophysiological interaction (gPPI) analysis (McLaren et al., 2012). A one-run GLM, which included six experimental conditions, one No-Response condition, and confounding variables of six parameters related to head motion, was newly estimated in each participant, and seed regions of the right hippocampus and right ATL in this model were identified by volumes of interest (VOI) with a sphere of 5-mm radius centered on the peak voxel (right hippocampus: $x = 29$, $y = -11$, $z = -25$, and right ATL: $x = 43$, $y = 7$, $z = -21$). These VOIs as seed regions were anatomically masked by a sphere ROI of 10-mm radius centered on each peak voxel, and by each ROI of the right hippocampus extracted from the AAL ROI package (Tzourio-Mazoyer et al., 2002) and the right ATL mentioned above.

In the functional connectivity analysis, we employed the gPPI toolbox.[2] This toolbox produces an individual-level (fixed-effect) model with three sets of columns, including: (1) condition-related regressors formed by convolving vectors of condition-related onsets with a canonical HRF; (2) time series of BOLD signals deconvolved from the seed region; and (3) PPI regressors as an interaction between (1) condition-related regressors and (2) time series of BOLD signals. Thus, the model in the present study included (1) condition-related regressors of the Source Hit-SAG, Source Hit-DAG, Item-Only Hit + Miss-SAG, Item-Only Hit + Miss-DAG, False Alarm-SAG + False Alarm-DAG, Correct Rejection, and No-Response condition; (2) BOLD signals of the right hippocampus or right ATL; and (3) PPI regressors of the Source Hit-SAG, Source Hit-DAG, Item-Only Hit + Miss-SAG, Item-Only Hit + Miss-DAG, False Alarm-SAG + Aalse Alarm-DAG, Correct Rejection, and No-Response condition. For each participant, the models for right hippocampal and right ATL seeds were estimated with confounding variables of six parameters related to head motion, and linear contrasts of PPI regressors were extracted as a *t*-contrast in each of Source Hit-SAG and Source Hit-DAG. Regions showing a significant effect in a *t*-contrast of PPI regressors reflected significant functional connectivity with each seed VOI. These *t*-contrasts of PPI regressors obtained in the individual-level analysis were applied to the group-level analysis. The right hippocampal VOI as a seed region was not significantly extracted from 6 young and 6 older adults. Thus, in the right hippocampal seed, PPI regressor contrasts obtained from 23 young and 18 older participants were analyzed in the group-level statistics. In the right ATL seed, significant VOI was not extracted from 3 young and 3 older adults; hence, PPI regressor contrasts identified in 26 young and 21 older adults were applied to the group-level analysis.

In the group-level (random-effect) analysis, *t*-contrast images reflecting significant functional connectivity with seed regions (right hippocampus and right ATL) in the Source Hit trials were compared between young and older adults by two-sample *t*-tests in each encoding condition (one-tailed). In this analysis, the height threshold at the voxel level was corrected for multiple comparisons in the whole-brain and each ROI of the dmPFC, precuneus, right pSTS, hippocampus, and right ATL regions (FWE, $p < 0.05$), which were explained above. Anatomical sites in all analyses were primarily defined using the SPM Anatomy toolbox (Eickhoff et al., 2005, 2006, 2007) and MRIcro.[3]

RESULTS

Behavioral Results

Behavioral results are summarized in **Table 2**. Memory performance based on the Source Hit rate vs. False Alarm rate in the retrieval task was analyzed by a two-way mixed ANOVA with factors of subgroup (Young and Old) and encoding condition

[2]www.nitrc.org/projects/gppi

[3]https://people.cas.sc.edu/rorden/mricro/mricro.html

TABLE 2 | Behavioral results.

	Young		Old	
	SAG (SD)	**DAG (SD)**	**SAG (SD)**	**DAG (SD)**
Accuracy (proportion)				
Source Hit	0.43 (0.12)	0.46 (0.10)	0.40 (0.08)	0.36 (0.11)
False Alarm	0.23 (0.10)	0.23 (0.10)	0.25 (0.11)	0.27 (0.12)
Source Hit vs. False Alarm	0.19 (0.13)	0.23 (0.12)	0.15 (0.14)	0.10 (0.13)
Response time (ms)				
Source Hit	2014.03 (429.18)	1999.74 (451.90)	2060.50 (349.39)	2118.34 (369.07)
False Alarm	2071.14 (441.15)	2151.46 (356.60)	2151.46 (356.60)	2188.76 (369.85)
Number of trials				
Source Hit	18.21 (4.11)	16.90 (4.93)	15.71 (3.29)	14.25 (4.20)
False Alarm	9.24 (3.94)	9.21 (3.74)	9.83 (4.51)	10.54 (4.89)

SAG, same age group; DAG, different age group; SD, standard deviation.

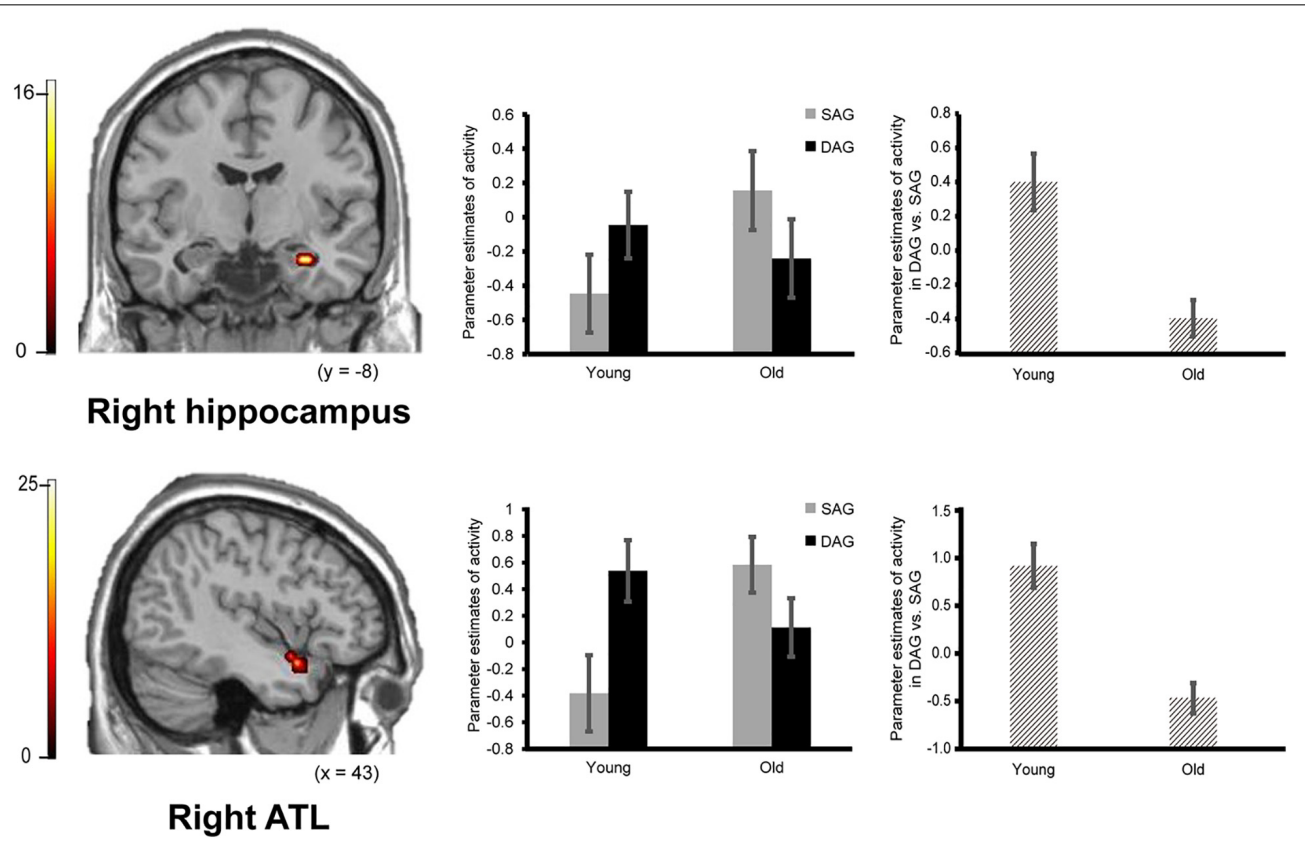

FIGURE 2 | Activation in the right hippocampus (29, −11, −25, $p = 0.026$ corrected by FWE) and right ATL (anterior temporal lobe: 43, 7, −21, $p = 0.003$ corrected by FWE) reflecting a significant interaction between encoding condition and subgroup, in which an age-dependent decrease in activation was identified in a contrast of DAG (different age group) vs. SAG (same age group). Parameter estimates of activity in the bar graphs were extracted from peak voxels in each region. Error bars represent standard errors. The bar graphs in the middle of this figure showed parameter estimates of activity related to the Source Hit trials in each encoding condition, and the bar graphs in the right side showed differences in parameter estimates of activity related to the Source Hit trials between the DAG and SAG conditions. Results in this analysis were identified in fMRI data from 29 young adults and 24 older adults. In this analysis, the height threshold at the voxel level was corrected for multiple comparisons (FWE, $p < 0.05$) in each region of interest (ROI) of the bilateral hippocampi and right ATL which were defined anatomically.

(SAG and DAG). ANOVA demonstrated a significant main effect of subgroup [$F_{(1,51)} = 9.47$, $p < 0.01$, $\eta_p{}^2 = 0.43$] and a significant interaction between subgroup and encoding condition [$F_{(1,51)} = 4.56$, $p < 0.05$, $\eta_p{}^2 = 0.30$]. However, a main effect of encoding condition was not significant [$F_{(1,51)} = 0.22$, $p = 0.64$, $\eta_p{}^2 = 0.07$]. *Post hoc* tests by the Bonferroni method revealed a significant difference between Young and Old in the DAG condition ($p < 0.05$).

Response time (RT) data (ms) during the Source Hit trials in the retrieval task were analyzed by a two-way mixed ANOVA with factors of subgroup (Young and Old) and encoding condition (SAG and DAG). There were no significant main effects of subgroup [$F_{(1,51)} = 0.58$, $p = 0.45$, $\eta_p{}^2 = 0.10$] and encoding condition [$F_{(1,51)} = 0.54$, $p = 0.47$, $\eta_p{}^2 = 0.10$]. The interaction between these factors was also not significant [$F_{(1,51)} = 1.47$, $p = 0.23$, $\eta_p{}^2 = 0.17$].

fMRI Results

Univariate Analysis of fMRI Data

Confirming our first prediction, the right hippocampus and right ATL showed age-dependent decreased activation during the successful retrieval of source memories encoded with the DAG people compared to those with the SAG people. However, age-dependent decreased activation in the successful retrieval of source memories related to the SAG people compared to the DAG people was not identified in any region (see **Figure 2**).

Successful retrieval activation of source memories was analyzed by a two-way ANOVA with subgroup (Young and Old) and encoding condition (SAG and DAG) as factors. In the first analysis, in which we examined regions showing age-dependent decreased activation in the successful retrieval of source memories encoded with the SAG people compared to the DAG people, an *F*-contrast reflecting a significant interaction between these factors was inclusively masked by a *t*-contrast of [(Source Hit-SAG vs. Source Hit-DAG in Young) vs. (Source Hit-SAG vs. Source Hit-DAG in Old)]. In this analysis, however, significant activation was not found in any regions in either whole-brain or ROI-based analyses. In the second analysis, which was conducted to find regions showing age-dependent decreased activation in the successful retrieval of source memories encoded with the DAG people compared to the SAG people, an *F*-contrast reflecting a significant interaction between these factors was inclusively masked by a *t*-contrast of [(Source Hit-DAG vs. Source Hit-SAG in Young) vs. (Source Hit-DAG vs. Source Hit-SAG in Old)]. This analysis of the predefined ROIs showed significant activation in the right hippocampus and right ATL (see **Figure 2**). In the whole-brain analysis, we did not find significant activation in any regions. Detailed results of the univariate analyses are summarized in **Table 3**.

Functional Connectivity Analysis of fMRI Data

Confirming our second prediction, functional connectivity of the right hippocampus or the right ATL with the right pSTS during the successful retrieval of source memories encoded with the DAG people significantly decreased in the Old subgroup compared to the Young subgroup (see **Figure 3**). However, functional connectivity between the hippocampus/ATL and the CMS regions during the successful retrieval of source memories encoded with the SAG people was not different between the Young and Old subgroups.

In the functional connectivity analysis with the right hippocampal and ATL seeds, in which significant activation was identified in the univariate analysis, age-dependent decreases in functional connectivity in the Source Hit trials were investigated by two-sample *t*-tests between the Young and Old subgroups in each condition (one-tailed). In the DAG condition, an age-dependent decrease in functional connectivity with the right hippocampal seed was found in the right middle temporal gyrus (pSTS) in the ROI-based analysis. However, the whole-brain analysis in this condition did not show any regions reflecting significant functional connectivity with the right hippocampal seed. The functional connectivity analysis with the right ATL seed in the DAG condition showed a significant aging effect on the right middle temporal gyrus (pSTS) in the ROI-based analysis and on the left calcarine region in the whole-brain analysis. In the SAG condition, we did not find any regions showing significant age-dependent decrease in functional connectivity with the right hippocampal seed or the right ATL seed in either whole-brain or ROI-based analyses. Detailed results of the functional connectivity analyses are summarized in **Table 4**.

TABLE 3 | Regions reflecting significant interactions between factors of encoding condition and subgroup.

			MNI coordinates				
Regions	**L/R**	**BA**	***x***	***y***	***z***	***Z* score**	**k**
Age-related decrease in SAG vs. DAG							
ROI-based analysis (dmPFC, precuneus, ATL, hippocampus)							
No significant activation was identified in any ROIs.							
Whole-brain analysis							
No significant activation was identified.							
Age-related decrease in DAG vs. SAG							
ROI-based analysis (pSTS, ATL, hippocampus)							
Hippocampus	R		29	−11	−25	3.61	3
ATL	R	38	43	7	−21	4.36	10
Whole-brain analysis							
No significant activation was identified.							

BA, Brodmann area; k, cluster size; L, left; R, right; ROI, region of interest; ATL, anterior temporal lobe; dmPFC, dorsomedial prefrontal cortex; pSTS, posterior superior temporal sulcus.

DISCUSSION

Two major findings emerged from the present study. First, the right hippocampus and right ATL showed an age-dependent decrease in activation during the successful retrieval of source memories encoded with the DAG people as outgroup members compared to the SAG people as ingroup members. In the successful retrieval of source memories encoded with the SAG people as ingroup members compared to the DAG people as outgroup members, however, an age-dependent decrease in activation was not identified in any region. Second, an age-dependent decrease in functional connectivity of the right hippocampus or the right ATL with the right pSTS was significant during the successful retrieval of source memories encoded with the DAG people as outgroup members, whereas a significant aging effect on functional connectivity with the right hippocampus or the right ATL was not found in any region during the successful retrieval of source memories encoded with the SAG people as ingroup members. These findings suggest that functional networks including the hippocampus related

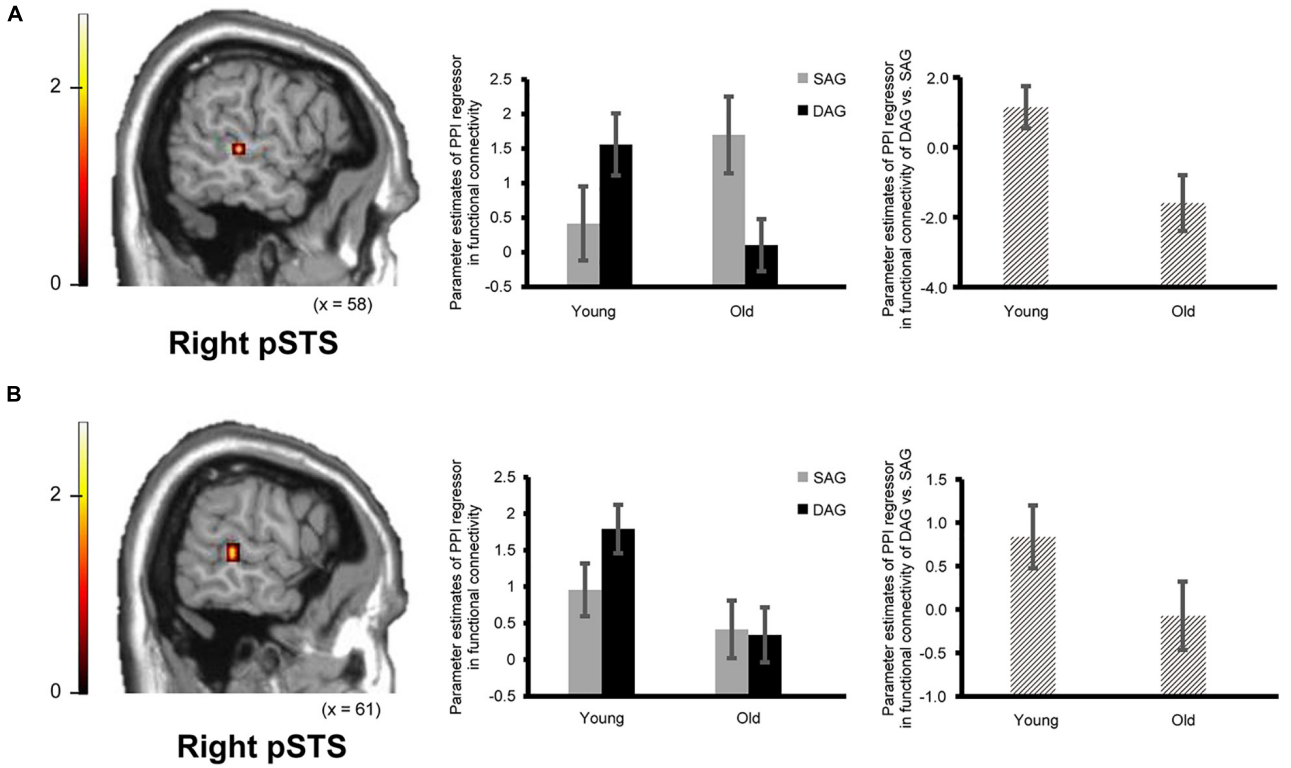

FIGURE 3 | Regions showing an age-dependent decrease in functional connectivity during the successful retrieval of source memories in the DAG conditions. **(A)** The right pSTS (posterior superior temporal sulcus: 61, –32, 0, $p = 0.045$ corrected by FWE) showing an age-dependent decrease in functional connectivity with the right hippocampal seed. Results in this analysis were identified in fMRI data from 23 young adults and 18 older adults, from whom a seed volume of interest (VOI) was extracted. **(B)** The right pSTS (64, –35, 4, $p = 0.024$ corrected by FWE) showing an age-dependent decrease in functional connectivity with the right ATL (anterior temporal lobe) seed. Results in this analysis were identified in fMRI data from 26 young adults and 21 older adults, from whom a seed VOI was extracted. Parameter estimates of PPI regressors in the bar graphs were extracted from peak voxels in each region of the pSTS showing significant functional connectivity. Error bars represent standard errors. The bar graphs in the middle of this figure showed parameter estimates of PPI regressors related to the Source Hit trials in each encoding condition, and the bar graphs in the right side showed differences in parameter estimates of PPI regressors related to the Source Hit trials between the DAG and SAG conditions. In these analyses, the height threshold at the voxel level was corrected for multiple comparisons (FWE, $p < 0.05$) in each ROI of the right pSTS, which was defined anatomically.

to the recollection process of episodic memories, the ATL related to the processing of social knowledge, and the pSTS related to the processing of other people could be impaired by aging in the recollection of memories associated with the DAG people as outgroup members. These findings are discussed in separate sections below.

Age-Dependent Decreased Activation During the Recollection of Memories Associated With Outgroup Members Categorized by Age-Group Membership

The first main finding of the present study was that activation in the right hippocampus and right ATL significantly decreased by aging during the successful recollection of source memories encoded with the DAG people as outgroup members compared to the SAG people as ingroup members. This finding suggests that the hippocampus and ATL, which contribute to the retrieval of associations between face and person-related social knowledge, are more involved in the recollection of memories associated with the DAG people as outgroup members than the recollection of memories associated with the SAG people as ingroup members and that the functionality of these regions is impaired by aging.

In the present study, we found that the right hippocampus and right ATL showed an age-dependent decrease in activation during the retrieval of source memories encoded with the DAG people as outgroup members compared to the SAG people as ingroup members. This finding is consistent with previous studies, which have demonstrated that recollection-related hippocampal activation significantly decreases with age (Shing et al., 2010; Koen and Yonelinas, 2014). For example, one fMRI study reported that activation in the hippocampus related to recollection significantly decreased in older adults compared to young adults, whereas activation related to familiarity in the entorhinal and perirhinal cortices was relatively preserved in older adults (Daselaar et al., 2006). In another fMRI study, an age-dependent decrease in hippocampal activation was identified in the recollection of autobiographical memories (St. Jacques et al., 2012). In addition, ATL activation

TABLE 4 | Regions showing age-related decrease in functional connectivity in each encoding condition.

Regions	L/R	BA	MNI coordinates			*Z* score	k
			x	*y*	*z*		
SAG condition							
Right hippocampus seed							
ROI-based analysis (dmPFC, precuneus, ATL)							
No significant activation was identified in any ROIs.							
Whole-brain analysis							
No significant activation was identified.							
Right ATL seed							
ROI-based analysis (dmPFC, precuneus, hippocampus)							
No significant activation was identified in any ROIs.							
Whole-brain analysis							
No significant activation was identified.							
DAG condition							
Right hippocampus seed							
ROI-based analysis (pSTS, ATL)							
Middle temporal gyrus	R		61	−32	0	2.35	1
Whole-brain analysis							
No significant activation was identified.							
Right ATL seed							
ROI-based analysis (pSTS, hippocampus)							
Middle temporal gyrus	R		64	−35	4	2.61	2
Whole-brain analysis							
Calcarine cortex	L		−24	−53	11	5.05	2

BA, Brodmann area; k, cluster size; L, left; R, right; ROI, region of interest; ATL, anterior temporal lobe; dmPFC, dorsomedial prefrontal cortex; pSTS, posterior superior temporal sulcus.

in the present study is consistent with a previous fMRI study that reported an age-dependent decline in correlations between hippocampal and ATL activation during the retrieval of associations between face and face-related social knowledge, including people's name and job title (Tsukiura et al., 2011). There is also functional neuroimaging evidence that ATL activation reflects the successful retrieval of person-related semantics as a type of social knowledge (Wang et al., 2017). The importance of the ATL in the processing of social knowledge has been consistently identified in functional neuroimaging studies and neuropsychological studies with brain-damaged patients (for review, see Olson et al., 2007; Olson et al., 2013). In the present study, the right hippocampus and right ATL showed significantly greater activation during the successful recollection of memories associated with the DAG people as outgroup members than that of memories associated with the SAG people as ingroup members, and activation in these regions significantly decreased by aging. Given our behavioral data that older adults showed worse retrieval performance than young adults only in the DAG condition, the hippocampus and ATL could play an important role in forming associations between face and social knowledge in the recollection of memories encoded with an outgroup member, and activation in these regions could be impaired by aging.

Age-Related Differences in Functional Connectivity During the Recollection of Memories Associated With Outgroup Members Categorized by Age-Group Membership

The second main finding of the present study was that an age-dependent decrease in functional connectivity was found between the hippocampus/ATL and the pSTS during the successful retrieval of source memories in the DAG condition. The finding suggests that functional networks including the hippocampus related to the retrieval of source memories, the pSTS related to social cognition of others, and the ATL related to the processing of social knowledge contribute to the recollection of memories associated with the DAG people as outgroup members, and that this mechanism related to the recollection of memories associated with outgroup members could decrease as a result of aging.

In the present study, functional connectivity of the right hippocampus or the right ATL with the right pSTS was significantly lower in older adults than in young adults during the successful retrieval of source memories encoded with the DAG people as outgroup members. This finding might be explained by evidence from previous studies showing that social cognition networks, including the pSTS, anterior cingulate cortex, ATL and temporal-parietal junction (TPJ), contribute to the mentalizing of others' mental states (for review, see Gallagher and Frith, 2003; Do and Telzer, 2019). For example, one fMRI study demonstrated that the ATL region, which showed greater activation during face processing than during scene or object processing, was functionally connected with the pSTS in resting-state fMRI scanning (Simmons et al., 2010). In another study, activation in the pSTS was modulated by a change in facial identity or expression (Fox et al., 2009). In addition, there is cognitive neuroscience evidence that the ATL is involved in the processing of person-related social knowledge (for review, see Olson et al., 2013), and that an age-dependent decrease in interacting mechanisms between the ATL and hippocampus are identified in the retrieval of social knowledge (name or job title) associated with faces (Tsukiura et al., 2011). Thus, functional connectivity between the ATL and pSTS in the DAG condition of the present study could reflect the mentalizing of others' mental states by referring to social knowledge about the DAG people, and an age-dependent decrease in functional networks including the hippocampus, ATL and pSTS suggests that the recollection of memory for the DAG people by the mentalizing of others' mental states with reference to social knowledge could be impaired in older adults.

No Aging Effect on Activation and Functional Connectivity During the Recollection of Memories Associated With Ingroup Members Categorized by Age-Group Membership

Inconsistent with our prediction, an age-dependent decrease in activation and functional connectivity was not identified in the

successful recollection of source memories encoded with the SAG people as ingroup members compared to the DAG people as outgroup members. This finding implies that the intergroup bias in memory is relatively preserved in older adults.

There is psychological evidence in which intergroup bias, such as the own-age bias, is commonly observed between young and older adults in several measures, including ingroup favoritism, perceived similarity, social distance, outgroup homogeneity, and self-stereotyping (Chasteen, 2005). The preserved own-age bias in older adults was also found in episodic memory research in which both young and older adults showed a significant own-age bias in name recall with better memory for names associated with faces of their own age, as compared to other-aged faces (Strickland-Hughes et al., 2020). In addition, one fMRI study reported that the amygdala showed significant activation during perceiving faces of the SAG people compared to the DAG people in both young and older groups (Wright et al., 2008). Behavioral data in the present study revealed that the Source Hit responses in the DAG condition were significantly lower in older adults than in young adults, whereas the Source Hit responses in the SAG condition were not different between young and older adults. Thus, the present finding that an age-dependent decrease in activation and functional connectivity during the successful recollection of memories in the SAG condition compared to the DAG condition was not identified in any regions could reflect the preserved intergroup bias related to the own-age bias during the recollection of source memories in older adults.

CONCLUSION

In the present study, using event-related fMRI, we investigated age-related differences in neural mechanisms underlying the successful recollection of memories encoded with the SAG people as ingroup members and with the DAG people as outgroup members. The results demonstrated that activation in the right hippocampus and right ATL significantly decreased with age during the successful source retrieval of memories encoded with the DAG people as outgroup members compared to the SAG people as ingroup members. However, no region showed age-dependent decreased activation during the successful source retrieval of memories associated with the SAG people as ingroup members compared to the DAG people as outgroup members. In addition, an age-dependent decrease in functional connectivity with the hippocampus or the ATL was significantly identified in the pSTS in the DAG condition, whereas functional connectivity with the hippocampus or the ATL in the SAG condition showed no significant aging effect on any regions. These findings suggest that functional networks including the hippocampus, ATL, and pSTS, which contribute to the recollection of memories for the DAG people by the mentalizing of others' mental states with reference to social knowledge, could be impaired in older adults compared to young adults, whereas neural mechanisms underlying the intergroup bias during the recollection of source memories for the SAG people could be relatively preserved in older adults.

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 Institutional Review Board (IRB) for Human Information Research of the Graduate School of Human and Environmental Studies, Kyoto University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

ET and TT designed the study and analyzed the data. ET conducted the fMRI experiment and wrote the original draft of the manuscript. TT reviewed and edited the manuscript and obtained funding for the experiment. Both authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the JSPS KAKENHI Grant Numbers JP16H06325 and JP20H05802. The funder was not involved in the study design, data collection, data analysis, interpretation of data, preparation of the manuscript, or decision to submit it for publication.

ACKNOWLEDGMENTS

We would like to thank Nobuhito Abe, Kohei Asano, and Aiko Murai for technical assistance in the MRI scanning, and Saeko Iwata and Reina Izumika for supporting the data analysis. The research experiments were conducted using an MRI scanner and related facilities at Kokoro Research Center, Kyoto University.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnbeh.2021.743064/full#supplementary-material

REFERENCES

Binney, R. J., Embleton, K. V., Jefferies, E., Parker, G. J. M., and Ralph, M. A. L. (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. *Cereb. Cortex* 20, 2728–2738. doi: 10.1093/cercor/bhq019

Chasteen, A. L. (2005). Seeing eye-to-eye: do intergroup biases operate similarly for younger and older adults? *Int. J. Aging Hum. Dev.* 61, 123–139. doi: 10.2190/07Q7-BWYT-NC9E-51FX

Cosseddu, M., Gazzina, S., Borroni, B., Padovani, A., and Gainotti, G. (2018). Multimodal face and voice recognition disorders in a case with unilateral right anterior temporal lobe atrophy. *Neuropsychology* 32, 920–930. doi: 10.1037/neu0000480

Costa, P. T., and MarcCrae, R. R. (1992). *Revised NEO Personality (NEO I-R) and NEO Five-Factory (NEO-FFI): Professional Manual*. Odessa, FL: Psychological Assessment Resources.

Daselaar, S. M., Fleck, M. S., Dobbins, I. G., Madden, D. J., and Cabeza, R. (2006). Effects of healthy aging on hippocampal and rhinal memory functions: an event-related fMRI study. *Cereb. Cortex* 16, 1771–1782. doi: 10.1093/cercor/bhj112

Dasgupta, S., Tyler, S. C., Wicks, J., Srinivasan, R., and Grossman, E. D. (2017). Network connectivity of the right STS in three social perception localizers. *J. Cogn. Neurosci.* 29, 221–234. doi: 10.1162/jocn_a_01054

Do, K. T., and Telzer, E. H. (2019). Corticostriatal connectivity is associated with the reduction of intergroup bias and greater impartial giving in youth. *Dev. Cogn. Neurosci.* 37:100628. doi: 10.1016/j.dcn.2019.100628

Dodell-Feder, D., Koster-Hale, J., Bedny, M., and Saxe, R. (2011). fMRI item analysis in a theory of mind task. *Neuroimage* 55, 705–712. doi: 10.1016/j.neuroimage.2010.12.040

Eickhoff, S. B., Heim, S., Zilles, K., and Amunts, K. (2006). Testing anatomically specified hypotheses in functional imaging using cytoarchitectonic maps. *Neuroimage* 32, 570–582. doi: 10.1016/j.neuroimage.2006.04.204

Eickhoff, S. B., Paus, T., Caspers, S., Grosbras, M. H., Evans, A. C., Zilles, K., et al. (2007). Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. *Neuroimage* 36, 511–521. doi: 10.1016/j.neuroimage.2007.03.060

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., et al. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25, 1325–1335. doi: 10.1016/j.neuroimage.2004.12.034

Fox, C. J., Moon, S. Y., Iaria, G., and Barton, J. J. S. (2009). The correlates of subjective perception of identity and expression in the face network: an fMRI adaptation study. *Neuroimage* 44, 569–580. doi: 10.1016/j.neuroimage.2008.09.011

Fujiwara, Y., Suzuki, H., Yasunaga, M., Sugiyama, M., Ijuin, M., Sakuma, N., et al. (2010). Brief screening tool for mild cognitive impairment in older Japanese: validation of the Japanese version of the Montreal Cognitive Assessment. *Geriatr. Gerontol. Int.* 10, 225–232. doi: 10.1111/j.1447-0594.2010.00585.x

Gainotti, G. (2013). Is the right anterior temporal variant of prosopagnosia a form of 'associative prosopagnosia' or a form of 'multimodal person recognition disorder'? *Neuropsychol. Rev.* 23, 99–110. doi: 10.1007/s11065-013-9232-7

Gallagher, H. L., and Frith, C. D. (2003). Functional imaging of 'theory of mind'. *Trends Cogn. Sci.* 7, 77–83. doi: 10.1016/S1364-6613(02)00025-6

Golby, A. J., Gabrieli, J. D. E., Chiao, J. Y., and Eberhardt, J. L. (2001). Differential responses in the fusiform region to same-race and other-race faces. *Nat. Neurosci.* 4, 845–850. doi: 10.1038/90565

Greven, I. M., and Ramsey, R. (2017). Neural network integration during the perception of in-group and out-group members. *Neuropsychologia* 106, 225–235. doi: 10.1016/j.neuropsychologia.2017.09.036

Hitokoto, H., and Uchida, Y. (2015). Interdependent happiness: theoretical importance and measurement validity. *J. Happiness Stud.* 16, 211–239. doi: 10.1007/s10902-014-9505-8

Ito, Y., Sagara, J., Ikeda, M., and Kawaura, Y. (2003). Reliability and validity of subjective well-being scale. *Shinrigaku Kenkyu* 74, 276–281. doi: 10.4992/jjpsy.74.276

Itsukushima, Y., Ishihara, O., Nagata, Y., and Koike, Y. (1991). Research of two–Chinese character word attributes: imagery, concreteness, and ease of learning. *Psychol. Res. Nihon Univ.* 12, 1–19.

Koen, J. D., and Yonelinas, A. P. (2014). The effects of healthy aging, amnestic mild cognitive impairment, and Alzheimer's disease on recollection and familiarity: a meta-analytic review. *Neuropsychol. Rev.* 24, 332–354. doi: 10.1007/s11065-014-9266-5

Masuda, Y., Tadaka, E., and Dai, Y. (2012). Reliability and validity of the Japanese version of the UCLA loneliness scale version 3 among the older population. *Japan Acad. Community Health Nurs.* 15, 25–32.

Mcadams, D. P., and Destaubin, E. (1992). A Theory of generativity and its assessment through self-report, behavioral acts, and narrative themes in autobiography. *J. Pers. Soc. Psychol.* 62, 1003–1015. doi: 10.1037/0022-3514.62.6.1003

McLaren, D. G., Ries, M. L., Xu, G. F., and Johnson, S. C. (2012). A generalized form of context-dependent psychophysiological interactions (gPPI): a comparison to standard approaches. *Neuroimage* 61, 1277–1286. doi: 10.1016/j.neuroimage.2012.03.068

Mimura, C., and Griffiths, P. (2007). A Japanese version of the Rosenberg Self-Esteem Scale: translation and equivalence assessment. *J. Psychosom. Res.* 62, 589–594. doi: 10.1016/j.jpsychores.2006.11.004

Molenberghs, P., and Louis, W. R. (2018). Insights from fMRI studies into ingroup bias. *Front. Psychol.* 9:1868. doi: 10.3389/fpsyg.2018.01868

Murray, R. J., Debbane, M., Fox, P. T., Bzdok, D., and Eickhoff, S. B. (2015). Functional connectivity mapping of regions associated with self- and other-processing. *Hum. Brain Mapp.* 36, 1304–1324. doi: 10.1002/hbm.22703

Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., et al. (2005). The montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* 53, 695–699. doi: 10.1111/j.1532-5415.2005.53221.x

Nicholls, M. E. R., Thomas, N. A., Loetscher, T., and Grimshaw, G. M. (2013). The Flinders Handedness survey (FLANDERS): a brief measure of skilled hand preference. *Cortex* 49, 2914–2926. doi: 10.1016/j.cortex.2013.02.002

Northoff, G., Heinzel, A., Greck, M., Bennpohl, F., Dobrowolny, H., and Panksepp, J. (2006). Self-referential processing in our brain–a meta-analysis of imaging studies on the self. *Neuroimage* 31, 440–457. doi: 10.1016/j.neuroimage.2005.12.002

Nyberg, L. (2017). Functional brain imaging of episodic memory decline in ageing. *J. Intern. Med.* 281, 65–74. doi: 10.1111/joim.12533

Okubo, M., Suzuki, H., and Nicholls, M. E. (2014). A Japanese version of the FLANDERS handedness questionnaire. *Jpn. J. Psychol.* 85, 474–481. doi: 10.4992/jjpsy.85.13235

Olson, I. R., McCoy, D., Klobusicky, E., and Ross, L. A. (2013). Social cognition and the anterior temporal lobes: a review and theoretical framework. *Soc. Cogn. Affect. Neurosci.* 8, 123–133. doi: 10.1093/scan/nss119

Olson, I. R., Ploaker, A., and Ezzyat, Y. (2007). The enigmatic temporal pole: a review of findings on social and emotional processing. *Brain* 130, 1718–1731. doi: 10.1093/brain/awm052

Patel, G. H., Sestieri, C., and Corbetta, M. (2019). The evolution of the temporoparietal junction and posterior superior temporal sulcus. *Cortex* 118, 38–50. doi: 10.1016/j.cortex.2019.01.026

Pitcher, D., Dilks, D. D., Saxe, R. R., Triantafyllou, C., and Kanwisher, N. (2011). Differential selectivity for dynamic versus static information in face-selective cortical regions. *Neuroimage* 56, 2356–2363. doi: 10.1016/j.neuroimage.2011.03.067

Polosan, M., Baciu, M., Cousin, E., Perrone, M., Pichat, C., and Bougerol, T. (2011). An fMRI study of the social competition in healthy subjects. *Brain Cogn.* 77, 401–411. doi: 10.1016/j.bandc.2011.08.018

Radloff, L. S. (1977). The CES-D scale: a self-report depression scale for research in the general population. *Appl. Psychol. Meas.* 1, 385–401. doi: 10.1177/014662167700100306

Rhodes, M. G., and Anastasi, J. S. (2012). The own-age bias in face recognition: a meta-analytic and theoretical review. *Psychol. Bull.* 138, 146–174. doi: 10.1037/a0025750

Rilling, J. K., Dagenais, J. E., Goldsmith, D. R., Glenn, A. L., and Pagnoni, G. (2008). Social cognitive neural networks during in-group and out-group interactions. *Neuroimage* 41, 1447–1461. doi: 10.1016/j.neuroimage.2008.03.044

Rosenberg, N. (1965). *Society and the Adolescent Self-Image*. Princeton, NJ: Princeton University Press.

Ross, L. A., Mccoy, D., Coslett, H. B., Olson, I. R., and Wolk, D. A. (2011). Improved proper name recall in aging after electrical stimulation of the anterior temporal lobes. *Front. Aging Neurosci.* 3:16. doi: 10.3389/fnagi.2011.00016

Russell, D., Peplau, L. A., and Ferguson, M. L. (1978). Developing a measure of loneliness. *J. Pers. Assess.* 42, 290–294. doi: 10.1207/s15327752jpa4203_11

Scheepers, D., Derks, B., Nieuwenhuis, S., Lelieveld, G. J., Van Nunspeet, F., Rombouts, S. A. R. B., et al. (2013). The neural correlates of in-group and self-face perception: is there overlap for high identifiers? *Front. Hum. Neurosci.* 7:528. doi: 10.3389/fnhum.2013.00528

Sell, H., and Nagpal, R. (1992). *Assessment of Subjective Well-Being: The Subjective Well-Being Inventory (SUBI)*. New Delhi: World Health Organization Regional Office for South-East Asia.

Shima, S. (1998). *The CES-D Scale Japanese Edition: User's Guide*. Tokyo: Chiba Test Center.

Shimonaka, Y., Nakazato, K., Gondo, Y., and Takayama, M. (1999). *NEO-PI-R, NEO-FFI Manual for the Japanese Version Big5 Personality Inventory*. Tokyo: Tokyo Shinri.

Shing, Y. L., Werkle-Bergner, M., Brehmer, Y., Muller, V., Li, S. C., and Lindenberger, U. (2010). Episodic memory across the lifespan: the contributions of associative and strategic components. *Neurosci. Biobehav. Rev.* 34, 1080–1091. doi: 10.1016/j.neubiorev.2009.11.002

Simmons, W. K., Reddish, M., Bellgowan, P. S., and Martin, A. (2010). The selectivity and functional connectivity of the anterior temporal lobes. *Cereb. Cortex* 20, 813–825. doi: 10.1093/cercor/bhp149

St. Jacques, P. L., Rubin, D. C., and Cabeza, R. (2012). Age-related effects on the neural correlates of autobiographical memory retrieval. *Neurobiol. Aging* 33, 1298–1310. doi: 10.1016/j.neurobiolaging.2010.11.007

Strickland-Hughes, C. M., Dillon, K. E., West, R. L., and Ebner, N. C. (2020). Own-age bias in face-name associations: evidence from memory and visual attention in younger and older adults. *Cognition* 200:104253. doi: 10.1016/j.cognition.2020.104253

Sugimoto, H., and Tsukiura, T. (2018). Contribution of the medial prefrontal cortex to social memory. *Brain Nerve* 70, 753–761. doi: 10.11477/mf.1416201076

Tabuchi, M., Nakagawa, T., Gondo, Y., and Komori, M. (2012). Development of the short version of generativity scale in old age. *J. Health Welf. Stat.* 59, 1–7.

Tromp, D., Dufour, A., Lithfous, S., Pebayle, T., and Despres, O. (2015). Episodic memory in normal aging and Alzheimer disease: insights from imaging and behavioral studies. *Ageing Res. Rev.* 24(Pt B), 232–262. doi: 10.1016/j.arr.2015.08.006

Tsukiura, T., Sekiguchi, A., Yomogida, Y., Nakagawa, S., Shigemune, Y., Kambara, T., et al. (2011). Effects of aging on hippocampal and anterior temporal activations during successful retrieval of memory for face-name associations. *J. Cogn. Neurosci.* 23, 200–213. doi: 10.1162/jocn.2010.21476

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. doi: 10.1006/nimg.2001.0978

Volz, K. G., Kessler, T., and von Cramon, D. Y. (2009). In-group as part of the self: in-group favoritism is mediated by medial prefrontal cortex activation. *Soc. Neurosci.* 4, 244–260. doi: 10.1080/17470910802553565

Wang, Y., Collins, J. A., Koski, J., Nugiel, T., Metoki, A., and Olson, I. R. (2017). Dynamic neural architecture for social knowledge retrieval. *Pro. Natl. Acad. Sci. U.S.A.* 114, E3305–E3314. doi: 10.1073/pnas.1621234114

Wright, C. I., Negreira, A., Gold, A. L., Britton, J. C., Williams, D., and Barrett, L. F. (2008). Neural correlates of novelty and face-age effects in young and elderly adults. *Neuroimage* 42, 956–968. doi: 10.1016/j.neuroimage.2008.05.015

Yamawaki, R., Nakamura, K., Aso, T., Shigemune, Y., Fukuyama, H., and Tsukiura, T. (2017). Remembering my friends: medial prefrontal and hippocampal contributions to the self-reference effect on face memories in a social context. *Hum. Brain Mapp.* 38, 4256–4269. doi: 10.1002/hbm.23662

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.

frontiers in Behavioral Neuroscience

ORIGINAL RESEARCH
published: 08 November 2021
doi: 10.3389/fnbeh.2021.749252

Prefrontal-Striatal Mechanisms of Behavioral Impulsivity During Consumption of Delayed Real Liquid Rewards

*Ayaka Misonou and Koji Jimura**

Department of Biosciences and Informatics, Keio University, Yokohama, Japan

Intertemporal choice involves the evaluation of future rewards and reflects behavioral impulsivity. After choosing a delayed reward in an intertemporal choice, a behavioral agent waits for, receives, and then consumes the reward. The current study focused on the consumption of the delayed reward and examined the neural mechanisms of behavioral impulsivity. In humans consuming delayed real liquid rewards in an intertemporal choice, the ventral striatum (VS) showed differential activity between anterior (aVS) and posterior (pVS) regions depending on the degree of behavioral impulsivity. Additionally, impulsive individuals showed activity in the anterior prefrontal cortex (aPFC). An analysis of task-related effective connectivity based on psychophysiological interaction (PPI) revealed that PPI was robust from the aPFC to pVS, but not in the opposite direction. On the other hand, strong bidirectional PPIs were observed between the aVS and pVS, but PPIs from the pVS to aVS were enhanced in impulsive individuals. These results suggest that behavioral impulsivity is reflected in aPFC-VS mechanisms during the consumption of delayed real liquid rewards.

Keywords: decision-making, reward consumption, delay discounting, primary reward, human

OPEN ACCESS

Edited by:
Yuji Naya,
Peking University, China

Reviewed by:
Kenji Matsumoto,
Tamagawa University, Japan
Takashi Tsukiura,
Kyoto University, Japan

****Correspondence:***
Koji Jimura
koji.jimura@gmail.com

Specialty section:
This article was submitted to Learning and Memory, a section of the Journal Frontiers in Behavioral Neuroscience

Received: *29 July 2021*
Accepted: *20 September 2021*
Published: *08 November 2021*

Citation:
Misonou A and Jimura K (2021) Prefrontal-Striatal Mechanisms of Behavioral Impulsivity During Consumption of Delayed Real Liquid Rewards. Front. Behav. Neurosci. 15:749252. doi: 10.3389/fnbeh.2021.749252

INTRODUCTION

Impulsivity is a behavioral pattern in which a behavioral agent persistently makes choices entailing a failure to achieve a long-term goal (Ainslie, 1975). In impulsive decisions, choices with optimal long-term optimal outcomes are overly discounted (Mischel et al., 1989), whereas those emphasizing short-term outcomes are overvalued (Levy and Glimcher, 2011).

One classical behavioral paradigm to evaluate impulsivity is an intertemporal choice, where a behavioral agent chooses between two alternatives that differ in outcome magnitude and time to the outcome (Rachlin et al., 1991; Keeney and Raiffa, 1993). Individuals choosing smaller rewards that are immediately available exhibit greater discounting of delayed rewards and are characterized as impulsive (Madden and Bickel, 2009). On the other hand, self-controlled (less impulsive) individuals are able to wait for a longer time showing lesser delay discounting to maximize attainment of long-term rewards (Rachlin, 2004).

The ventral striatum (VS) is implicated as a core neural site involved in impulsive decision-making (McClure et al., 2004, 2007; Tanaka et al., 2004, 2020; Jimura et al., 2013), whereas the prefrontal cortex (PFC) is associated with less impulsive (or self-controlled) decision-making (McClure et al., 2004, 2007; Shamosh et al., 2008; Hare et al., 2009; Jimura et al., 2013, 2018; Tanaka et al., 2020). These involvements of the PFC and VS raise the possibility that a PFC-VS mechanism plays an important role in determining impulsivity in value-based decision-making.

Prior human studies examining the neural mechanisms of intertemporal choice behavior have mainly focused on the choice period (McClure et al., 2004, 2007; Tanaka et al., 2004; Hariri et al., 2006; Kable and Glimcher, 2007; Peters and Buechel, 2010), with a few exceptions examining temporal changes in brain activity while future outcomes were anticipated (Berns et al., 2006; Jimura et al., 2013; McGuire and Kable, 2015; Tanaka et al., 2020). However, it remains unclear how the PFC-VS mechanism is involved while a behavioral agent is receiving a reward after having waited for it, despite the collective evidence that the VS and PFC are associated with behavioral impulsivity and self-control.

Intertemporal choice paradigms using delayed real liquid rewards (Jimura et al., 2009, 2011, 2013; Tanaka et al., 2020) could provide a unique opportunity to examine brain mechanisms involved in the direct experiences of delayed rewards. In these paradigms, humans make choices between two alternatives, one larger amount of liquid reward delayed by tens of seconds, and a smaller amount of liquid reward available immediately (**Figure 1A**). After making a choice, the participants immediately experience the delay and then consume the liquid reward (**Figure 1B**). Using functional MRI, we continuously measured brain activity while participants performed the paradigms (Jimura et al., 2013; Tanaka et al., 2020). Whereas our prior fMRI analyses focused on choice and delay periods (Jimura et al., 2013; Tanaka et al., 2020), in the current study, we focused on the consumption period and examined the brain mechanisms underlying impulsive choice by analyzing fMRI data while humans consumed the delayed real liquid rewards (Jimura et al., 2013). We first evaluated head movements and image quality during the drinking period, and then examined brain activity in the VS and PFC. A particular analysis focused on the anterior prefrontal cortex implicated in reward anticipation during the delay period (Jimura et al., 2013; Tanaka et al., 2020), aiming to examine prefrontal-striatum mechanisms consistently involved through entire task events in our intertemporal choice task. Finally, we assessed task-related functional connectivity between the VS and PFC.

MATERIALS AND METHODS

Participants

Participants (N = 43; mean age, 23.0 years; range, 18–35 years; 20 male, 23 female) were right-handed and free from any history of psychiatric or neurological disorders. Each participant provided written informed consent after additional screening for physical or medical conditions that would affect their eligibility for fMRI. The study protocol was approved in accordance with guidelines instituted by the Washington University Human Research Protection Office, and data were collected by the senior author (KJ) at Washington University in St. Louis. Participants were compensated for their participation ($10 per h for the behavioral session, $25 per h for the fMRI session). Of the 45 participants recruited into the study, two were eliminated due to the small number of choices (<10) for the delayed option in the fMRI session.

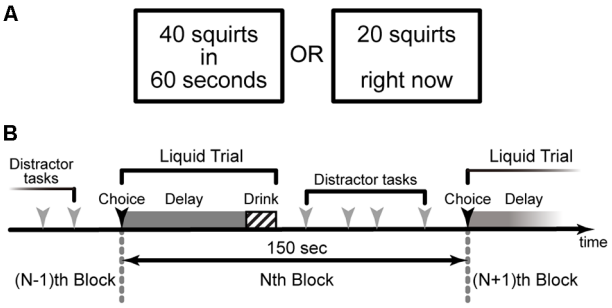

FIGURE 1 | The behavioral paradigm of the intertemporal choice task for a delayed real liquid reward (Jimura et al., 2013). **(A)** Human participants made a choice between a larger amount of liquid available after a delay of 10 s of seconds and a smaller amount of liquid available immediately. **(B)** In each trial, participants consumed the real liquid rewards after experiencing a specified delay.

Dataset

We analyzed a data set collected in an fMRI experiment of an intertemporal choice task involving real liquid rewards where human participants directly experienced choice, delay, and consumption of the rewards (Jimura et al., 2013; **Figure 1**). In this experiment, participants performed the intertemporal decision-making task (**Figure 1**) in two separate (behavioral and fMRI) sessions.

The analyses of the choice and delay periods were published previously (Jimura et al., 2013), and were not analyzed in the current study. The current study analyzed the data while participants consumed the liquid rewards only, which were not analyzed in the previous study (Jimura et al., 2013).

Behavioral Session Procedure

The behavioral session aimed to measure individuals' delay discounting of real liquid rewards (Jimura et al., 2009, 2011, 2013). Prior to the behavioral session, participants were asked to choose one favorite drink that would serve as the reward from a list consisting of apple, orange, grape, grapefruit, and cranberry juices, lemonade, and water. No participants requested to change the reward drink in the fMRI session.

At the beginning of each trial, two alternatives were presented on the left and right sides of the screen, respectively: one involved a larger reward (20 or 40 squirts) available after a delay (10, 30, or 60 s), while the other consisted of a variable smaller amount available immediately (**Figure 1A**). Participants were instructed to press one of two corresponding response buttons to indicate their preference.

During the delay, a fixation cross was presented on the center of the screen. At the time of reward delivery, participants saw a visual message indicating the reward was ready. Importantly, participants were able to control the rate of liquid flow. Reward delivery continued as long as the button was held down; if the button was released, delivery paused and then resumed when the button was pressed again. During reward delivery, the amount remaining (in squirts) was displayed below a red horizontal bar whose length corresponded to the number of squirts still available. After the participant finished drinking, a fixation cross was presented.

To estimate individuals' delay-discounting rates, the current study used three delay conditions (10, 30, 60 s) for the larger amount (40 squirts), and two delay conditions (10, 30 s) for the smaller amount (20 squirts; Jimura et al., 2009). On the first trial of each delay condition, the choice was between a larger delayed amount and an immediate reward that was half of the delayed amount. For each delay condition, the amount of the immediate reward after the first trial was adjusted based on the participant's preceding choice. If the participant had chosen the smaller, immediate reward on the preceding trial, then the amount of the immediate reward was decreased by half (i.e., 10 and five squirts for the 40- and 20-squirt conditions, respectively); if the participant had chosen the larger, delayed reward on the preceding trial, then the amount of the immediate reward was increased by half (Jimura et al., 2009, 2011). The adjustment amount was five squirts in the third trial in the 40-squirt condition. The subjective value of the delayed reward was estimated to be 1 ml (i.e., 2.5 squirts) more or less than the amount of the immediate reward available in the last trial (third and second trials in the 40- and 20-squirt conditions, respectively), depending on whether the delayed or immediate reward had been chosen on that trial.

After the behavioral session, the participants practiced drinking liquid rewards in a supine position with a mock scanner setup. When drinking liquid rewards, they were encouraged to use jaw movements and mouth muscles for swallowing, but not to move their heads.

fMRI Session Procedure

During fMRI scanning, participants performed an intertemporal decision-making task that was similar to that of the behavioral session. The primary difference was that the choice options for each trial were prespecified (rather than adjusted across the session), but set in an individualized manner based on a discounting profile estimated from the behavioral session. Three conditions (60 s/40 squirts, 30 s/40 squirts, 30 s/20 squirts) were used to measure brain activity during the delay period. The value of the immediate reward was systematically manipulated so that across trials, its value was smaller than the subjective value of the delayed reward, estimated for each participant based on their choice profile in the behavioral session. This manipulation of the immediate reward amount biased decisions toward delayed options, as the reward value was always smaller than the subjective value of the delayed reward, providing more opportunity to measure brain activity during consumption of delayed rewards (Jimura et al., 2013). When drinking the liquid rewards, the participants were instructed to use jaw movements and mouth muscles without moving their heads.

Imaging Procedure

Both anatomical and functional images were available from each participant. High-resolution anatomical images were acquired using an MP-RAGE T1-weighted sequence [repetition time (TR), 9.7 s; echo time (TE), 4.0 ms; flip angle (FA), 10°; slice thickness, 1 mm; in-plane resolution, 1×1 mm^2]. Functional [blood oxygen level-dependent (BOLD)] images were acquired using a gradient echo-planar imaging sequence (TR, 2.0 s; TE, 27 ms; FA, 90°; slice thickness, 4 mm; in-plane resolution, 4×4 mm^2; 34 slices) in parallel to the anterior–posterior commissure line, allowing complete brain coverage at a high signal-to-noise ratio. Each functional run involved 512 volume acquisitions.

Assessment of Impulsivity

For each participant, the degree of behavioral impulsivity was quantified by calculating the area under the discounting curve (AuC; Myerson et al., 2001; Sellitto et al., 2010; Jimura et al., 2011, 2013; Tanaka et al., 2020). The AuC represents the area under the observed subjective values at a given delay; more specifically, the AuC was calculated as the sum of the trapezoidal areas under the indifference points normalized by the amount and delay (Myerson et al., 2001). Both subjective value and delay were normalized for the purposes of calculating the AuC, which, as a result, ranged between 0.0 (maximally steep discounting) and 1.0 (no discounting). It has been argued that the AuC is the best measure of delay discounting for use in individual difference analyses, because it is theoretically neutral (i.e., assumption-free) and also psychometrically reliable (Myerson et al., 2001).

Each participant was classified into one of three groups, namely steep (STP), shallow (SHL), and intermediate (INT) discounting, based on their AuC values. The groups were identical to those analyzed in the previous study (Jimura et al., 2013).

Image Analysis Procedure

Image Preprocessing

Imaging data were analyzed using SPM12[1]. All functional images were first temporally aligned across the brain volume, corrected for movement using rigid-body rotation and translation correction, and then registered to the participant's anatomical images to correct for movement between the anatomical and function scans. Participants' anatomical images were transformed into standardized MNI atlas space. The functional images were then registered to the reference brain using the alignment parameters derived for the anatomical scans. The data were next resampled into 2-mm isotropic voxels and spatially smoothed with an 8-mm full-width at half-maximum Gaussian kernel.

General Linear Model

A general linear model (GLM) approach was used to separately estimate parameter values for each event occurring during the task. Consumption of liquid rewards after the delay period was

[1] https://www.fil.ion.ucl.ac.uk/spm/

encoded as an epoch that started from the press of the button to begin drinking (i.e., initiation of pump movement) until the time at which all the liquid rewards were infused into the participants' mouths (i.e., cessation of pump movement). As we focused on the consumption of delayed rewards, consumption periods for immediate rewards after participants chose immediate options were coded separately but similarly, and not analyzed in the current study. Choice and delay periods and distractor tasks were also included in the GLM as in the previous study (Jimura et al., 2013). All events were convolved with a canonical hemodynamic response function (HRF). In order to reduce potential confounds of head movements derived from jaw movements during drinking, head motion estimation parameters were also included in the GLM as nuisance regressors.

The parameter estimates of the consumption of delayed rewards were collected from all participants and then submitted to a group-level GLM analysis treating the participants as a random effect. To examine the correlation between AuC and the parameter estimates across participants, the AuC values of individual participants were z-scored (i.e., demeaned and divided by the standard deviation), and then included in the GLM. Additionally, for each head-movement axis, the maximum value of the movement parameters was calculated along the temporal dimension, and then z-scored across participants. The maximum movement parameters for six axes were included in the GLM as nuisance regressors to minimize potential confounds derived from head motions. Thus, the group-level GLM involved eight regressors (constant, AuC, and movement values for six axes). Z-scoring AuC and movement parameters orthogonalized these parameters and the constant regressor (group-mean effect).

During consumption, because participants were able to press and release the button to regulate liquid flow, imaging data could be confounded by the repetitive button presses. However, as shown in **Figure 2**, head motion during button press is almost absent. We thus believe that button-press-derived head motions are not major confounds. Another possible confound is BOLD signal reflecting the motor execution. Importantly, as noted above, participants received a practice session after the behavioral session to drink liquid rewards in a supine position using a mock scanner. No participants had difficulty drinking the rewards. The practices enabled the participants to drink the reward without pausing liquid flows, and thus, repetitive button presses were almost absent during the drinking period. This entails that the regressor coding the button presses become almost linear to the drink-period regressor. Thus, simultaneous event coding of button press and drinking would produce significant multicollinearity. Then, we only coded drinking events in our GLM analysis to avoid statistical artifacts. We also acknowledge that the activation maps in **Figure 3** involved finger movements.

Definitions of Regions of Interest

Because the current study focused on the mechanisms in the VS and anterior prefrontal cortex (aPFC; see also "Introduction" and "Results" sections), a region-of-interest (ROI) approach was used. ROIs were defined independently of the activation data during liquid consumption that were analyzed in the current study.

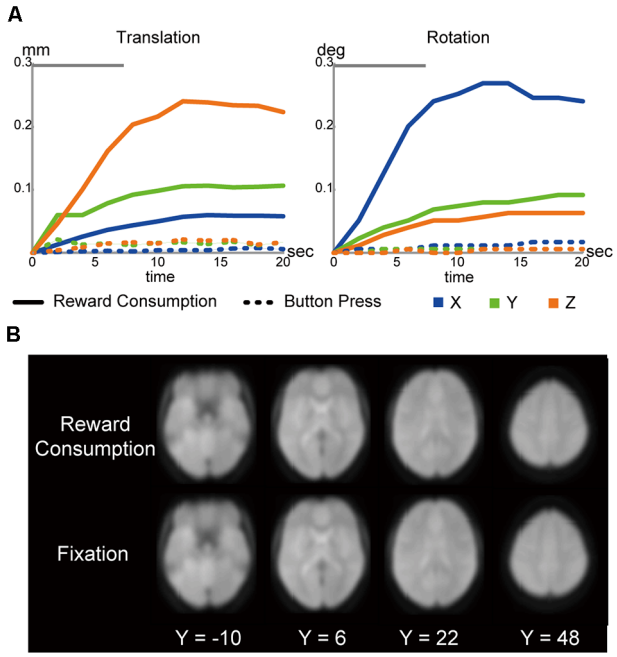

FIGURE 2 | (A) Head motions during liquid consumptions (solid line) and button presses (dotted line; *left*: translation; *right*: rotation). The gray bars on the top indicate the mean drinking duration. **(B)** Functional images during liquid consumptions and fixation periods. The images were normalized into the MNI space, averaged across participants, and then shown in transverse sections. The levels of sections are indicated by Z levels at the bottom.

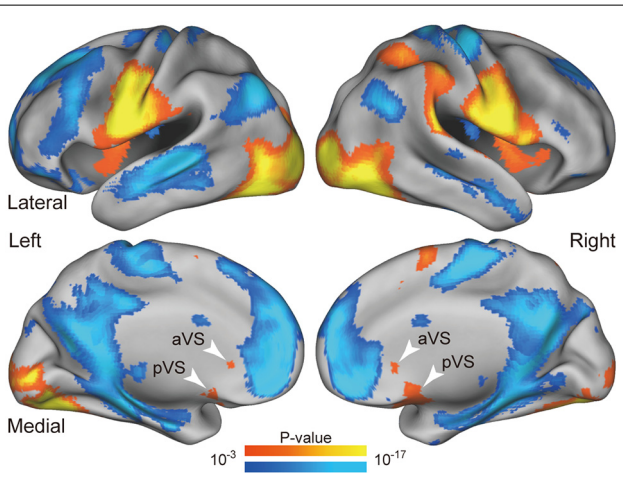

FIGURE 3 | Whole-brain statistical maps for activations during liquid consumption. Statistical maps of the activations are overlaid on a 3D surface of the standard brain. The hot and cool colors indicate positive and negative activity relative to the fixation periods, respectively. The white arrowheads indicate positive activity in the anterior and posterior VS regions. aVS: anterior ventral striatum; pVS: posterior ventral striatum.

The VS ROIs were defined anatomically as spheres with 8-mm radius, centered at the bilateral anterior and posterior ends of the VS in the Harvard-Oxford MNI atlas; the spheres were further masked by the anatomical VS regions. We defined

the VS ROIs for anterior and posterior parts separately, given strong activation with distinct peaks during consumption of liquid rewards, as shown in **Figure 3**. The anterior and posterior ROIs were labeled as the anterior VS (aVS) and posterior VS (pVS), respectively.

aPFC ROIs were defined as spheres with an 8-mm radius that were centered at bilateral aPFC coordinates showing an anticipatory utility effect during the delay period, as reported in our previous study analyzing the identical data set [coordinates: (28, 54, −7), (−31, 55, −7); Table 1 in Jimura et al., 2013]. These bilateral aPFC regions also showed the anticipatory utility effect in our recent study (Tanaka et al., 2020). For exploratory analysis, a small-volume correction approach was used, and statistical significance levels of the peak were corrected for multiple comparisons within the aPFC ROIs using voxel-level family-wise error rates.

Psychophysiological Interaction (PPI) Analysis

A set of PPI analyses (Friston et al., 1997) was performed to examine task-related functional connectivity among aPFC and VS regions. The seed regions for the aPFC, aVS, and pVS in each hemisphere (i.e., six ROIs in total) were identical to the ROIs defined above.

For each of the six ROIs, PPI effects were first calculated as implemented in SPM12. Then, single-level statistical analysis was performed based on a standard GLM analysis for each ROI. The GLM models included the PPI and nuisance effects (i.e., the timecourse of MRI signals in the ROI, the main effect of the condition of interest convolved with a canonical HRF, head-movement parameters for the axes, and timecourse of MRI signals for white matter, cerebrospinal fluid, and the whole-brain).

For each seed ROI, the estimated PPIs were extracted for all target ROIs (i.e., five ROIs). Thus, 30 PPIs (six seeds × five targets) were calculated for each participant. Then, these PPIs were collected from all participants, and group-level statistics were calculated for PPIs from each pair of seed and target ROIs. For statistical testing, PPIs between seed and target regions were averaged across contralateral and ipsilateral hemispheres, as we did not observe strong hemispheric asymmetry in PPIs (see "Results" section). Then, the significance of the PPI strength was tested by the one-sample *t*-test. P-values were corrected for multiple comparisons based on Bonferroni correction.

RESULTS

Behavioral Results

Participants performed 19.5 ± 2.6 (mean ± SD) trials, and choose the delayed alternative in 83.0 ± 12.6% of the trials. The number of trials where the delayed alternative was chosen did not correlate with behavioral impulsivity (AuC) measured in the behavioral session (see "Materials and Methods" section) [$r = 0.15$, $t_{(41)} = 0.97$, $P = 0.33$]. The mean drinking duration was 7.5 ± 2.1 s.

Participants were classified into three groups based on AuC values reflecting the delay-discounting pattern (see "Materials and Methods"section), as in the previous study (Jimura et al., 2013: steep discounters (N = 15; high impulsivity), shallow discounters (N = 15; low impulsivity) and an intermediate group (N = 13).

Head Movements During Drinking

The liquid was delivered from outside the scanner room through a plastic tube, which enabled participants to consume the liquid during fMRI administration (see "Materials and Methods" section). However, it is well known that head movements during fMRI lead to significant artifacts and signal losses in images. To evaluate whether our data were contaminated by the motion-derived artifacts and signal losses, we first assessed head movements and MRI images while participants were drinking liquid rewards.

As shown in **Figure 2A**, head movements were greater during liquid consumptions than during button presses made in a money discounting distractor task performed in the same scanning sessions (**Figure 1B**; Jimura et al., 2013, 2018). However, the absolute magnitude of head motion during consumption was small, with maximum translations of (0.06, 0.11, 0.24 mm) ± (0.02, 0.03, 0.09 mm; mean ± SD), and maximum rotations of (0.27, 0.09, 0.06) ± (0.09, 0.03, 0.02) degrees, along the *x*-, *y*-, and *z*-axes, respectively. Moreover, the maximum head movement parameters in the six axes (see "Materials and Methods" section) did not correlate with the AuC ($|r|s < 0.28$; $|t|s < 1.88$; Ps > 0.07, uncorrected).

Nonetheless, it is known that jaw movements can yield significant instability in echoplanar images, as reported in nonhuman primate scanning (Keliris et al., 2007). However, the instability seemed absent in the current study, as we observed comparable image quality during liquid consumption and the fixation periods (**Figure 2B**).

Given these results, we felt confident in assuming that movement-derived contamination during liquid consumption was less obvious in the current study than the previous study (Keliris et al., 2007). Our recent study also showed similar results (Tanaka et al., 2020).

Imaging Results

Brain Activity During Liquid Consumption

Figure 3 shows brain activity during consumption of delayed liquid rewards. Robust activations were observed in the primary gustatory cortex, primary motor cortex related to the jaw, and primary visual cortex, as the maps reflect various effects including jaw movements, swallowing, gustatory perception, and visual perception. These prominent activations validated the absence of major contaminations due to movement-derived artifacts and indicated that the data obtained during the consumption period could be used in substantive analyses. Importantly, robust activations were observed in the anterior and posterior parts of the bilateral VS (white arrows in **Figure 3**).

Ventral Striatal Activity and Impulsivity

As we observed stronger activity in the anterior and posterior VS (aVS and pVS, respectively) during consumption (**Figure 3**), we examined the activity in the aVS and pVS in each discounting group. We anatomically defined ROIs in the aVS and pVS, and

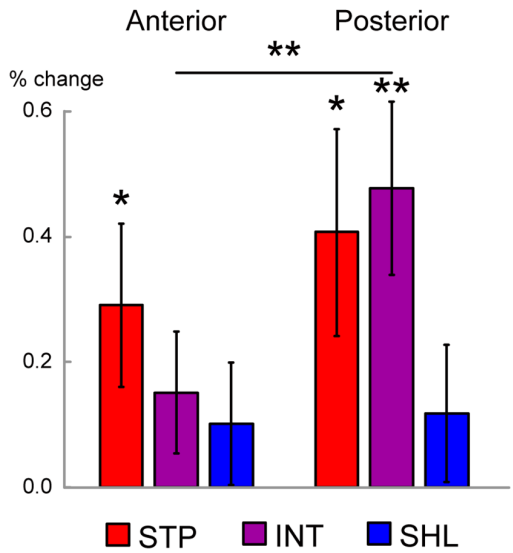

FIGURE 4 | Regions of interest (ROI) analysis in the ventral striatum (VS). The magnitude of MRI signals is shown during the liquid consumption period in the anterior and posterior VS ROIs were defined anatomically. Signal magnitudes were calculated for each of the discounting groups (STP: steep; INT: intermediate; SHL: shallow), and averaged across hemispheres. Error bars indicate SEM. Levels of statistical significance are indicated by the number of asterisks (**$P < 0.01$; *$P < 0.05$).

evaluated consumption activity in these ROIs (see "Materials and Methods" section).

As shown in **Figure 4**, steep discounters (high impulsivity) showed significant activation in the aVS, [$t_{(14)} = 2.2$; $P < 0.05$]. On the other hand, both steep and intermediate discounters showed significant activation in the pVS [steep: $t_{(14)} = 2.5$; $P < 0.05$; intermediate: $t_{(12)} = 3.5$; $P < 0.01$]. Interestingly, in the intermediate discounters, the activation was greater in the pVS than the aVS [$t_{(12)} = 3.1$; $P < 0.01$]. On the other hand, in both the aVS and pVS, significant activation was absent in shallow discounters (low impulsivity). These results suggest that the aVS and pVS are differently involved in liquid reward consumption depending on the degree of impulsivity.

Prefrontal Activity and Impulsivity

Aiming to examine neural mechanisms consistently involved through task events, we asked how the anterior prefrontal region related to reward anticipation during the delay (Jimura et al., 2013; Tanaka et al., 2020) was involved during consumption.

We first explored brain regions showing significant activation during consumption within aPFC ROIs. The ROIs were defined based on our previous study (Jimura et al., 2013; see also "Materials and Methods" section). As shown in **Figure 5** (*left*), aPFC regions showed strong activation bilaterally [$P < 0.05$, corrected for multiple comparison based on voxel-level family-wise error rate; left: (−32, 62, −6), $z = 2.6$; right: (26, 58, −1), $z = 3.0$].

We next examined the correlation between the consumption period activity and the degree of delay discounting estimated in a separate behavioral session (see "Materials and Methods" section). Specifically, we explored aPFC ROIs showing the correlation between AuC and brain activity during consumptions. As shown in **Figure 5** (*middle*), negative correlations were observed in the bilateral aPFC [$P < 0.05$, corrected for multiple comparison based on voxel-level family-wise error rate; left: (−28, 52, −8), $z = 2.5$; right: (32, 48), −6, $z = 2.4$]. Importantly, these aPFC regions were identified within the identical ROIs involving regions showing strong activation. The negative correlations indicate greater activity in steep discounters (high impulsivity; **Figure 5** *right*). In impulsive individuals, the aPFC activity was almost significant [$t_{(14)} = 2.0$, $P = 0.065$], possibly due to the small sample size of each discounting group, and strong activation was absent when averaging across all participants [$t_{(42)} = 0.52$, $P = 0.62$].

Interestingly, in the previous studies, the aPFC regions exhibited an anticipatory utility effect during the delay period and the effect was enhanced in shallow discounters (Jimura et al., 2013; Tanaka et al., 2020), whereas in the current study, the consumption period activity was reduced in shallow discounters (**Figure 5** *left*; see "Discussion" section for details).

Psychophysiological Interaction

The results related to consumption period activity and its relation to behavioral impulsivity suggest that the aPFC and VS play an important role during the consumption of real liquid rewards. We, therefore, examined task-related functional connectivity between these regions based on psychophysiological interactions (PPIs; Friston et al., 1997; see "Materials and Methods" section).

Figure 6A shows PPIs between the aPFC, aVS, and pVS. For each pair of ROIs, PPIs appear to covary between ipsilateral and contralateral hemispheres, and obvious hemispheric asymmetries look absent. Thus, PPIs were averaged across hemispheres, and statistical testing was performed. From the aPFC, PPI was strong towards aVS [$t_{(42)} = 3.1$; $P < 0.05$, Bonferroni corrected], but strong PPI was not observed in the opposite direction, suggesting top-down signaling from the aPFC to aVS. On the other hand, PPIs were robust bidirectionally between the aVS and pVS [aVS to pVS: $t_{(42)} = 3.1$; $P < 0.05$, Bonferroni corrected; pVS to aVS: $t_{(42)} = 4.6$; $P < 0.0001$, Bonferroni corrected].

In order to examine whether the PPIs were dependent on impulsivity, the bidirectional PPI matrix was inspected for the steep, intermediate, shallow discounting groups. As shown in **Figure 6B**, in steep discounters, there was strong PPI from the pVS to aVS [$t_{(42)} = 3.7$; $P < 0.001$, Bonferroni corrected], but not in the opposite direction. Such unidirectional strong PPI was absent in the other two groups.

DISCUSSION

The current study provides new insights regarding prefrontal-striatal mechanisms of intertemporal choice by focusing on brain activity and task-related functional connectivity while humans consumed real liquid rewards delayed by tens of seconds. Impulsivity was associated with activity enhancement in the aPFC and VS, and the activation magnitudes in the VS

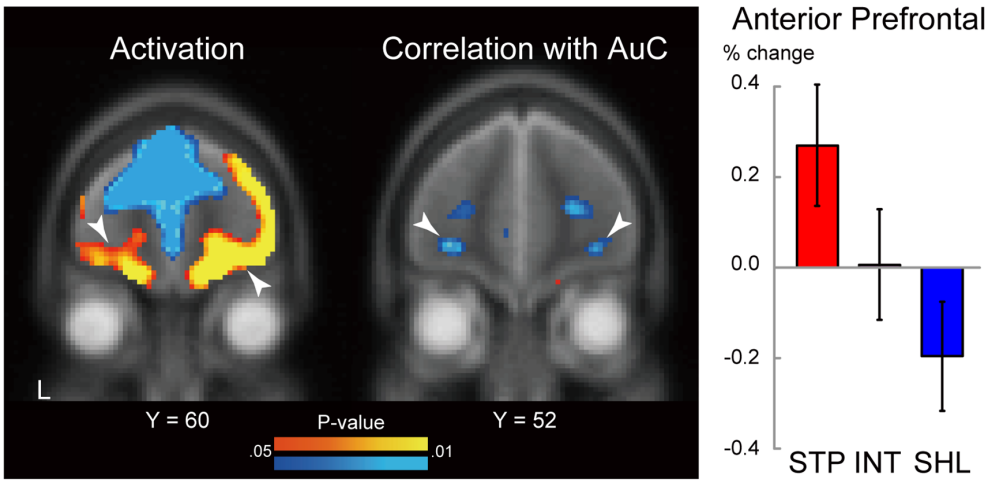

FIGURE 5 | Statistical map for brain activity during liquid consumption (*left*). The level of the section is indicated by the Y coordinate of the MNI space. The threshold of the map was $P < 0.05$ (uncorrected) for display purposes. White arrow heads indicate correlations in the anterior prefrontal cortex. L: left. Hot and cool colors indicate positive and negative activation, respectively. Statistical map for correlation between behavioral impulsivity and brain activity during liquid consumption (*middle*). Behavioral impulsivity is quantified as the area under the curve of the subjective value of the delayed reward. The format is similar to those on the left. Hot and cool colors indicate positive and negative correlation, respectively. MRI signal magnitudes in the aPFC ROIs were calculated for each discounting group and averaged across hemispheres (*right*). Regions of interest were defined based on the previous study (Jimura et al., 2013). The formatting is similar to that in **Figure 4**.

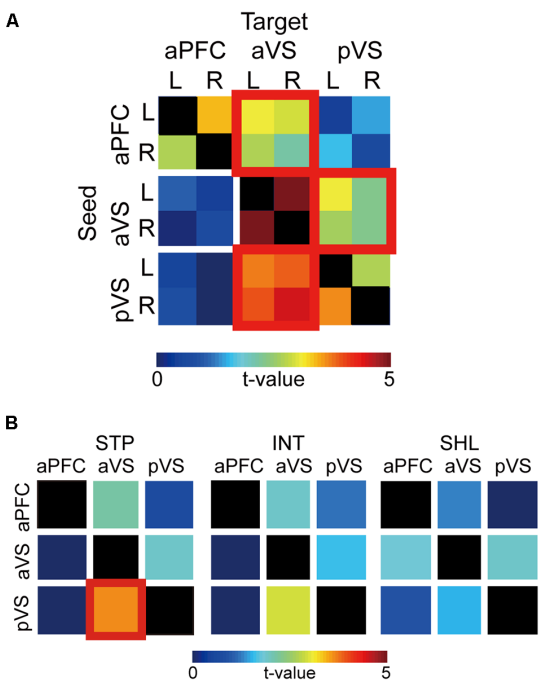

FIGURE 6 | Psychophysiological interaction (PPI) analysis. **(A)** Group level *t*-values of PPIs from seed to target regions are color-coded according to the bar at the bottom. The columns and rows indicate seeds and targets, respectively. aPFC: anterior prefrontal cortex; aVS: anterior ventral striatum; pVS: posterior ventral striatum. Significant PPIs are indicated by red squares. L: left; R: right. **(B)** PPIs for each discounting group. Formatting is similar to that in panel **(A)**.

differed between the anterior and posterior regions depending on impulsivity. PPI was robust from the aPFC to aVS, but not in the opposite direction, suggesting top-down signaling from aPFC to aVS. On the other hand, bidirectional PPIs were observed between the aVS and pVS, with enhanced PPI from the pVS to aVS in impulsive individuals (steep discounters). These findings suggest that prefrontal and striatal mechanisms are involved in reward consumption, reflecting behavioral impulsivity in decision-making.

A Putative Prefrontal-Striatal Model of Impulsivity During Reward Consumption

The current results highlight functional segregation and integration in the prefrontal cortex and the ventral striatum during the consumption of delayed liquid rewards. **Figure 7** summarizes our results and illustrates activity magnitudes and signal flows between the aPFC, aVS, and pVS for three levels of impulsivity.

In the aPFC, activation is greater in impulsive individuals, resulting in enhanced signaling toward the aVS and an elevation of its activity. On the other hand, the pVS transmits reward-related signals to the aVS, and then aVS activity is further amplified in impulsive individuals. Interestingly, the signals from both the aPFC and pVS appear to be associated with both the magnitude of aVS activation and the level of impulsivity.

These mechanisms can be interpreted as showing that the reward-related signals from the aPFC and pVS are aggregated into the aVS, which is critical for the degree of behavioral impulsivity. These prefrontal-striatum mechanisms are compatible with those suggested by prior studies analyzing connectivity between aPFC and VS (Diekhof and Gruber, 2010; Jimura et al., 2013; Tanaka et al., 2020).

A diffusion tensor imaging study of the human VS showed that both the anterior and posterior parts of the VS are connected to the anterior ventral part of the PFC and the orbitofrontal

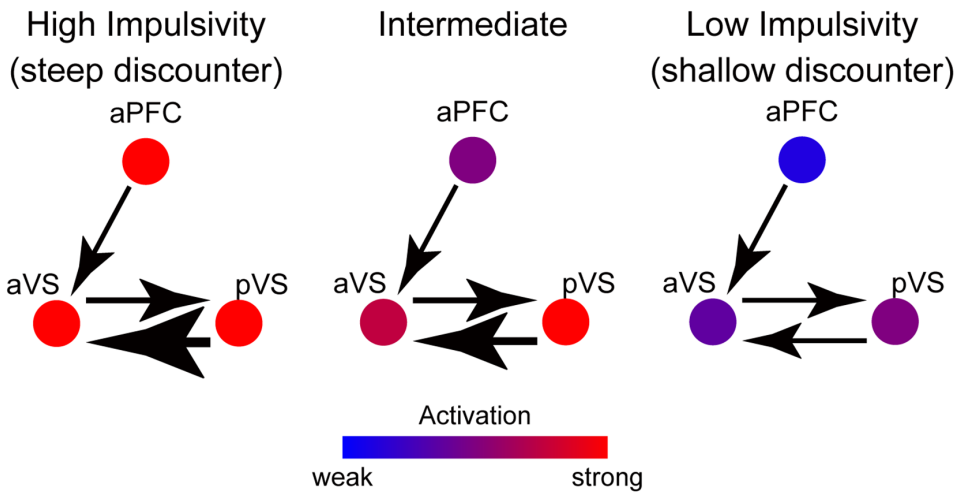

FIGURE 7 | A schematic path diagram for putative functional mechanisms between the aPFC, aVS, and pVS during reward consumption. The colors of the circles indicate the activation magnitude in these regions, according to the color bar at the bottom. The thickness of each arrow indicates the strength of the connectivity from region to region, and the direction of each arrowhead indicates the signal flow direction. In highly impulsive individuals, strong aPFC and pVS signals are sent to the aVS, enhancing aVS activity, whereas, in minimally impulsive individuals, negative aPFC activation reduces aVS activity.

cortex, including the aPFC in the current study. Interestingly, the anterior part is also connected to more dorsal parts of the PFC (Tziortzi et al., 2014). The dorsal parts of the PFC are involved in working memory and choice difficulty in intertemporal choice, which is enhanced in self-controlled individuals (Jimura et al., 2018). The anatomical connection between the anterior VS and dorsal PFC may regulate the aPFC-aVS-pVS mechanisms during consumption of liquid rewards, as illustrated by the reduced activity in the intermediate discounting group (**Figure 4**).

aPFC-VS Mechanisms in Distinct Behavioral Phases

The current study focused on the consumption phase of intertemporal choice, whereas previous studies analyzed delay and choice phases (Jimura et al., 2013; Tanaka et al., 2020). These collective results are derived from a single dataset, allowing us to speculate regarding possible functional mechanisms involved in the distinct phases of intertemporal choice behavior: choice, delay, and consumption.

In all three phases, impulsivity in decision-making was associated with increased VS activation, consistent with previous reports (Tanaka et al., 2004; Hariri et al., 2006; Kable and Glimcher, 2007; Ballard and Knutson, 2009; Pine et al., 2009). On the other hand, aPFC activation differed among the phases. During the choice and early delay periods, self-controlled (less impulsive) individuals exhibited increased aPFC activity (Jimura et al., 2013; Tanaka et al., 2020). In contrast, in the current study, highly impulsive individuals exhibited increased aPFC activity during the consumption period. A possible unified explanation of these findings is that the aPFC is associated with currently available utility, i.e., future anticipation during delay and reward consumption during drinking. The former may be more valuated in self-controlled individuals to maximize future reward attainment, whereas the latter may be evaluated in impulsive individuals when consuming a reward after a delay.

Greater aPFC activation during a choice period in less impulsive individuals may reflect resistance to impulsive choice, which reduces VS activity (Diekhof and Gruber, 2010). During the delay period, aPFC regions may encode an anticipatory utility signal associated with delayed rewards, which is the extra utility derived from the pleasure of waiting for a reward delivered in the future (Loewenstein, 1987; Berns et al., 2006, 2007; Peters and Buechel, 2010; Jimura et al., 2013; Tanaka et al., 2020).

On the other hand, as mentioned above, highly impulsive individuals showed greater aPFC activation during the consumption period, suggesting marked current utility when individuals consume liquid rewards. Because highly impulsive individuals do not prefer to wait to obtain a larger reward, the utility of a liquid reward would become greater upon completion of the delay, eliciting greater aPFC activation when consuming delayed rewards.

Another possibility is that the aPFC encodes current pleasure related to reward attainment, as suggested by aPFC regions showing anticipatory utility effect that is thought to reflect pleasure of waiting (Loewenstein, 1987; Jimura et al., 2013; Tanaka et al., 2020). Thus, the receipt of a delayed reward may provide greater pleasure for highly impulsive individuals. Alternatively, while consuming the reward, highly impulsive individuals may retrieve episodic information about past experiences of rewards, resulting in greater aPFC activation. However, additional evidence is needed to directly support the role of the aPFC in relation to pleasure.

CONCLUSION

The current study addressed a unique question, how the prefrontal cortex and ventral striatum are involved

while humans consumed delayed real liquid rewards. We found that the anterior prefrontal cortex, anterior ventral striatum, and posterior ventral striatum constitute a functional network, which is modulated by behavioral impulsivity. Our results highlight a prefrontal-striatal mechanism of behavioral impulsivity during reward consumption.

DATA AVAILABILITY STATEMENT

The datasets are available from the corresponding author on reasonable request. Requests to access these datasets should be directed to Koji Jimura, koji.jimura@gmail.com.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Washington University in St. Louis, USA. Patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

KJ designed research. AM and KJ performed research, analyzed data, wrote the article. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by Kakenhi (JSPS; 26350986, 26120711, 17K01989, 17H05957, and 19H04914), grants from Uehara Memorial Foundation, and Takeda Science Foundation to KJ.

ACKNOWLEDGMENTS

We thank Dr. Todd S. Braver for data sharing and scientific advice.

REFERENCES

Ainslie, G. (1975). Specious reward - behavioral theory of impulsiveness and impulse control. *Psychol. Bull.* 82, 463–496. doi: 10.1037/h0076860

Ballard, K., and Knutson, B. (2009). Dissociable neural representations of future reward magnitude and delay during temporal discounting. *Neuroimage* 45, 143–150. doi: 10.1016/j.neuroimage.2008.11.004

Berns, G. S., Chappelow, J., Cekic, M., Zink, C. F., Pagnoni, G., and Martin-Skurski, M. E. (2006). Neurobiological substrates of dread. *Science* 312, 754–758. doi: 10.1126/science.1123721

Berns, G. S., Laibson, D., and Loewenstein, G. (2007). Intertemporal choice - toward an integrative framework. *Trends. Cogn. Sci.* 11, 482–488. doi: 10.1016/j.tics.2007.08.011

Diekhof, E. K., and Gruber, O. (2010). When desire collides with reason: functional interactions between anteroventral prefrontal cortex and nucleus accumbens underlie the human ability to resist impulsive desires. *J. Neurosci.* 30, 1488–1493. doi: 10.1523/JNEUROSCI.4690-09.2010

Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., and Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage* 6, 218–229. doi: 10.1006/nimg.1997.0291

Hare, T. A., Camerer, C. F., and Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. *Science* 324, 646–648. doi: 10.1126/science.1168450

Hariri, A. R., Brown, S. M., Williamson, D. E., Flory, J. D., de Wit, H., and Manuck, S. B. (2006). Preference for immediate over delayed rewards is associated with magnitude of ventral striatal activity. *J. Neurosci.* 26, 13213–13217. doi: 10.1523/JNEUROSCI.3446-06.2006

Jimura, K., Chushak, M. S., and Braver, T. S. (2013). Impulsivity and self-control during intertemporal decision making linked to the neural dynamics of reward value representation. *J. Neurosci.* 33, 344–357. doi: 10.1523/JNEUROSCI.0919-12.2013

Jimura, K., Chushak, M. S., Westbrook, A., and Braver, T. S. (2018). Intertemporal decision-making involves prefrontal control mechanisms associated with working memory. *Cereb Cortex* 28, 1105–1116. doi: 10.1093/cercor/bhx015

Jimura, K., Myerson, J., Hilgard, J., Braver, T. S., and Green, L. (2009). Are people really more patient than other animals? Evidence from human discounting of real liquid rewards. *Psychon. Bull. Rev.* 16, 1071–1075. doi: 10.3758/PBR.16.6.1071

Jimura, K., Myerson, J., Hilgard, J., Keighley, J., Braver, T. S., and Green, L. (2011). Domain independence and stability in young and older adults' discounting of delayed rewards. *Behav. Proc.* 87, 253–259. doi: 10.1016/j.beproc.2011.04.006

Kable, J. W., and Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10, 1625–1633. doi: 10.1038/nn2007

Keeney, R. L., and Raiffa, H. (1993). *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. New York: Wiley.

Keliris, G. A., Shmuel, A., Ku, S.-P., Pfeuffer, J., Oeltermann, A., Steudel, T., et al. (2007). Robust controlled functional MRI in alert monkeys at high magnetic field: effects of jaw and body movements. *Neuroimage* 36, 550–570. doi: 10.1016/j.neuroimage.2007.02.057

Levy, D. J., and Glimcher, P. W. (2011). Comparing apples and oranges: using reward-specific and reward-general subjective value representation in the brain. *J. Neurosci.* 31, 14693–14707. doi: 10.1523/JNEUROSCI.2218-11.2011

Loewenstein, G. (1987). Anticipation and the valuation of delayed consumption. *Econ. J.* 97, 666–684. doi: 10.2307/2232929

Madden, G. J., and Bickel, W. K. (2009). *Impulsivity: The Behavioral and Neurological Science of Discounting*. Washington, DC: American Psychological Association.

McClure, S. M., Ericson, K. M., Laibson, D. I., Loewenstein, G., and Cohen, J. D. (2007). Time discounting for primary rewards. *J. Neurosci.* 27, 5796–5804. doi: 10.1523/JNEUROSCI.4246-06.2007

McClure, S. M., Laibson, D. I., Loewenstein, G., and Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science* 306, 503–507. doi: 10.1126/science.1100907

McGuire, J. T., and Kable, J. W. (2015). Medial prefrontal cortical activity reflects dynamic re-evaluation during voluntary persistence. *Nat. Neurosci.* 18, 760–766. doi: 10.1038/nn.3994

Mischel, W., Shoda, Y., and Rodriguez, M. L. (1989). Delay of gratification in children. *Science* 244, 933–938. doi: 10.1126/science.2658056

Myerson, J., Green, L., and Warusawitharana, M. (2001). Area under the curve as a measure of discounting. *J. Exp. Anal. Behav.* 76, 235–243. doi: 10.1901/jeab.2001.76-235

Peters, J., and Buechel, C. (2010). Episodic future thinking reduces reward delay discounting through an enhancement of prefrontal-mediotemporal interactions. *Neuron* 66, 138–148. doi: 10.1016/j.neuron.2010.03.026

Pine, A., Seymour, B., Roiser, J. P., Bossaerts, P., Friston, K. J., Curran, H. V., et al. (2009). Encoding of marginal utility across time in the human brain. *J. Neurosci.* 29, 9575–9581. doi: 10.1523/JNEUROSCI.1126-09.2009

Rachlin, H. (2004). *The Science of Self-control*. Cambridge: Harvard University Press.

Rachlin, H., Raineri, A., and Cross, D. (1991). Subjective-probability and delay. *J. Exp. Anal. Behav.* 55, 233–244. doi: 10.1901/jeab.1991.55-233

Sellitto, M., Ciaramelli, E., and di Pellegrino, G. (2010). Myopic discounting of future rewards after medial orbitofrontal damage in humans. *J. Neurosci.* 30, 16429–16436. doi: 10.1523/JNEUROSCI.2516-10.2010

Shamosh, N. A., DeYoung, C. G., Green, A. E., Reis, D. L., Johnson, M. R., Conway, A. R. A., et al. (2008). Individual differences in delay discounting relation to intelligence, working memory and anterior prefrontal cortex. *Psychol. Sci.* 19, 904–911. doi: 10.1111/j.1467-9280.2008.02175.x

Tanaka, D., Aoki, R., Suzuki, S., Takeda, M., Nakahara, K., and Jimura, K. (2020). Self-controlled choice arises from dynamics prefrontal signals that enable future anticipation. *J. Neurosci.* 40, 9736–9750. doi: 10.1523/JNEUROSCI.1702-20.2020

Tanaka, S. C., Doya, K., Okada, G., Ueda, K., Okamoto, Y., and Yamawaki, S. (2004). Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat. Neurosci.* 7, 887–893. doi: 10.1038/nn1279

Tziortzi, A. C., Haber, S. N., Searle, G. E., Tsoumpas, C., Long, C. J., Shotbolt, P., et al. (2014). Connectivity-based functional analysis of dopamine release in the striatum using diffusion-weighted MRI and positron emission tomography. *Cereb. Cortex* 24, 1165–1177. doi: 10.1093/cercor/bhs397

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 in Human Neuroscience

ORIGINAL RESEARCH
published: 26 November 2021
doi: 10.3389/fnhum.2021.777464

Distillation of Regional Activity Reveals Hidden Content of Neural Information in Visual Processing

Trung Quang Pham[1], *Shota Nishiyama*[1,2,3], *Norihiro Sadato*[1,4] *and Junichi Chikazoe*[1,3]*

[1] Section of Brain Function Information, Supportive Center for Brain Research, National Institute for Physiological Sciences, Okazaki, Japan, [2] Aichi Institute of Technology Graduate School of Business Administration and Computer Science, Toyota, Japan, [3] Araya Inc., Tokyo, Japan, [4] Division of Cerebral Integration, National Institute for Physiological Sciences, Okazaki, Japan

Multivoxel pattern analysis (MVPA) has become a standard tool for decoding mental states from brain activity patterns. Recent studies have demonstrated that MVPA can be applied to decode activity patterns of a certain region from those of the other regions. By applying a similar region-to-region decoding technique, we examined whether the information represented in the visual areas can be explained by those represented in the other visual areas. We first predicted the brain activity patterns of an area on the visual pathway from the others, then subtracted the predicted patterns from their originals. Subsequently, the visual features were derived from these residuals. During the visual perception task, the elimination of the top-down signals enhanced the simple visual features represented in the early visual cortices. By contrast, the elimination of the bottom-up signals enhanced the complex visual features represented in the higher visual cortices. The directions of such modulation effects varied across visual perception/imagery tasks, indicating that the information flow across the visual cortices is dynamically altered, reflecting the contents of visual processing. These results demonstrated that the distillation approach is a useful tool to estimate the hidden content of information conveyed across brain regions.

Keywords: MVPA, decoding, machine learning, fMRI, visions

OPEN ACCESS

Edited by:
Kuniyoshi L. Sakai,
The University of Tokyo, Japan

Reviewed by:
Michael W. Cole,
Rutgers University, Newark, United States
Chunlin Li,
Capital Medical University, China

***Correspondence:**
Junichi Chikazoe
j.chikazoe@gmail.com

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

Received: *15 September 2021*
Accepted: *09 November 2021*
Published: *26 November 2021*

Citation:
Pham TQ, Nishiyama S, Sadato N and Chikazoe J (2021) Distillation of Regional Activity Reveals Hidden Content of Neural Information in Visual Processing. Front. Hum. Neurosci. 15:777464. doi: 10.3389/fnhum.2021.777464

1. INTRODUCTION

Brain decoding has drawn interest from neuroscientists for decades. Decoding gives meaning to the activity patterns inside the brain, thus providing a potential for reverse engineering in order to understand how the brain organizes and stores information. Recent studies have broadly utilized the multi-voxel pattern analysis (MVPA) of functional magnetic resonance imaging (fMRI) images as a standard tool to decipher what people are seeing (Haxby et al., 2001; Kamitani and Tong, 2005; Horikawa and Kamitani, 2017), hearing (Hoefle et al., 2018), imagining (Stokes et al., 2009; Reddy et al., 2010; Cichy et al., 2012), and dreaming (Horikawa et al., 2013).

In terms of targeted perception, vision has been the preferred candidate due to its simplicity. Visual processing, particularly visual object recognition, is a well-established hierarchical organization in both anatomical and functional aspects (Felleman and Van Essen, 1991). A recent study (Horikawa and Kamitani, 2017) presented a decoding approach for generic decoding of visual features in both perception and imagery tasks. The authors suggested that the mental imagery is a

type of top-down processing, whereas mental perception is a bottom-up process. Interplay between top-down and bottom-up processing helps sharpen the neural representation of stimuli (Abdelhack and Kamitani, 2018). However, the top-down signals also cause bias in the early visual-sensitive area (Kok et al., 2012, 2013). Therefore, to unveil the "true pattern" reflecting the received visual stimuli, one should eliminate the influence of top-down signals.

In this study, MVPA of fMRI images was used to distill the unsullied pattern of activity in a region of interest (ROI). We assume the prediction of a low-level ROI based on the activity of a high-level ROI to specifically represent its top-down signals from that specific one, and the prediction of a high-level ROI based on the activity of a low-level signals to represent bottom-up signals. Using the open-access data obtained from (Horikawa and Kamitani, 2017), we demonstrated a region-to-region decoding technique in which the top-down/bottom-up signals at an ROI (target) are linearly integrated from the activity of the other regions (seeds). Thereafter, we examined the prediction of visual features of observed stimuli before and after eliminating the top-down/bottom-up signals during the perception and imagery tasks. Finally, we compared the magnitude of distillation effects between all possible seed–target pairs associated with the visual processing.

2. MATERIALS AND METHODS

2.1. Data and Preprocessing

We used the preprocessed task fMRI data of 5 subjects in the publicly accessible Generic Object Decoding dataset (https://github.com/KamitaniLab/GenericObjectDecoding). This dataset was used to replicate Horikawa et al.' paper (Horikawa and Kamitani, 2017). MRI data were collected using 3.0-Tesla Siemens MAGNETOM Trio A Tim scanner from the ATR Brain Activity Imaging Center. An interleaved T2∗-weighted gradient-echo plannar imaging (EPI) scan was performed [repetition time (TR), 3,000 ms; echo time (TE), 30 ms; flip angle, 80 deg; field of view [FOV], 192 × 192 mm^2]. T1-weighted magnetization-prepared rapid acquisition gradient-echo fine-structural images of the entire head were also acquired (TR, 2,250 ms; TE, 3.06 ms; TI, 900 ms; flip angle, 9 deg, FOV, 256 × 256 mm^2; voxel size, 1.0 × 1.0 × 1.0 mm^3.

The fMRI data underwent three-dimensional motion correction using the SPM5 software (http://www.fil.ion.ucl.ac.uk/spm). Data were then coregistered with the whole-head high-resolution anatomical images. The coregistered data were then reinterpolated using 3 × 3 × 3 mm^3 voxels. After within-run linear trend removal, voxel amplitudes were normalized relative to the mean activity of the entire time course within each run. To estimate the brain activity associated with each trial, the normalized voxel activity was then averaged within each 9-s stimulus block (image presentation experiment) or within each 15-s imagery period (imagery experiment), after shifting the delay the data by 3 s to compensate for hemodynamic delays.

This dataset consists of 1,200 training, 1,750 test (perception), and 500 test (imagery) trials for each subject. Visual images were collected from the online image database ImageNet (Deng et al., 2009). Two hundred representative object categories were selected as stimuli in the visual presentation experiment. In the training image session, a total of 1,200 images from 150 object categories (eight images from each category) were presented only once. In the test image session, a total of 50 images from 50 object categories (one image from each category) were presented 35 times each. Care was taken to avoid misuse of the categories for the test session during the training session. In the imagery experiment, the subjects were asked to visually imagine images from one of the 50 categories that were presented in the test image session of the image presentation experiment.

2.2. Region-to-Region Decoding

To estimate the information flow from a region to a region, we calculated the fine-grained topographic connectivity between regions (Heinzle et al., 2011). In this analysis, a single voxel activity in the target region was modeled by a weighted linear summation of all the voxel activities in the seed region. Considering that the activity of voxels in the same ROI is highly correlated, a ridge regression analysis was employed for weight estimation, but not ordinary least squares analysis. The ridge parameter was optimized such that the prediction performance in the validation dataset is maximized. For this purpose, we divided the training dataset into 600 training and 600 validation trials. In test dataset, the voxel activity in the target region was predicted through the estimated weights computed using the optimal ridge parameter.

To evaluate the region-to-region decoding performance, the average coefficients of determination (R^2) among all the voxels in the target ROI were calculated. For comparison, functional connectivity was also calculated between regions. As the present dataset reflects task-related activity, the method proposed by Rissman et al. (2004) was used.

2.3. Bottom-Up/Top-Down Signal Elimination

We hypothesize that the observed activity of visual cortices reflects both bottom-up and top-down signals conveyed between the visual pathways. The bottom-up/top-down signals can be approximated using linear predictions through the observation of other ROIs. Prediction of a targeted ROI derived from a seed ROI can be expressed as follows.

$$\mathbf{X}_{seed \rightarrow target} \approx \mathbf{a} \times \mathbf{X}^{*}_{seed} + \mathbf{b}$$

where **a** and **b** denote the parameters of the linear regression. The $\mathbf{X}^{*}_{seed}$ denotes the observed representation of a signal at the seed ROI. Then, the representation of the signal at the target ROI can be expressed as follows

$$\mathbf{X}^{*}_{target} = \mathbf{X}_{seed \rightarrow target} + \mathbf{X}_{latent\_factor}$$

where $\mathbf{X}^{*}_{target}$ denotes the observed representation of the signal at the target ROI, and the $\mathbf{X}_{latent\_factor}$ represents "hidden" content at the target ROI which was subsequently used for visual feature prediction.

These expressions suggest that the activity of the primary visual cortex would directly reflect retinal input if the top-down signal from the higher visual cortex (such as the fusiform face area, FFA) could be appropriately eliminated. Considering this, the activity explained by the seed region (e.g., FFA) was eliminated from the target region (e.g., V1). The former activity is estimated using the region-to-region decoding technique described above.

2.4. Visual Feature Prediction

We tested 13 candidates of visual features, including a convolutional neural network (CNN1–CNN8) (Krizhevsky et al., 2012), HMAX model (HMAX1–HMAX3) (Riesenhuber and Poggio, 1999; Serre et al., 2007; Mutch and Lowe, 2008), GIST (Oliva and Torralba, 2001), and scale-invariant feature transform (SIFT) (Lowe, 1999) in combination with the bag of features (BOF) (Csurka et al., 2004). All visual features are continuous data. Among these, the multi-layered models (CNN and HMAX) represent the hierarchical processing of human visual systems. GIST provides a low-dimensional representation of a scene, specified for scene recognition. SIFT + BOF is similar to GIST but is designed for object recognition.

Visual feature vectors of seen objects were predicted from the activity patterns of each ROI, based on a linear regression function. To build the prediction model, we used the code available on Horikawa et al.'s website (https://github.com/KamitaniLab/GenericObjectDecoding). The sparse linear regression (SLR; http://www.cns.atr.jp/cbi/sparse_estimation/index.html) (Bishop, 2006) was used for automatically selecting the important features for prediction. The regression function can be expressed as follows:

$$y(\mathbf{x}) = \sum_{i=1}^{d} w_i x_i + w_0$$

where x_i denotes the scalar value of the voxel i, w_i denotes the weight of voxel i, w_0 denotes the bias, and d denotes the number of voxels in an fMRI sample $\mathbf{x}$. For weight estimation, we adopted the variational Bayesian automatic relevance determination model (Sato, 2001; Tipping, 2001; Horikawa et al., 2013).

Hence, the weights of the regression function can be estimated by evaluating the following joint posterior probability of $\mathbf{w}$:

$$P(\mathbf{w}, \boldsymbol{\alpha}, \beta | \mathbf{X}, \boldsymbol{t_l}) = \frac{P(\boldsymbol{t_l}|\mathbf{X}, \mathbf{w}, \beta) P_0(\mathbf{w}|\boldsymbol{\alpha}) P_0(\boldsymbol{\alpha}) P_0(\beta)}{\int d\mathbf{w} d\boldsymbol{\alpha} d\beta P(\boldsymbol{t_l}, \mathbf{w}, \boldsymbol{\alpha}, \beta | \mathbf{X})}$$

where $\boldsymbol{t_l}$ denotes the target variable of the l^{th} component of a visual feature explained by the $y(\mathbf{x})$ with additive Gaussian noise; $\mathbf{w}$, the weight vector of regression function; $\boldsymbol{\alpha}$, the weight precision parameters; and β, the noise precision parameter. The learning algorithm involves the maximization of the product of the marginal likelihood and the prior probabilities of $\mathbf{w}$, $\boldsymbol{\alpha}$, and β.

We trained linear regression models that predicted the feature vectors of the individual feature types/layers for seen objects of the fMRI samples during the training session. For the test dataset, fMRI samples corresponding to the same categories (35 samples in the test image session and 10 samples in the imagery session) were averaged across trials to increase the signal-to-noise ratio of the fMRI signals. Using the trained models, feature vectors of seen/imagined objects from averaged fMRI samples were predicted to construct one predicted feature vector for each test category. Model fitting and prediction were conducted for each feature unit. A total of 100 feature units were randomly selected for each visual feature. As a metric of decoding accuracy, we calculated the correlation coefficient between true and predicted feature values of the 50 test images.

Correlation coefficients of 100 units from five participants were calculated, providing 100 × 5 correlation coefficients in

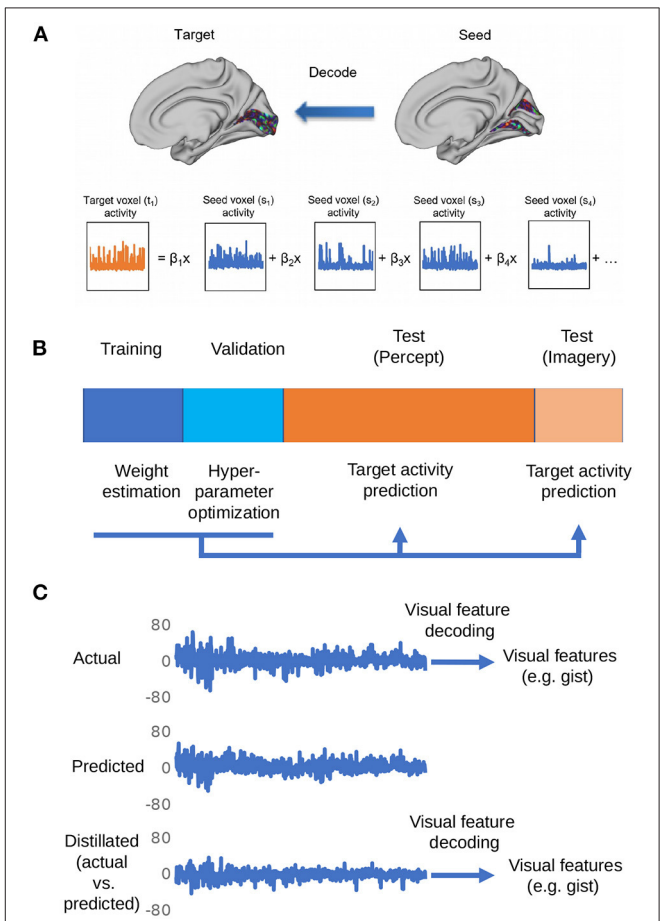

FIGURE 1 | Depiction of the procedures for the proposed distillation analysis. **(A)** Voxel activity in the target region is predicted by a weighted linear summation of all the voxel activities in the seed region. By repeating this procedure for all the voxels in the target region, the whole activity pattern of the target region is predicted. **(B)** The dataset is divided into four subsets: training, validation, perception, and imagery tests. The training data set is used to estimate weights for region-to-region decoding. The validation data set is used for the optimization of hyperparameters (ridge parameter). The weights estimated by the optimization of hyperparameters was used to predict the target region activity in perception and imagery test datasets. **(C)** The representative actual/predicted/distilled activity of a voxel in the perception test data.

each feature type/layer for each ROI. To evaluate the effect of the bottom-up/top-down signal elimination, the prediction modes were built before and after the signal elimination. The significance of the signal elimination effect was examined using a paired t-test. The correlation coefficient was preferred because we focused on feature decoding where the pattern across feature units is more important than the absolute value of a single unit. The mean absolute error (MAE) and mean squared error (MSE) were additionally measured for validating the findings derived from correlation analysis.

Figure 1 shows the overall procedure of our analysis. There are three steps in total: region-to-region prediction, top-down/bottom-up signal elimination, and visual feature predictions.

3. RESULTS

3.1. Region-to-Region Decoding

First, we calculated the functional connectivity between ROIs associated with the visual processing. **Figure 2A** shows the connectivity matrix between the ROIs associated with visual object recognition, including the lower visual cortices (V1-V4), the lateral occipital complex (LOC), fusiform face area (FFA), and parahippocampal place area (PPA). The nearby regions, for instance, V1 and V2, exhibited a strong connectivity (Pearson's correlation, mean $r = 0.96$) whereas that between the distant regions such as V1 and PPA, was weaker ($r = 0.66$).

Using each ROI as a seed, a linear ridge regression was performed to predict the activity of all the other ROIs. **Figure 1C** shows an example of the predicted activity at V1 based on the activity of V2 and its actual activity. We employed the optimal ridge parameter which best predicted the activity in the validation dataset of each seed–target combination for each subject. Similar to the connectivity, the R^2 was high between nearby regions and low between distant regions (**Figures 2B,C**). Particularly, the R^2 in imagery test was relatively lower than those in perception test, suggesting that the effectiveness of region-to-region decoding may differed according to behavioral task.

3.2. Distillation Analysis

The brain activity at a specific region (target region) was subtracted from its predicted activity based on the seed region. Here, the two poles of the visual object recognition were selected, i.e., V1 and FFA. A decoder (Horikawa and Kamitani, 2017) was used to predict the value of the visual features using the multi-voxel fMRI signals of these ROIs. Subsequently, the quality of the prediction was evaluated based on its correlation with the original visual feature.

In the image perception task, the correlation between the GIST descriptors predicted by the V1 activity and the original signal significantly increased after eliminating the top-down signals from the FFA [**Figure 3A**; two-sided t-test after Fisher's z-transform, $t_{(499)} = 3.01$, $p < 0.05$ [uncorrected]]. Interestingly, a similar increment was also observed in the correlation between the GIST predicted by the FFA and the original after subtracting the bottom-up signals [$t_{(499)} = 22.67$, $p < 0.001$ [uncorrected]]. However, an opposite effect (negative effect) in the imagery task. The correlation between the predicted GIST descriptors and the original one declined at both V1 [$t_{(499)} = -5.22$, $p < 0.001$ [uncorrected]] and FFA [$t_{(499)} = -6.52$, $p < 0.001$ [uncorrected]] after distillation (**Figure 3B**). These results indicated that during the imagery task, the visual features at these cortices was similar, whereas they were dissimilar during the image perception task.

3.3. Effects of Distillation According to the Regional Seed–Target Pair

To investigate the effect of distillation in general, all possible seed–target pairs were analyzed. Subsequently, the difference in the correlation coefficient before and after distillation were arranged into a 7×7 matrix for every visual feature (**Figure 4**). The diagonals were omitted since they represent the self-distillation, which is the scope of the current analysis. In this matrix, the upper triangle represents the effect of the top-down signals whereas the lower triangle represents the effect of the bottom-up signals. The positive values indicate that the representations of a visual feature were enhanced by eliminating the modulation of the seed region, and vice versa.

Enhancement of visual feature representations in V1 were observed after eliminating the modulation from FFA for CNN2, HMAX1–HMAX3, and GIST in the image-based perception task (**Figure 4**). Conversely, the negative effects indicate that the representations of a visual feature were diminished by eliminating the modulation of the seed regions. Such effects were

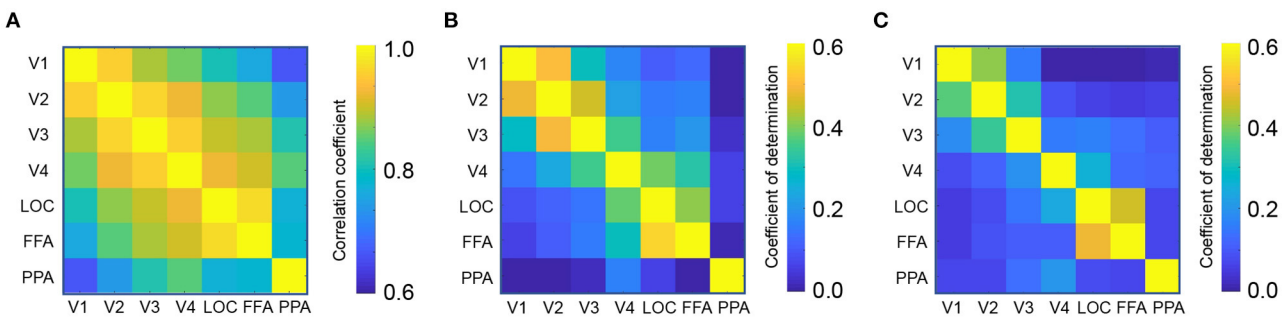

FIGURE 2 | Region-to-region decoding. **(A)** Connectivity matrix between visual object recognition related regions. **(B)** Coefficient of determination in the perception test. **(C)** Coefficient of determination in the imagery test.

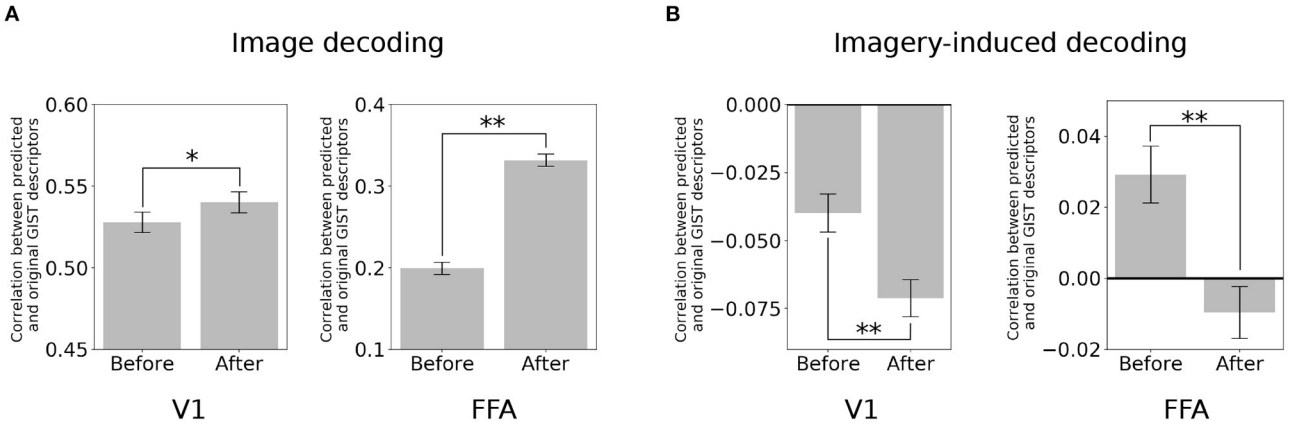

FIGURE 3 | Distillation analysis between V1 and FFA. **(A)** An example of V1 and FFA after distillation in predicting the GIST descriptors. Correlation between predicted and original GIST descriptors in the image-based perception task. **(B)** Correlation between predicted and original GIST descriptors in the class-based imagery task. Error bars represent standard error of measurement (SEM). * uncorrected $p < 0.05$, ** uncorrected $p < 0.001$; two-sided paired *t*-test after Fisher's z-transform; $n = 500$.

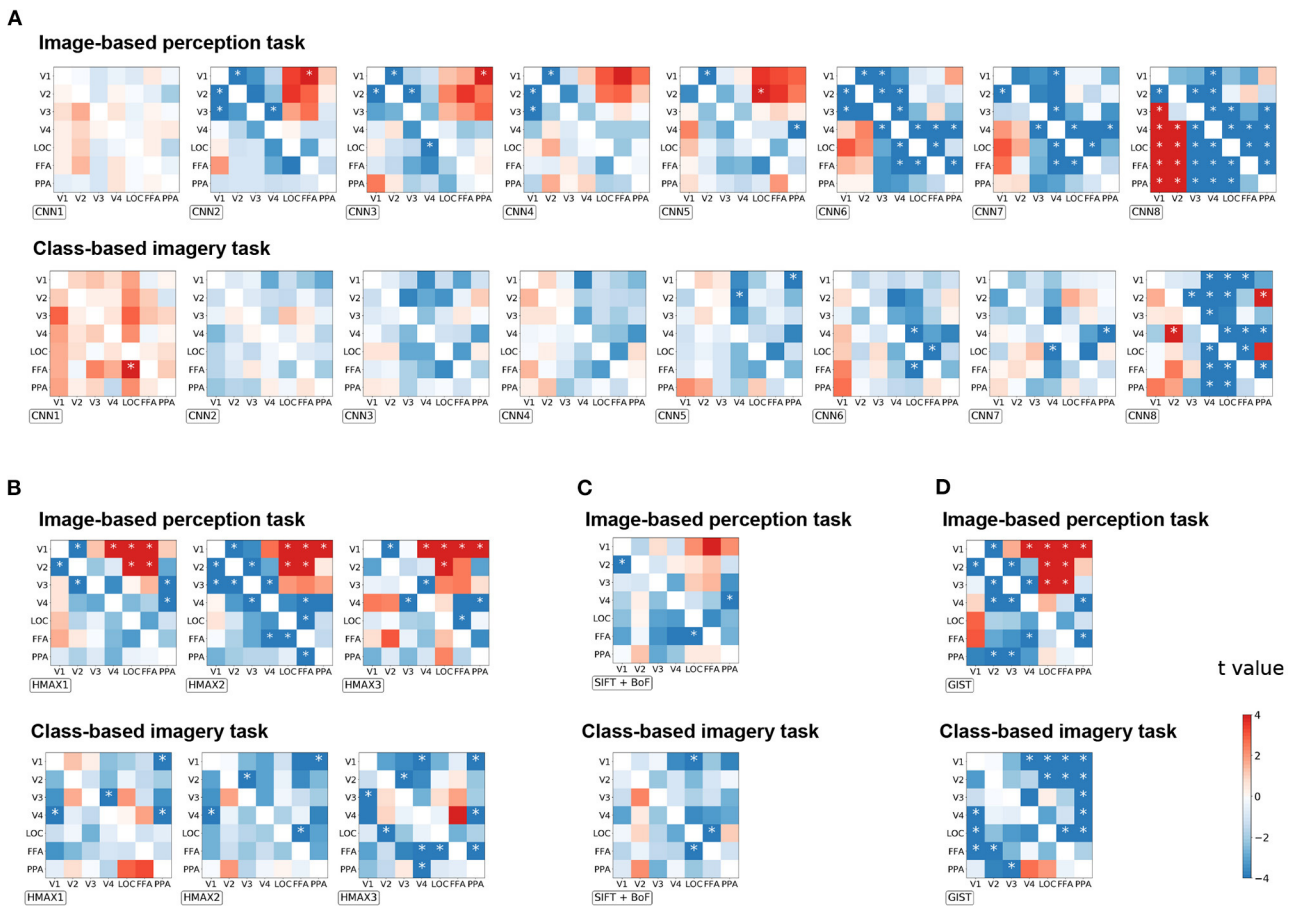

FIGURE 4 | The effect of distillation according to each seed–target pair for image-based perception task and class-based imagery tasks across all groups of visual features. **(A)** CNN visual features, **(B)** HMAX visual features, **(C)** SIFT + BOF, and **(D)** GIST. The difference of the correlation coefficients before and after distillation is arranged into a 7 × 7 matrix for every visual feature. For each group, the matrices are arranged as their complexity increased (from left to right). The diagonals are omitted since they represent the self-distillation which is not within the scope of this study. The horizontal axis represents the seed ROIs, whereas the vertical axis represents the target ROIs. The color bar indicates the t-value of the difference between the corresponding seed–target pair. *$p < 0.05$ after Bonferroni correction for multiple comparisons (*FWE* < 5%); two-sided paired *t*-test after Fisher's z-transform.

expected to be observed in the combination of brain regions that share the common information. Accordingly, the distillation for the nearby regional pairs caused a negative effect (the blue area) as the nearby regions express strong connectivity as shown in **Figure 2A**. Interestingly, for complex visual features such as CNN6–8, the negative effects were observed in the pairs of higher visual cortices, such as V4, LOC, FFA, and PPA. These negative effects suggest mutual dependence across the higher visual cortices during the processing of complex visual features.

As expected in the imagery task, almost all pairs had a negative effect, suggesting the existence of a close interaction between the visual cortices during this task (**Figure 4B**). Specifically, a negative effect was prominent in the upper triangle, indicating that top-down modulation is crucial for visual feature representations during the imagery task. Furthermore, an opposite effect in the imagery task compared to that in the image-based perception task was prominent in CNN8 and GIST.

Since the correlation coefficient is not sensitive to the scaling of data and may be biased under some circumstance (Poldrack et al., 2020), the distillation effect was validated by measuring the MAE (**Supplementary Figure 1**) and MSE (**Supplementary Figure 2**). The matrices are arranged in a manner similarly to that of **Figure 4**. A similar effect of top-down/bottom-up distillation was observed between the MAE/MSE analysis and the correlation analysis. Several exceptions include HMAX3 and GIST visual features in the imagery task. The difference may be due to the fact that the region-to-region decoding in imagery tasks was more difficult, as the coefficient of determinations were lower than those in perception task (**Figure 2C**).

To quantitatively measure the effect of top-down and bottom-up distillation, the mean of the 3 × 3 squares at the upper right and lower left of the effect matrices were measured, respectively (**Figure 5A**). As the complexity of the features increased, the effect of top-down distillation decreased (**Figure 5B**), even becoming reversed effect after CNN6. In contrast, the effect of bottom-up distillation slowly increased. This phenomenon was prominent in the case of the image-based perception task. In the case of the class-based imagery task, the predicted features after top-down distillation worsened as the complexity increased.

4. DISCUSSION

We have demonstrated a region-to-region decoding technique capable of predicting the neural activity at one region based on

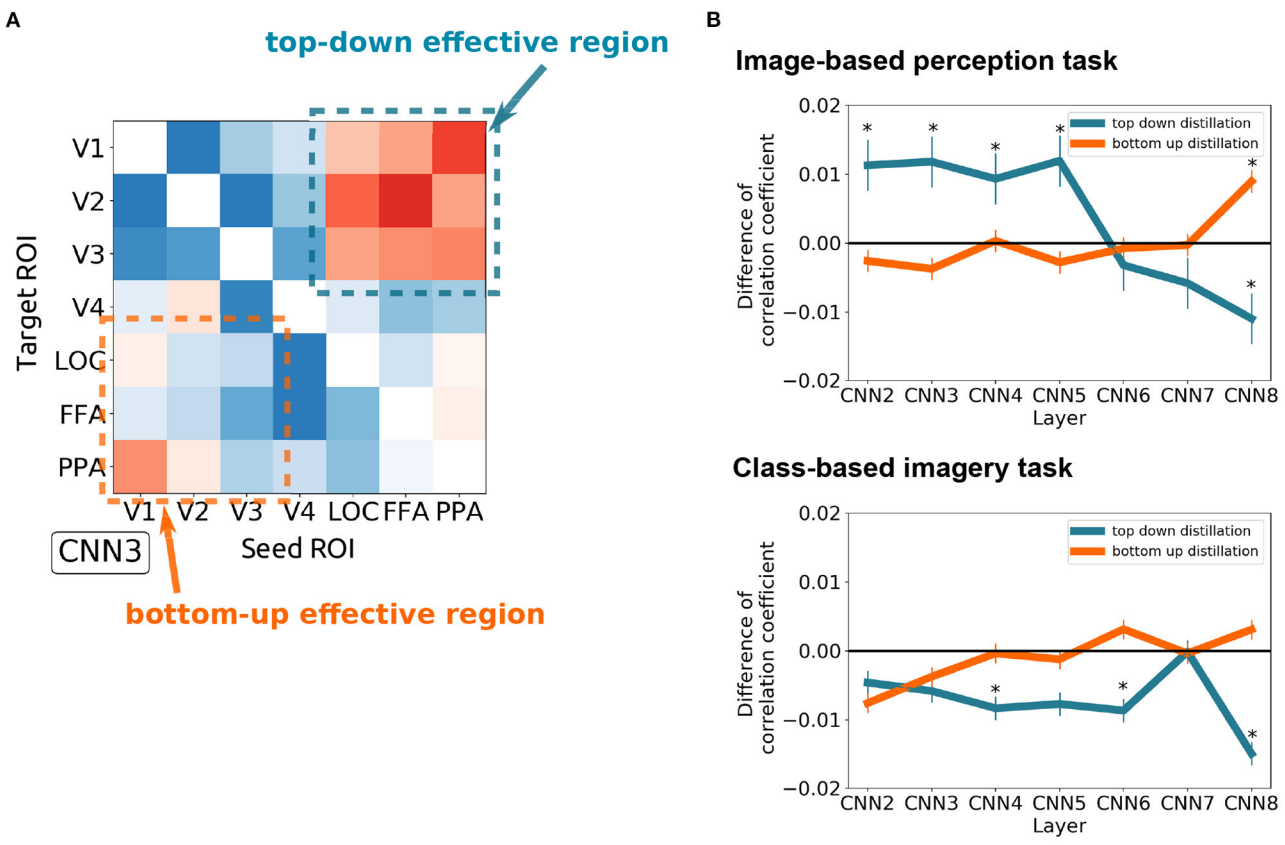

FIGURE 5 | The difference between the top-down distillation and bottom-up distillation. **(A)** An example of the CNN3 's effect matrix illustrating the measured regions. The mean of the 3 × 3 squares at the upper right and lower left of the effect matrices were measured. **(B)** The variance of the two kinds of distillation as the feature complexity increases (from CNN2 to CNN8) in case of image-based perception task (upper) and class-based imagery task (lower). Error bars represent SEM. *$p < 0.05$ after Bonferroni correction for multiple comparisons ($FWE < 5\%$); two sided one sample t-test after Fisher's z-transform.

the neural activity of another (**Figure 2**). By eliminating the top-down/bottom-up signals from the original signals (distillation approach), a significant change in the prediction of visual features from brain activity was found. Further analysis revealed three characteristics of the distillation approach.

First, the effectiveness of the distillation approach depends on the connectivity between the regional seed–target pairs. The seed–target pairs that exhibited weak connectivity were more suitable for distillation due to their distinct representations of information. The representations of information were alike between those with strong connectivity, hence, the distillation of top-down/bottom-up signals eliminated their original signals. Second, the distillation approach is also dependent on the type of task, as the imagery task evoked an effect opposed to that of the image-based perception task. Finally, the distillation approach specifies the direction and content of the conveyed information. For example, during an image-based perception task, representations of relatively simpler visual features such as CNN2–5, in V1 were enhanced by eliminating top-down modulations from FFA or PPA, whereas such effects were not observed for CNN6–8. Taken together, the distillation approach is a novel tool for estimating the direction and content of information conveyed across brain regions.

The difference between the visual feature matrices in **Figure 4** is expected since the complexity of visual features were found to increase in deeper layers. Previous studies (Horikawa and Kamitani, 2017) synthesized preferred images for each layer using the activation maximization technique, and found the increasing complexity, from simple edge detector representations to complex shapes and textures such as object parts. Conversely, the HMAX model was originally built to mimic the hierarchical processing of the ventral visual pathway. As mentioned in the original paper (Riesenhuber and Poggio, 1999), the HMAX model alternates layers of units by combining simple filters into more complex ones. Finally, SIFT + BoF and GIST are usually considered low-level features designed for object and scene recognition, respectively, as introduced in section 2.

Given our prior knowledge of vision, the FFA is expected to encode the information of facial expressions, and the PPA is expected to encode scene information. Such essential information is undoubtedly delivered from the lower visual cortex (such as V1) via the hierarchical bottom-up processing in the ventral stream. Therefore, they vanished after the bottom-up signals from V1 were distilled (**Figures 3**, **4**). Given a particular case of PPA, the scene information might be removed after distillation of the bottom-up signal, decreasing the decoding performance of the GIST in the PPA. Similarly, in the case of the FFA, the decoding performance of the SIFT + BoF which is related to encoding of facial expressions, decreased. Other information available at FFA and PPA but not directly related to the bottom-up signals, would be maintained. The decoding performance of other information such as GIST, HMAX1–3, and CNN8 increased at FFA. Thus, this result is in line with those of several previous studies in which FFAs were shown to hold information about non-face objects (Haxby et al., 2001; Kanwisher, 2017).

4.1. Forward-Backward Interaction Between the Low and High Visual Cortices During Imagery Task

Interestingly, our results show that the information of the visual features was lost after the distillation in the imagery task, suggesting a similarity between the representations of the low-level/high-level regions and their corresponding top-down/bottom-up signals. Due to the lack of concrete "actual" features from visual stimuli, the neural representation of the lower visual cortex aggressively employed the top-down signals. Neural representations at the high-level cortex were reinforced via bottom-up feedback. The quality of such reconstruction depends on the vividness of the imagined object or scene and on the available time for imagery.

Our results are in line with several previous studies (Stokes et al., 2009; Horikawa and Kamitani, 2017; Abdelhack and Kamitani, 2018), which depicts the mental imagery as a type of top-down processing. Previous studies reported the common neural representations during both perception and mental imagery (Albers et al., 2009, 2013; Stokes et al., 2009; Reddy et al., 2010; Cichy et al., 2012; Xing et al., 2013; Johnson and Johnson, 2014; Naselaris et al., 2015; Horikawa and Kamitani, 2017). Our analysis revealed the different compositions of these neural representations concerning their utilization of top-down/bottom-up signals.

4.2. Significance and Limitations of the Distillation Analysis

Whereas the functional connectivity and effective connectivity analyses focus on estimating the strength of regional connectivity, the distillation approach specifies the content of the conveyed information between regions as well as interregional connectivity. **Figure 5** demonstrates that the top-down modulation during the visual perception task could diminish the simpler visual features (i.e., CNN2–5) represented in the early visual cortices (i.e., V1–V3). With the distillation approach, the effect of top-down modulation can be analytically eliminated, resulting in the enhancement of simpler visual feature representations in the early visual cortices. It should be noted that the effects of distillation were not observed for any visual features. Rather, these effects were specific to the visual features of interest as well as the combination of the seed and target.

The current approach has two limitations. First, the distillation results can be artificial if the actual connectivity between the seed and target does not exist. Hence, results of the distillation analysis should be interpreted with caution to avoid misinterpretation. *In this study, both the correlation coefficient and the additional MAE/MSE were used as the metrics to evaluate the distillation results. A result was considered reliable if it was consistent across these metrics, i.e., the effect of distillation in perception task, and across CNN visual feature in imagery task.* The current results were derived from our prior assumptions, wherein the prediction of a low-level ROI based on the activity of a high-level region represents its top-down signals from that specific region, and vice versa. Second, the recurrent effect

was not considered (i.e., the effect caused by the circulation of information, from FFA to V1 to FFA), because its complexity may cause artifacts. Furthermore, we assume that the recurrent effect would be relatively smaller than the direct modulation at the moment of observation and, thus, could be negligible. To specify the precise modulatory effect between regions (e.g., from V1 to FFA), the recurrent effect should be included in the future model.

4.3. Potential Uses of the Distillation Analysis

The development of brain-computer interfaces (BCIs) will benefit from our distillation approach for precise decoding. Incorporating with the now mature image reconstruction by deep learning (Shen et al., 2019), our approach may reconstruct what people see even though it is not visually recognizable. One could also consider repeating the distillation analysis to obtain a better representation of the information at an ROI. Furthermore, the distillation analysis is applicable to the decoding of other sensory modalities, such as auditory and haptics feedback. Cross-distillation between different modalities would help us to gain better insights into their intervention in future work.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://github.com/KamitaniLab/GenericObjectDecoding.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethical Committee of the National Institute for Physiological Sciences of Japan. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

NS and JC contributed to the design and provided the conception and overall guidance for the project. SN, TP, and JC contributed to the data analysis and interpretation. TP and JC contributed to the initial drafting of the manuscript. All authors contributed to the writing, revision, and approval of the manuscript.

FUNDING

This work was supported by a grant from Japan Society for the Promotion of Science (JSPS) KAKENHI to TP (Grant Number 19K20390), JSPS KAKENHI to JC (Grant Number 21H02806, 18H05017, and 17H06033), and a grant from Japan Agency for Medical Research and Development (AMED) to JC (Grant Number JP21dm0207086) and NS (JP18dm0107152 and JP20dm0307005).

ACKNOWLEDGMENTS

We thank Y. Koyama for technical collaboration and discussion. Computational resources were provided by the Data Integration and Analysis Facility, National Institute for Basic Biology, and the Research Center for Computational Science.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnhum.2021.777464/full#supplementary-material

REFERENCES

Abdelhack, M., and Kamitani, Y. (2018). Sharpening of hierarchical visual feature representations of blurred images. *eNeuro* 5. doi: 10.1523/ENEURO.0443-17.2018

Albers, A., Kok, P., Toni, I., Dijkerman, H., and de Lange, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature* 458, 1476–4687. doi: 10.1038/nature07832

Albers, A., Kok, P., Toni, I., Dijkerman, H., and de Lange, F. (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

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Berlin: Springer-Verlag.

Cichy, R. M., Heinzle, J., and Haynes, J.-D. (2012). Imagery and perception share cortical representations of content and location. *Cereb. Cortex* 22, 372–380. doi: 10.1093/cercor/bhr106

Csurka, G., Dance, C. R., Fan, L., Willamowski, J., and Bray, C. (2004). "Visual categorization with bags of keypoints," in *Workshop on Statistical Learning in Computer Vision, ECCV* (Prague), 1–22.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. (2009). "ImageNet: a large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition* (Miami, FL: IEEE), 248–255. doi: 10.1109/CVPR.2009.5206848

Felleman, D. J., and Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1, 1–47. doi: 10.1093/cercor/1.1.1

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., and Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293, 2425–2430. doi: 10.1126/science.1063736

Heinzle, J., Kahnt, T., and Haynes, J.-D. (2011). Topographically specific functional connectivity between visual field maps in the human brain. *Neuroimage* 56, 1426–1436. doi: 10.1016/j.neuroimage.2011.02.077

Hoefle, S., Engel, A., Basilio, R., Alluri, V., Toiviainen, P., Cagy, M., et al. (2018). Identifying musical pieces from fmri data using encoding and decoding models. *Sci. Rep.* 8, 2045–2322. doi: 10.1038/s41598-018-20732-3

Horikawa, T., and Kamitani, Y. (2017). Generic decoding of seen and imagined objects using hierarchical visual features. *Nat. Commun.* 8:15037. doi: 10.1038/ncomms15037

Horikawa, T., Tamaki, M., Miyawaki, Y., and Kamitani, Y. (2013). Neural decoding of visual imagery during sleep. *Science* 340, 639–642. doi: 10.1126/science.1234330

Johnson, M., and Johnson, M. (2014). Decoding individual natural scene representations during perception and imagery. *Front. Hum. Neurosci.* 8:59. doi: 10.3389/fnhum.2014.00059

Kamitani, Y., and Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nat. Neurosci.* 8, 1546–1726. doi: 10.1038/nn1444

Kanwisher, N. (2017). The quest for the ffa and where it led. *J. Neurosci.* 37, 1056–1061. doi: 10.1523/JNEUROSCI.1706-16.2016

Kok, P., Brouwer, G. J., van Gerven, M. A., and de Lange, F. P. (2013). Prior expectations bias sensory representations in visual cortex. *J. Neurosci.*33, 16275–16284. doi: 10.1523/JNEUROSCI.0742-13.2013

Kok, P., Jehee, J. F. M., and de Lange, F. P. (2012). Less is more: expectation sharpens representations in the primary visual cortex. *Neuron* 75, 265–270. doi: 10.1016/j.neuron.2012.04.034

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). "ImageNet classification with deep convolutional neural networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1, NIPS'12* (Red Hook, NY: Curran Associates Inc.), 1097–1105.

Lowe, D. (1999). "Object recognition from local scale-invariant features," in *Proceedings of the Seventh IEEE International Conference on Computer Vision* (Kerkyra), 1150–1157. doi: 10.1109/ICCV.1999.790410

Mutch, J., and Lowe, D. G. (2008). Object class recognition and localization using sparse features with limited receptive fields. *Int. J. Comput. Vision* 80, 1573–1405. doi: 10.1007/s11263-007-0118-0

Naselaris, T., Olman, C. A., Stansbury, D. E., Ugurbil, K., and Gallant, J. L. (2015). A voxel-wise encoding model for early visual areas decodes mental images of remembered scenes. *Neuroimage* 105, 215–228. doi: 10.1016/j.neuroimage.2014.10.018

Oliva, A., and Torralba, A. (2001). Modeling the shape of the scene: a holistic representation of the spatial envelope. *Int. J. Comput. Vision* 42, 1573–1405. doi: 10.1023/A:1011139631724

Poldrack, R. A., Huckins, G., and Varoquaux, G. (2020). Establishment of best practices for evidence for prediction: a review. *JAMA Psychiatry* 77, 534–540. doi: 10.1001/jamapsychiatry.2019.3671

Reddy, L., Tsuchiya, N., and Serre, T. (2010). Reading the mind's eye: decoding category information during mental imagery. *Neuroimage* 50, 818–825. doi: 10.1016/j.neuroimage.2009.11.084

Riesenhuber, M., and Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nat. Neurosci.* 2, 1546–1726. doi: 10.1038/14819

Rissman, J., Gazzaley, A., and D'Esposito, M. (2004). Measuring functional connectivity during distinct stages of a cognitive task. *Neuroimage* 23, 752–763. doi: 10.1016/j.neuroimage.2004.06.035

Sato, M.-A. (2001). Online model selection based on the variational bayes. *Neural Comput.* 13, 1649–1681. doi: 10.1162/089976601750265045

Serre, T., Wolf, L., Bileschi, S., Riesenhuber, M., and Poggio, T. (2007). Robust object recognition with cortex-like mechanisms. *IEEE Trans. Pattern Anal. Mach. Intell.* 29, 411–426. doi: 10.1109/TPAMI.2007.56

Shen, G., Dwivedi, K., Majima, K., Horikawa, T., and Kamitani, Y. (2019). End-to-end deep image reconstruction from human brain activity. *Front. Comput. Neurosci.* 13:21. doi: 10.3389/fncom.2019.00021

Stokes, M., Thompson, R., Cusack, R., and Duncan, J. (2009). Top-down activation of shape-specific population codes in visual cortex during mental imagery. *J. Neurosci.* 29, 1565–1572. doi: 10.1523/JNEUROSCI.4657-08.2009

Tipping, M. E. (2001). Sparse bayesian learning and the relevance vector machine. *J. Mach. Learn. Res.* 1, 211–244. doi: 10.1162/15324430152748236

Xing, Y., Ledgeway, T., McGraw, P. V., and Schluppeck, D. (2013). Decoding working memory of stimulus contrast in early visual cortex. *J. Neurosci.* 33, 10301–10311. doi: 10.1523/JNEUROSCI.3754-12.2013

Conflict of Interest: JC and SN are employed by Araya Inc.

The remaining 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.

frontiers in Behavioral Neuroscience

REVIEW
published: 06 December 2021
doi: 10.3389/fnbeh.2021.756801

Reunification of Object and View-Center Background Information in the Primate Medial Temporal Lobe

He Chen[1]* *and Yuji Naya*[1,2,3]*

[1] School of Psychological and Cognitive Sciences, Peking University, Beijing, China, [2] IDG/McGovern Institute for Brain Research, Peking University, Beijing, China, [3] Beijing Key Laboratory of Behavioral and Mental Health, Faculty of Science, College of Psychology and Cognitive Sciences, Peking University, Beijing, China

Recent work has shown that the medial temporal lobe (MTL), including the hippocampus (HPC) and its surrounding limbic cortices, plays a role in scene perception in addition to episodic memory. The two basic factors of scene perception are the object ("what") and location ("where"). In this review, we first summarize the anatomical knowledge related to visual inputs to the MTL and physiological studies examining object-related information processed along the ventral pathway briefly. Thereafter, we discuss the space-related information, the processing of which was unclear, presumably because of its multiple aspects and a lack of appropriate task paradigm in contrast to object-related information. Based on recent electrophysiological studies using non-human primates and the existing literature, we proposed the "reunification theory," which explains brain mechanisms which construct object-location signals at each gaze. In this reunification theory, the ventral pathway signals a large-scale background image of the retina at each gaze position. This view-center background signal reflects the first person's perspective and specifies the allocentric location in the environment by similarity matching between images. The spatially invariant object signal and view-center background signal, both of which are derived from the same retinal image, are integrated again (i.e., reunification) along the ventral pathway-MTL stream, particularly in the perirhinal cortex. The conjunctive signal, which represents a particular object at a particular location, may play a role in scene perception in the HPC as a key constituent element of an entire scene.

Keywords: macaque monkey, medial temporal lobe, perirhinal cortex, inferotemporal cortex, ventral pathway, figure-ground segmentation, relational space, view-center background

OPEN ACCESS

Edited by:
Inah Lee,
Seoul National University,
South Korea

Reviewed by:
Sang Ah Lee,
Seoul National University,
South Korea
Sylvia Wirth,
Centre National de la Recherche Scientifique (CNRS), France

***Correspondence:**
He Chen
hechen2013@pku.edu.cn
Yuji Naya
yujin@pku.edu.cn

Specialty section:
This article was submitted to Learning and Memory, a section of the journal Frontiers in Behavioral Neuroscience

Received: *11 August 2021*
Accepted: *15 November 2021*
Published: *06 December 2021*

Citation:
Chen H and Naya Y (2021) Reunification of Object and View-Center Background Information in the Primate Medial Temporal Lobe. Front. Behav. Neurosci. 15:756801. doi: 10.3389/fnbeh.2021.756801

INTRODUCTION

Scene perception is a cognitive function used to construct a mental representation of the external world. The scene construction of primates, including humans, depends on the visual modality, which allows us to know "what" we are currently looking at, and "where" it is in the environment (**Figure 1**). Traditionally, the two aforementioned types of information are processed in different brain pathways (the two-stream theory) (Mishkin and Ungerleider, 1982; Haxby et al., 1991). The

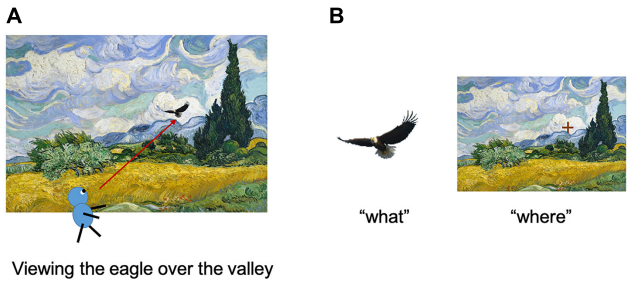

FIGURE 1 | Encoding of items and their location in a scene. **(A)** Assume that a subject is viewing the eagle in the wheat field. **(B)** The item information of the eagle (left) and its location in the environment (right) were acquired using this process. Adapted with permission from Chen and Naya (2020b).

first is called the ventral pathway, which deals with object-related information ("what"), while the other is called the dorsal pathway, which deals with space-related information ("where"). Previously, item encoding in the ventral pathway was considered to be accompanied by a loss of spatial information, such as the retinal position. Although this conventional theory is still dominant in neuroscience research, recent studies have reported that the ventral pathway likely represents space-related information in addition to object-related information (Nowicka and Ringo, 2000; Sobotka et al., 2002; Lehky et al., 2008; Kornblith et al., 2013; Vaziri et al., 2014; Connor and Knierim, 2017). However, researchers are yet to suggest an underlying substance for space-related information in the ventral pathway. The possible reasons for this may include a lack of appropriate task paradigms, which allows us to understand its unique implication distinct from that being extensively investigated in the "where" pathway.

In this review, we first briefly summarize the anatomical projections from the two streams of visual association areas to the medial temporal lobe (MTL). We then list the coding properties of the item signals and recently reported space-related information in the ventral pathway-MTL stream. Finally, we discuss our recent studies that investigated the neuronal representations of object and space-related information along the ventral pathway-MTL stream using a newly devised short-term memory task, called the item-location-retention (ILR) paradigm (Chen and Naya, 2020a,b). According to the preceding literature and our recent studies using the ILR paradigm, we propose a new hypothesis (reunification theory), which includes two conceptual advances. First, in addition to object information, the ventral pathway-MTL stream carries background information on the retina, including parafoveal vision, at each gaze. This view-center background signal could explain most of the space-related information in the ventral pathway-MTL stream reported by previous studies (Georges-François et al., 1999; Kornblith et al., 2013; Vaziri et al., 2014; Hong et al., 2016). Second, information on an object's position is acquired along the ventral pathway-MTL stream by a constructive process integrating the object and view-center background signals, which are derived from the same retinal image and separately processed along the ventral pathway. The reunification theory may provide a foundation for understanding the scene construction process to support the visual perception, as well as episodic memory.

ANATOMY OF THE PRIMATE TEMPORAL LOBE

According to the two-stream theory, item- and space-related processing from the primary visual cortex reaches the inferotemporal (IT) cortex (composed of TEO and TE) and posterior parietal cortex, through V4 and the middle temporal area (MT), along the ventral and dorsal pathways in macaques (i.e., occipito-temporal and occipito-parietal paths), respectively. These two pathways connect to the hippocampus (HPC) via the surrounding MTL cortical regions (Albright and Stoner, 2002; Kravitz et al., 2011, 2013).

Along the ventral pathway-MTL stream, the signal in V4 propagates to TEO and succeeds in TE (Distler et al., 1993; Saleem et al., 1993). Subsequently, the signal in TE propagates to the MTL through the PRC and reaches the HPC via the entorhinal cortex (ERC) (i.e., V4-TEO-TE-PRC-ERC-HPC) (Van Hoesen and Pandya, 1975; Suzuki and Amaral, 1994a; Lavenex and Amaral, 2000; Squire et al., 2004). Moreover, V4 connects ventromedially to the posterior subregion of the PHC (TFO) (Ungerleider et al., 2008), which then connects to its anterior subregion (TF/TH) (Suzuki and Amaral, 1994b). The signal in the anterior PHC propagates directly (Suzuki and Amaral, 1994b; Ho and Burwell, 2014) and indirectly (via the PRC) (Suzuki and Amaral, 1994a; Lavenex et al., 2004) to the ERC before the HPC (i.e., V4-PHC-[PRC]-ERC-HPC). It should be noted here that the PHC receives inputs from the early stages of the ventral pathway in addition to those from the dorsal pathway (see below).

Along the dorsal pathway-MTL stream, the signal in the MT propagates to the inferior parietal lobule (IPL) (Rozzi et al., 2006) and enters the MTL directly through the PHC (Cavada and Goldman-Rakic, 1989), or indirectly via the posterior cingulate cortex (PCC) and retrosplenial cortex (RSC) [i.e., MT-IPL-(PCC-RSC)-PHC-ERC-HPC] (Van Hoesen and Pandya, 1975; Vogt and Pandya, 1987; Suzuki and Amaral, 1994b; Morris et al., 1999; Kobayashi and Amaral, 2003, 2007; Kondo et al., 2005).

Separate visual processing of the item and space in the occipito-temporal and occipito-parietal paths in humans was revealed by functional brain imaging and patient studies (Haxby et al., 1991; Jeannerod et al., 1994; Grill-Spector et al., 2001). The connectivity of the ventral pathway (ventrolateral temporal lobe-PRC-HPC) (Kahn et al., 2008; Libby et al., 2012) and dorsal pathway-MTL streams (IPL-PCC-RSC-PHC-HPC) (Rushworth et al., 2006; Kahn et al., 2008; Margulies et al., 2009; Vincent et al., 2010; Libby et al., 2012) in humans is generally comparable with that in macaques (Kravitz et al., 2011, 2013; Ranganath and Ritchey, 2012; Libby et al., 2014). The two-stream theory has also been applied to rodent studies, especially those investigating the MTL system. In consistent with the primates, the rodent PRC-lateral ERC-HPC and postrhinal cortex (rodent homolog of PHC)-medial ERC-HPC circuits have been suggested to process visual items and space information in

parallel (Burwell et al., 1995; Burwell and Amaral, 1998; Witter et al., 2000; Eichenbaum, 2006; Furtak et al., 2007).

OBJECT CODING IN THE VENTRAL PATHWAY

Neurons in the higher visual areas receive inputs from each of the earlier brain areas with smaller receptive fields, either directly or indirectly. After simple algebraic operations in the early visual areas, hierarchically organized information processing realizes neurons displaying comparatively more complex response properties to objects along the ventral pathway (Felleman and Van Essen, 1991). The complexity of neuronal responses reportedly increases with the size of the receptive field in non-human primates. While the receptive field is 0.5–2° near the fovea in V1, it is typically 2–10° in V4. The receptive fields of the IT cortex neurons are enlarged further (10–30°) and substantially cover bilateral portions of the visual field (Hubel and Wiesel, 1968; Kobatake and Tanaka, 1994; Roe et al., 2012). Thus, while V1 neurons distinguish the orientation, spatial position, and movement direction of a small stick, IT cortex neurons respond to a large complex shape, containing multiple visual features, the selectivity of which does not depend on the size of the stimuli or retinal position (Schwartz et al., 1983; Kobatake and Tanaka, 1994).

Rather than a simple pixel-wise representation of the retinal image, we perceive a visual object by supposedly representing its inner spatial configurations, regardless of its actual size and position on the retina (Connor and Knierim, 2017). The transformation from retinal representation to relational representation proceeds sequentially along the anatomical hierarchy of the visual areas. For example, a spatial relationship among the points along an extended contour on the retina is combined to construct orientation tuning in V1. Changes in the orientation (e.g., abrupt for corners and gradual for curves) then construct curvature tuning in V4. After the curvatures are assembled into local configurations (e.g., eyes), their spatial relationship constructs coherent object tuning (e.g., faces) in the IT cortex. Therefore, relational coding is specialized for object processing at the expense of the loss of the absolute retinal position of the perceived object.

MNEMONIC EFFECTS ON OBJECT CODING IN THE MEDIAL TEMPORAL LOBE

While TE is located at the last stage of visual object perception (Miyashita, 1993; Sheinberg and Logothetis, 1997), the PRC is located at the entrance of the MTL and serves as a hub with converging inputs from a wide range of unimodal and polymodal association areas, including TE (Suzuki and Naya, 2014; Miyashita, 2019). A functional double dissociation between these two adjacent brain regions was reported by lesion studies using two distinct tasks requiring sensory perception (e.g., a color discrimination task) and recognition (e.g., delayed non-matching task) (Buckley et al., 1997). While TE is more responsible for visual perception, the PRC is critical for memory functions, such as item recognition (Buffalo et al., 1999; Buckley et al., 2001; Baxter, 2009).

Neurophysiological studies have further revealed different neural operations between the PRC and TE, particularly in the visual pair-association (PA) paradigm (Naya et al., 2001, 2003), which examines the semantic association memory of visual objects (Sakai and Miyashita, 1991). In the PA task, a visual object was presented as a cue stimulus, and after a delay period, monkeys were required to choose a particular visual object that had been assigned as a paired associate of the cue stimulus and should be retrieved from long-term memory. During the PA task, a substantial number of neurons in both the PRC and TE showed item-selective responses to cue stimuli (Naya et al., 2003). Among the item-selective neurons, some neurons showed a correlated response to the visual objects of the same pairs during the cue stimulus presentation ("pair-coding neuron") (Sakai and Miyashita, 1991). The proportion of pair-coding neurons dramatically increased when the visual signal was transmitted forward, from TE (4.9% of the item-selective neurons) to the PRC (33%) (Naya et al., 2003). In addition to the pair-coding neurons, the PA task revealed a separate group of neurons that represented the to-be-retrieved target, the paired associate of the cue stimulus ("pair-recall neurons") (Sakai and Miyashita, 1991). The memory retrieval signal appeared in the PRC, even during the cue stimulus presentation (∼200 ms after cue onset), while neurons in TE were gradually recruited to display the retrieved visual objects (∼500 ms) after the emergence of the memory-retrieval signal in the PRC (Naya et al., 2001). These findings suggested backward spreading of the memory-retrieval signal from the PRC to TE, which was supported by a following study simultaneously recording from the PRC and TE (Takeda et al., 2015). Together, the series of neurophysiological studies using the PA paradigm showed that neurons in the PRC contribute to item-item association memory, while those in TE provide the PRC with item signals and receive to-be-retrieved item information from the PRC.

SCENE-SELECTIVE RESPONSE IN THE VENTRAL PATHWAY-MEDIAL TEMPORAL LOBE STREAM

The brain regions responsible for scene processing were explored in human neuroimaging studies, which compared BOLD signals when human subjects viewed scene-like stimuli and object-like stimuli (e.g., face). The scene-selective regions included the human transverse occipital sulcus (TOS), RSC, and parahippocampal place area (PPA) regions (Epstein and Kanwisher, 1998a,b; Epstein et al., 2003, 2008). The aggregated anatomical localization of these regions along the dorsal pathway-MTL stream has led to the conclusion that the dorsal pathway-MTL stream is exclusively involved in scene processing for memory-guided behavior, in particular, spatial navigation (Ranganath and Ritchey, 2012; Epstein et al., 2017).

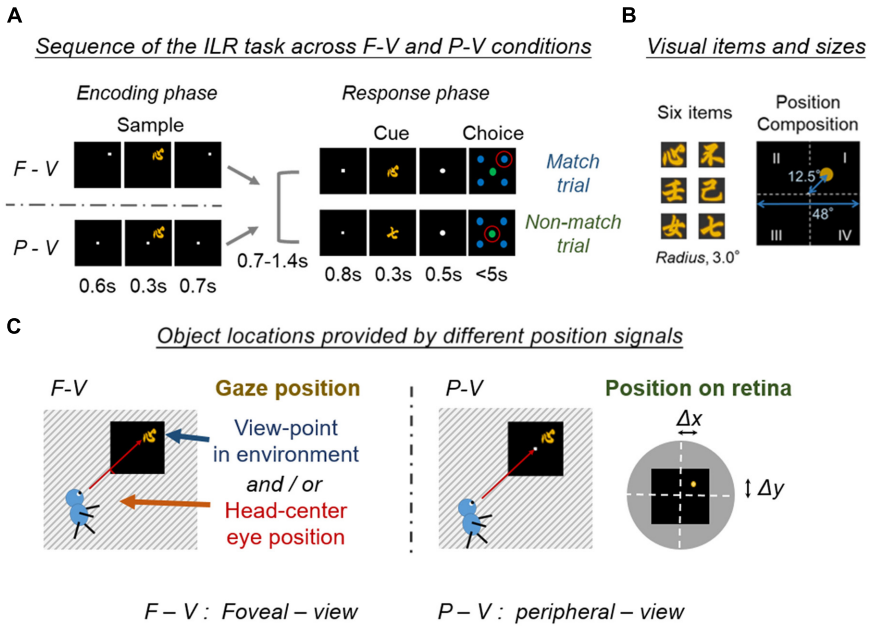

FIGURE 2 | Encoding of the location and item in the two viewing conditions. **(A)** Schematic diagram of the location and item encoding the F-V and P-V conditions of the ILR task. The cue stimulus was the same as the sample stimulus in the match trial (Top), while the two stimuli differed in the non-match trial (Bottom). Red circles indicate the correct answers. Adapted with permission from Chen and Naya (2020b). F-V, foveal-view; P-V, peripheral-view; ILR, item-location-retention. **(B)** Six visual item stimuli, and a spatial composition during the sample period. Adapted with permission from Chen and Naya (2020a). **(C)** Object locations provided by different position signals between the F-V and P-V conditions. The task screen is indicated by a black square and the surrounding environment is indicated by a stripe. The object location was provided by the gaze position in the F-V condition, while the object location was provided by the retinal position in the P-V condition. Note that the gaze position in the F-V condition can be defined by either the eye position in the head-center coordinate or the viewpoint in the environment. The object position on the retina could be defined as a shift from the center of the fovea (Δx, Δy). White dashed lines indicate the horizontal and vertical meridians of the visual field. The yellow dot indicates the center of the sample stimulus. Black square indicates the task screen. The gray disk indicates the surrounding environment in the visual field.

However, recent electrophysiological studies testing non-human primates as subjects have mentioned that single neurons in the ventral pathway can also code spatial information. For example, by combining electrophysiology with functional magnetic resonance imaging, Kornblith et al. (2013) identified two scene-selective brain regions, referred to as the lateral place patch (LPP) and the medial place patch (MPP). The LPP is located in the occipitotemporal sulcus and corresponded to TEOv, while the MPP is located in the posterior subdivision (TFO) of the PHC. Both regions receive strong connections from V4 (see section "The Anatomy of the Primate Temporal Lobe," above). According to the location proximity and selectivity to scene-like stimuli, these two brain regions have been suggested as homolog areas of the human PPA, which is located in the PHC (Epstein and Julian, 2013). These findings suggest that the response of the human PPA could be explained by the inputs from the ventral pathway-MTL stream, as well as by inputs from the dorsal pathway-MTL stream.

In addition to the ventral part of the temporal lobe just anterior to V4, another brain region was identified as showing spatial information by Vaziri et al. (2014). In the macaque TEd, the majority of neurons responded strongly to large-scale environmental stimuli, in contrast to the weak response to object-sized stimuli. These scene-selective areas in the macaque temporal lobe (i.e., LPP, MPP, and TEd) receive visual signals directly or indirectly from V4 (see section "The Anatomy of the Primate Temporal Lobe," above), which is suggested as the beginning point of figure-ground segmentation (Roe et al., 2012). We hypothesize that in addition to a figure, which is processed as an object, the ventral pathway would process a background, which might have been observed as a scene-selective response in previous studies (Epstein and Kanwisher, 1998a; Kornblith et al., 2013; Vaziri et al., 2014). Consistent with this idea, recent rodent studies have reported that the PRC of the ventral pathway-MTL stream contributes to spatial or visual scene processing as well as conventional item processing (Fiorilli et al., 2021; Lee et al., 2021).

LOCATION IN THE SCENE

In addition to the scene-selective responses, spatial information was observed in the ventral pathway for object position by applying a population decoding method to multi-unit recording (Hong et al., 2016). In this experiment, the subjects kept fixating at a central spot and passively viewed the object stimuli that were sequentially presented at different locations in the visual field. To examine the potential spatial information, the researchers gathered the multi-unit firing rates and conducted population decoding using a linear classifier. The success rate of decoding for the object position was higher in the IT cortex than in V4. Thus, the population decoding results suggest that the spatial

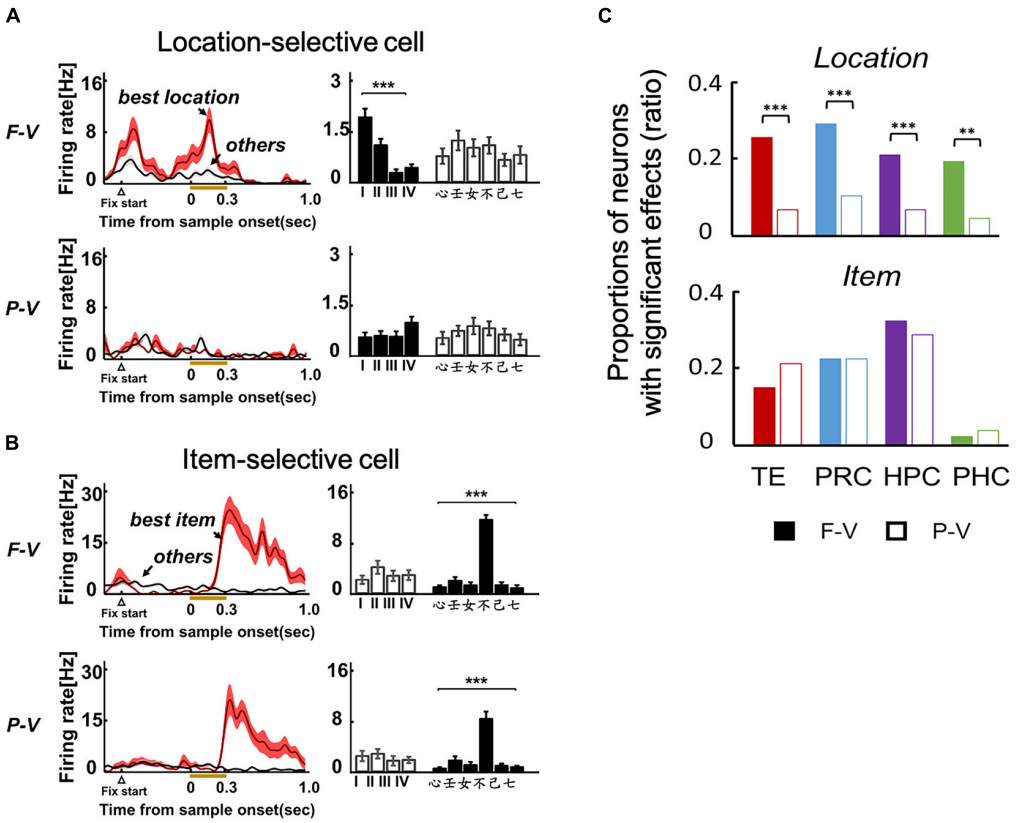

FIGURE 3 | Location and item selective responses in the two viewing conditions of the ILR task. **(A)** Example of a location-selective cell from TE in the F-V (Top) and P-V (Bottom) conditions of the ILR task. The neuron showed location-selective responses only in the F-V condition. **(B)** Example of an item-selective cell from TE. The neuron showed the same preferred item between the F-V and P-V conditions. ***$P < 0.0001$. Two-way ANOVA with interaction for each cell and view condition. **(C)** Proportions of location-selective cells (Top) and item-selective cells (Bottom) during the sample period in the F-V (Filled bars) and P-V (Open bars) conditions in the ILR task. **$P = 0.0011$, $X^2 = 13.24$, d.f. = 1 for PHC. ***$P < 0.0001$, $X^2 = 20.67$, 20.01, and 30.82 for TE, PRC, and HPC, respectively. *P* values were corrected by Bonferroni corrections among the four recording regions. Adapted with permission from Chen and Naya (2020b).

information was augmented at the late stage of the visual ventral pathway relative to its early stage, while spatial information could not be detected at the single neuron level.

Similarly, many previous visual neuroscience studies have examined the effect of object position while the subjects' eye position was kept at a central spot and the stimuli were presented in their peripheral visual field [peripheral-view (P-V) condition design]. These studies reported the spatially invariant object representation at least at the single neuron level (Schwartz et al., 1983; Desimone et al., 1984; Miyashita and Chang, 1988). However, when we look at an object in real life, we usually move our eyes toward it and automatically obtain its location in the surrounding environment.

To understand the possible neural patterns closer to everyday behavior, Chen and Naya (2020a; 2020b) adopted a foveal-view (F-V) condition design for the ILR task, which required subjects to encode the identity of a sample stimulus object and its location in each trial (**Figure 2**). In the F-V condition, subjects fixated on a white square presented within one of the four quadrants of a display. After fixation, one of the six visual objects was presented in the same quadrant as the sample stimulus. After this encoding phase, the response phase was initiated with a fixation dot presented at the center of the screen. One of the visual objects was then presented at the center as a cue stimulus. When the cue stimulus was the same as the sample stimulus, the subject was required to manually answer the sample position (i.e., match trial). Otherwise, the subject was required to choose the disk in the center, regardless of the sample position (i.e., non-match trial). Thus, the ILR task required subjects to encode and retain the identity and location of a sample object stimulus.

A substantial number of neurons (20–30%) exhibited location-selective responses in TE of the ventral pathway (**Figure 3A**), as well as the PRC, HPC, and PHC of the MTL in the F-V condition of the ILR task (**Figure 3C**, top). Neurons in these brain areas, except for the PHC, also showed item-selective responses (**Figures 3B,C**, bottom). The selectivity to the location and item was also examined in the P-V condition, in which the subjects maintained fixation to the center spot and the sample object was presented in the peripheral visual field (**Figure 2**). Although the ILR task required the subjects to acquire the same task-relevant information (e.g., "ﾉ乚" and "Position I") for the following response in both view conditions, the location-selective responses were substantially diminished in the P-V condition for TE (**Figure 3A**) and the MTL areas (**Figure 3C**, top). These results indicate that the ventral pathway-MTL stream signals an object position even at the single neuron

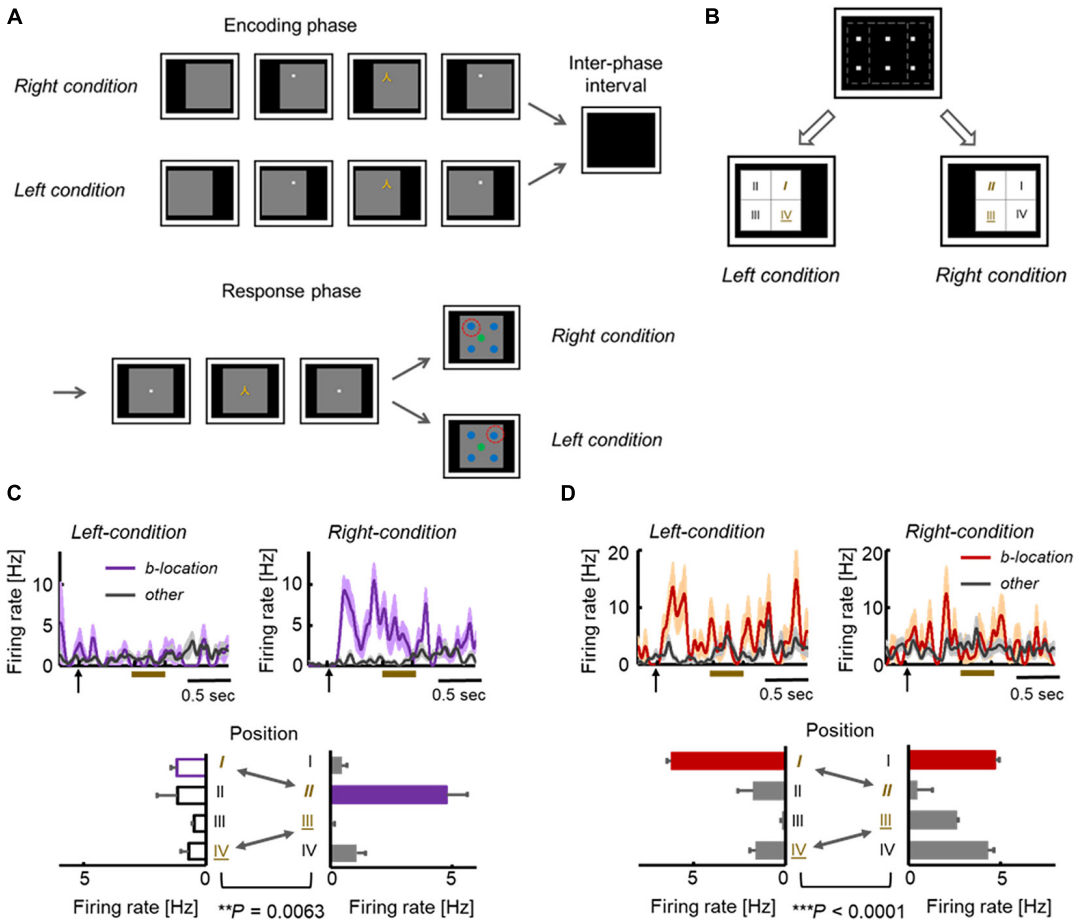

FIGURE 4 | Location-selective responses in conditional-type ILR Task. **(A)** Sequences in the conditional-type ILR task showing the left and right conditions. The time parameters are the same as those in the standard ILR task, except for the initial presentation of the gray square (0.5 s). **(B)** Spatial configuration of the stimulus positions in the left and right conditions. The first and fourth quadrants in the left condition ("I" and "IV") were also used as the second and third quadrants in the right condition ("II" and "III"). **(C)** Example of an HPC neuron in which responses to the same fixation position differed between the two conditions. **(D)** Example of a TE neuron showing significantly different responses between the two conditions. Adapted with permission from Chen and Naya (2020a).

level when subjects look at an object by their foveal vision, because neurons exhibit robust activities that are selective to gaze positions. In contrast, item selectivity did not differ between the two view conditions (**Figures 3B,C**, bottom), confirming the spatially invariant representation of object information.

The gaze-dependent activity could be explained by the following factors (**Figure 2C**). One type of factor is derived from the somatosensory (proprioception) and motor systems, which are directly related to the internal control of eye movements. The other type is derived from the visual system, which provides information on a retinal image at each gaze position. To dissociate the two types of gaze-related effects, additional experiments were conducted by modifying the F-V condition of the ILR task (i.e., conditional-type ILR task). In the modified version of this task, a large gray square was first presented on the right or left side of the display. Subsequently, the fixation dot and sample stimulus were presented sequentially during the encoding phase in the same way as in the F-V condition (**Figures 4A,B**). After the inter-phase interval with a blank screen, a large gray square was presented in the center of the display in both right and left encoding conditions, and the subject was required to answer the sample position relative to the large gray square in the match trials.

Across TE and the MTL areas, neurons do not necessarily represent the gaze positions themselves. For example, **Figure 4C** shows an HPC neuron that exhibited location-selective activity only in the right condition. This neuron exhibited the strongest response to the second quadrant in the right condition. However, the same neuron showed only negligible responses to the first quadrant in the left condition, although the two were at the same positions in the head-center coordinates (**Figure 4B**). Some other neurons showed preferred responses to the same quadrant in both the left and right conditions (**Figure 4D**). In short, the responses of these neurons were related with the visual inputs, rather than the head-center gaze positions. As shown by these examples, the location selective activity in the F-V condition could not be explained by the gaze positions themselves. Presumably, neurons in the primate temporal lobe signal the

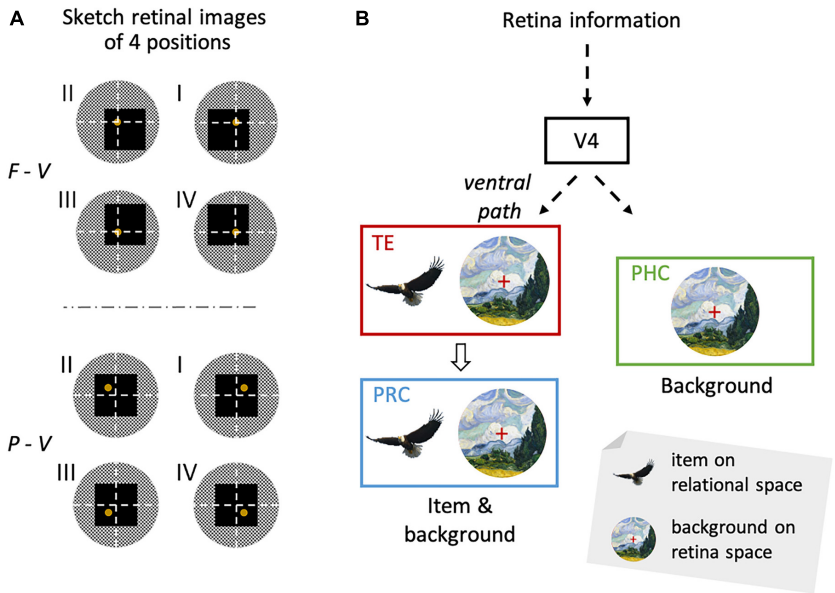

FIGURE 5 | Location signal reflecting background information and the item signal from the same retinal image. **(A)** Schematic diagram of visual inputs to the retinae during the sample period. White dashed lines indicate the horizontal and vertical meridians of the visual field. Yellow disks indicate the sample stimuli (i.e., yellow Chinese characters). Black squares indicate the task screen. Gray circles indicate the surrounding environment. Please note the larger change in the retinal images across different quadrants of the sample stimulus in the F-V condition compared with the P-V conditions. **(B)** The visual information from the retina reaches TE through the ventral pathway. In TE, there are not only item signals insensitive to the retinal position information, but also location signals provided by the background information sensitive to the retinal position information. The information has then been transferred to the PRC. While the PRC has two types of signals, the PHC has only the location signal. Adapted with permission from Chen and Naya (2020b). F-V, foveal-view; P-V, peripheral-view; ILR, item-location-retention.

retinal image reflecting background information (including the parafoveal vision), which would specify the current gaze position and object position in the F-V condition (**Figure 5A**, top). The overall shift in the retinal images across the different gaze positions would result in a substantial number of neurons exhibiting location-selective activity in the F-V condition, while the local change of the retinal image across the different sample positions would allow only a small number of neurons to show location-selective activity in the P-V condition (**Figures 3C**, **5A**). We refer to the information of the retinal image specifying gaze position as the "view-center background signal." Previous studies have elucidated the effect of eye position in the ventral pathway (Nowicka and Ringo, 2000; Lehky et al., 2008). Interestingly, Norwicka and Ringo revealed the effect of eye position on the responses of IT neurons under both light and dark conditions. However, separate neural populations exhibited eye position-sensitive responses between the two conditions (i.e., the presence and absence of visual inputs). Therefore, some IT neurons could be driven by eye-position-relevant inputs from non-visual modalities, but the visual input may be dominant in the primate IT cortex and their eye position selective responses could be explained by the view-center background, at least when it is available (i.e., light condition).

To examine the task dependence of the view-center background signal, Chen and Naya (2020a) performed another independent experiment free of memory demand using a passive-encoding task (Chen and Naya, 2020a). They found that the view-center background signal, which is accompanied by gaze behavior, is automatically encoded into the MTL along the ventral pathway. In contrast, the item signal was not detected in the passive-encoding task, regardless of the viewing conditions. The loss of item-selective response can presumably be attributed to a combination of the features of man-made complex stimuli (i.e., Chinese characters) and the passive-encoding task, which would not allow the sample stimulus to be segmented from the background as an object.

The view-center background signal is necessarily an egocentric spatial representation but may also provide us with allocentric spatial information of a target location in an environment. Rolls (1999) previously reported the HPC "view cells," which exhibited selective responses to a particular target location where a monkey gazed. This view-location-selective response was not sensitive to the monkey's self-position in a room, and consequently could not be explained by the gaze position itself. It can be assumed that the location-selective responses of view cells could be explained by the view-center background signal for the following reasons: the view-center background of a particular target location would be relatively similar across different self-positions compared with that of a different target location. For example, the retinal image changes according to the distance and direction between the retina and light sources, which depend on the subject's movement, but the relationship among segmental images of the neighboring retinal positions is generally maintained as long as the subject views the same target location. In contrast, different target view locations would produce different view-center background

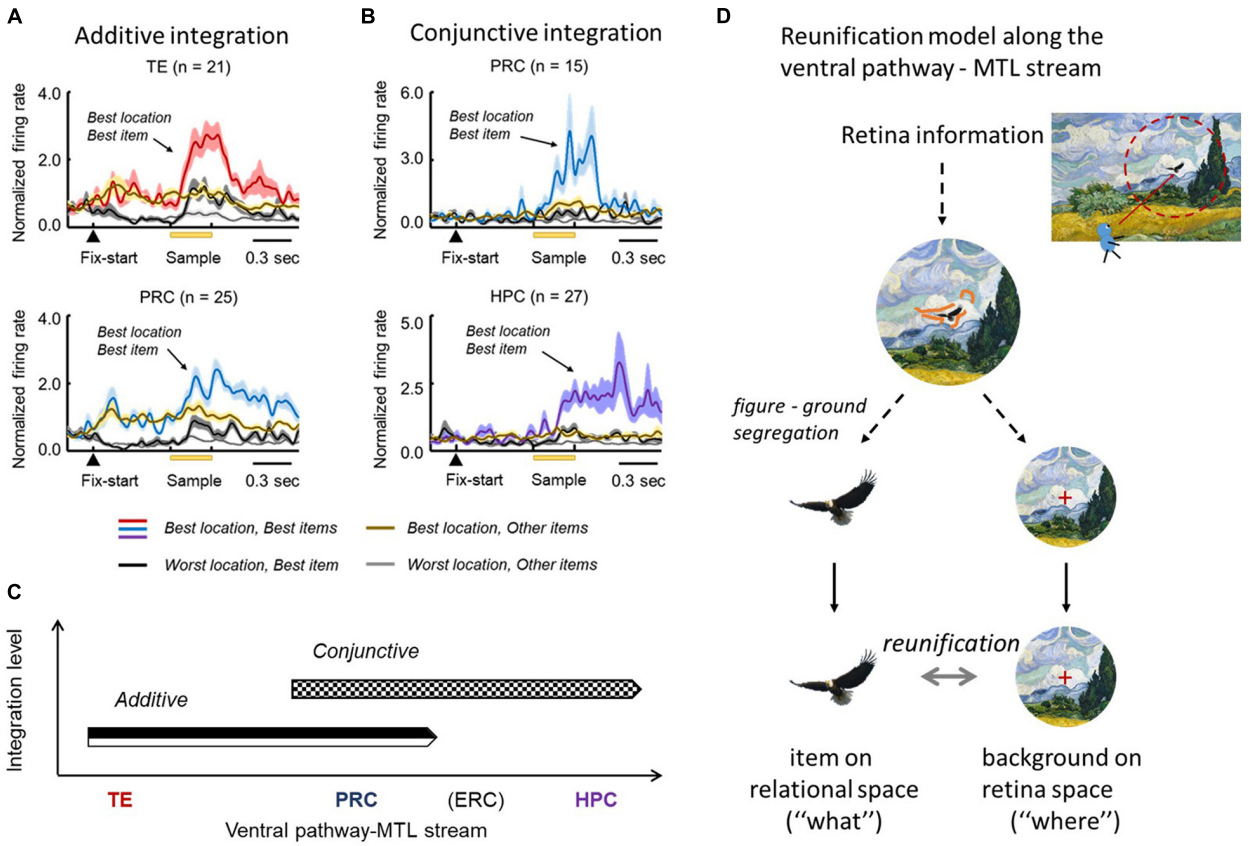

FIGURE 6 | Stepwise linkage of the location and item signals along the TE-PRC-HPC stream. **(A)** Population-averaged responses of neurons displaying additive integration in TE and the PRC in the F-V condition of the ILR task. These neurons started the location-selective responses after fixation, and the item-selective response was added after the sample presentation. **(B)** Population-averaged responses of neurons displaying conjunctive integration. Please note that the responses in the best location and best item demonstrate a robust increase, following the sample presentation in both the PRC and HPC. Adapted with permission from Chen and Naya (2020a). F-V, foveal-view; ILR, item-location-retention. **(C)** Additive and conjunctive types of integration process occurred sequentially along the ventral pathway-MTL stream. **(D)** Reunification model. Through the ventral pathway, the visual information from the retina first went through the figure-ground segregation process, where the item and location signals were segregated. After the two signals were separately processed, they were integrated again (i.e., "reunification"), at the single-unit level. The reunification process may be supported by the additive and conjunctive integrations along the ventral pathway—MTL stream.

signals. Thus, the existence of a view-center background signal reconciles the first person's perspective with the allocentric representation of the space.

The view-center background signal in the ventral pathway is considered as one aspect of space-related information, while the dorsal pathway is related with other aspects of space-related information such as a direction of gaze, a position with respect to head and a self-location in an environment (Olson et al., 1996; Ranganath and Ritchey, 2012). The different aspects of spatial processing in the ventral and dorsal pathways may require a modification of the "two-stream theory" but still acknowledge it (Goodale and Milner, 1992, 2018).

REUNIFICATION OF ITEM AND VIEW-CENTER BACKGROUND SIGNALS

In the F-V condition of the ILR task, the co-existence of item and view-center background signals (**Figures 3C**, **5B**) prompted the ventral pathway-MTL stream to integrate them at the single neuron level. In TE and the PRC, there were significantly larger numbers of neurons that exhibited both signals in an additive manner (**Figure 6A**). There were also other types of integration neurons in the PRC and HPC (**Figure 6B**). The second type of neuron exhibited activities selective to both the item and location of a sample object, only after the stimulus was presented. In other words, this type of integration neuron did not show location-selective activity before the sample stimulus presentation, even though the subjects gazed at a position where a sample stimulus would be presented. The neurons represented a combination of the identity and location of the sample stimulus rather than the two signals.

The two types of integrations proceeded sequentially along the ventral pathway-MTL stream (**Figure 6C**). Here, we propose a new theory that explains the neural mechanisms that encode the identity and location of an object in a scene (**Figure 6D**). After figure-ground segmentation, which reportedly occurs at V4 (Roe et al., 2012), the object information and background information

are separately processed along the ventral pathway, and the two signals are then reunified ("reunification model") to signal a particular item at a particular location. The representations of the object-related information and the space-related information by different neural ensembles in the reunification model contrast with the population coding which has rapidly prevailed in computational neuroscience, particularly after the invention of "deep learning" (Hong et al., 2016). Along the ventral pathway-MTL stream, the PRC may play a key role in the integration process to develop an additive representation into the conjunctive representation. We hypothesize that the object-location signal carried by the PRC conjunctive integration neurons would be transmitted to the HPC, in which additional association would proceed. The prominent integration effect of the PRC for perceptual processing may underlie the formation of long-term association memory in the PRC (see section "Mnemonic Effects on Object Coding in the MTL" above) (Suzuki and Naya, 2014).

In contrast to the reunification model for the object location relative to the background, Aggelopoulos and Rolls (2005) suggested that neurons in the macaque IT cortex encoded object locations relative to other objects presented at once in a passive fixation task. They reported that receptive fields of the IT neurons were reduced into the center of visual field as well as became more asymmetric in the presence of multiple objects compared with a condition when a single object was presented by itself. The reduced size of receptive fields may be explained by mutual inhibitions among neurons that were excited by the neighboring objects on the retinae (Aggelopoulos and Rolls, 2005) and/or by a lack of selective attentions to individual objects due to the competitions among the objects, which may influence the figure-ground segmentation. The processing of object information on the relational space might be an active process depending on the selective attention to a target either voluntarily (e.g., ILR task) or involuntarily (e.g., passive-encoding task with a single object). It would be useful to test the reunification model in a complex natural scene containing multiple objects in future studies.

SCENE CONSTRUCTION

The reunification model explains the perception or encoding of an object and space in a single snapshot of view. In each separate view, the TE-PRC-HPC stream provides the HPC with a conjunctive signal. However, multiple saccades are necessary to construct an entire scene in daily life. The across-saccades coordination might be mediated by the PHC, which receives inputs from the dorsal pathway and shows modulations of neuronal activity by eye-position/movement (Olson et al., 1996). We hypothesize that multiple shots of view-center backgrounds would be combined across saccades along the dorsal pathway-MTL stream. Furthermore, the combined space representation is integrated with the conjunctive signal from the TE-PRC-HPC stream in the HPC to construct a coherent scene, including objects from the first person's perspective on the perception and encoding of episodic memory (Baxter, 2009; Suzuki, 2009), which would contribute to a subsequent recollection-based recognition (Eichenbaum et al., 2007). Eventually, an intrinsic relationship among the multiple snapshots of views may also serve as an allocentric cognitive map in the HPC, which has often been reported in the context of spatial navigation that requires information about the environment (Vogt et al., 1992; Iaria et al., 2007; Burgess, 2008; Epstein, 2008; Kravitz et al., 2011; Zhang and Naya, 2020).

CONCLUSION

In contrast to the conventional two-stream theory, the ventral pathway carries as much space-related information as object-related information. The space-related information is likely substantiated by a large background image projected onto the retina. This view-center background image may not only provide the first person's perspective, but also specify a viewing location in an environment to provide allocentric spatial information. The object signal and the view-center background signal are transmitted along the ventral pathway-PRC-(ERC)-HPC stream when a subject is looking at an object by foveal vision. The two signals are integrated step-by-step to represent a particular object at a particular location along the stream, particularly in the PRC. We refer to this neural model as the "reunification theory" because the two signals are derived from the same retinal image. The object-location information based on the view-center background signal in the ventral pathway-MTL stream is more consistent with the distinction of the ventral/dorsal pathways based on their corresponding behavioral functions ("perception" and "action") (Goodale and Milner, 1992, 2018) than with that based on their typical representational contents ("what" and "where") (Mishkin and Ungerleider, 1982; Haxby et al., 1991). Accordingly, the reunification theory does not contradict but specifies the signals in the two-stream theory. Although the reunification theory can explain the perceptual process for a single snapshot of view, multiple saccades are required to perceive the entire scene. We hypothesized that the PHC could combine multiple view-center background images across saccades because it receives the space-related information from the dorsal pathway in addition to the view-center background signal. Future studies should investigate the neural mechanisms responsible for the construction of the entire background in the PHC, as well as for the construction of the entire scene, including objects in the HPC, which would support scene perception and episodic memory.

AUTHOR CONTRIBUTIONS

HC and YN wrote the manuscript. Both authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by the National Natural Science Foundation of China (Grants 31421003 and 31871139 to YN), the Fundamental Research Funds for the Central Universities, PKU (7100602954 to YN), and Peking University Boya Postdoctoral Fellowship (to HC).

REFERENCES

Aggelopoulos, N. C., and Rolls, E. T. (2005). Scene perception: inferior temporal cortex neurons encode the positions of different objects in the scene. *Eur. J. Neurosci.* 22, 2903–2916. doi: 10.1111/j.1460-9568.2005.04487.x

Albright, T. D., and Stoner, G. R. (2002). Contextual influences on visual processing. *Annu. Rev. Neurosci.* 25, 339–379. doi: 10.1146/annurev.neuro.25.112701.142900

Baxter, M. G. (2009). Involvement of medial temporal lobe structures in memory and perception. *Neuron* 61, 667–677. doi: 10.1016/j.neuron.2009.02.007

Buckley, M. J., Booth, M. C. A., Rolls, E. T., and Gaffan, D. (2001). Selective perceptual impairments after perirhinal cortex ablation. *J. Neurosci.* 21, 9824–9836. doi: 10.1523/JNEUROSCI.21-24-09824.2001

Buckley, M. J., Gaffan, D., and Murray, E. A. (1997). Functional double dissociation between two inferior temporal cortical areas: perirhinal *Cortex versus* middle temporal gyrus. *J. Neurophysiol.* 77, 587–598. doi: 10.1152/jn.1997.77.2.587

Buffalo, E. A., Ramus, S. J., Clark, R. E., Teng, E., Squire, L. R., and Zola, S. M. (1999). Dissociation between the effects of damage to perirhinal cortex and area TE. *Learn. Mem.* 6, 572–599. doi: 10.1101/lm.6.6.572

Burgess, N. (2008). Spatial cognition and the brain. *Ann. N. Y. Acad. Sci.* 1124, 77–97. doi: 10.1196/annals.1440.002

Burwell, R. D., and Amaral, D. G. (1998). Cortical afferents of the perirhinal, postrhinal, and entorhinal cortices of the rat. *J. Comp. Neurol.* 398, 179–205.

Burwell, R. D., Witter, M. P., and Amaral, D. G. (1995). Perirhinal and postrhinal cortices of the rat: a review of the neuroanatomical literature and comparison with findings from the monkey brain. *Hippocampus* 5, 390–408. doi: 10.1002/hipo.450050503

Cavada, C., and Goldman-Rakic, P. S. (1989). Posterior parietal cortex in rhesus monkey: I. Parcellation of areas based on distinctive limbic and sensory corticocortical connections. *J. Comp. Neurol.* 287, 393–421. doi: 10.1002/cne.902870402

Chen, H., and Naya, Y. (2020b). Forward processing of object-location association from the ventral stream to medial temporal lobe in nonhuman primates. *Cereb. Cortex* 30, 1260–1271. doi: 10.1093/cercor/bhz164

Chen, H., and Naya, Y. (2020a). Automatic encoding of a view-centered background image in the macaque temporal lobe. *Cereb. Cortex* 30, 6270–6283. doi: 10.1093/cercor/bhaa183

Connor, C. E., and Knierim, J. J. (2017). Integration of objects and space in perception and memory. *Nat. Neurosci.* 20, 1493–1503. doi: 10.1038/nn.4657

Desimone, R., Albright, T. D., Gross, C. G., and Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *J. Neurosci.* 4, 2051–2062. doi: 10.1523/JNEUROSCI.04-08-02051.1984

Distler, C., Boussaoud, D., Desimone, R., and Ungerleider, L. G. (1993). Cortical connections of inferior temporal area TEO in macaque monkeys. *J. Comp. Neurol.* 334, 125–150. doi: 10.1002/cne.903340111

Eichenbaum, H. (2006). Remembering: functional organization of the declarative memory system. *Curr. Biol.* 16, R643–R645. doi: 10.1016/j.cub.2006.07.026

Eichenbaum, H., Yonelinas, A. P., and Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annu. Rev. Neurosci* .30, 123–152. doi: 10.1146/annurev.neuro.30.051606.094328

Epstein, R. A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends Cogn. Sci.* 12, 388–396. doi: 10.1016/j.tics.2008.07.004

Epstein, R. A., and Julian, J. B. (2013). Scene areas in humans and macaques. *Neuron* 79, 615–617. doi: 10.1016/j.neuron.2013.08.001

Epstein, R. A., and Kanwisher, N. (1998a). A cortical representation of the local visual environment. *Nature* 392, 598–601. doi: 10.1038/33402

Epstein, R. A., and Kanwisher, N. (1998b). The parahippocampal place area: a cortical representation of the local visual environment. *Neuroimage* 7, S341–S341. doi: 10.1016/S1053-8119(18)31174-1

Epstein, R. A., Graham, K. S., and Downing, P. E. (2003). Viewpoint-specific scene representations in human parahippocampal cortex. *Neuron* 37, 865–876. doi: 10.1016/S0896-6273(03)00117-X

Epstein, R. A., Parker, W. E., and Feiler, A. M. (2008). Two kinds of fMRI repetition suppression? Evidence for dissociable neural mechanisms. *J. Neurophysiol.* 99, 2877–2886. doi: 10.1152/jn.90376.2008

Epstein, R. A., Patai, E. Z., Julian, J. B., and Spiers, H. J. (2017). The cognitive map in humans: spatial navigation and beyond. *Nat. Neurosci.* 20, 1504–1513. doi: 10.1038/nn.4656

Felleman, D. J., and Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1, 1–47. doi: 10.1093/cercor/1.1.1-a

Fiorilli, J., Bos, J. J., Grande, X., Lim, J., Düzel, E., and Pennartz, C. M. A. (2021). Reconciling the object and spatial processing views of the perirhinal cortex through task-relevant unitization. *Hippocampus* 31, 737–755. doi: 10.1002/hipo.23304

Furtak, S. C., Wei, S. M., Agster, K. L., and Burwell, R. D. (2007). Functional neuroanatomy of the parahippocampal region in the rat: the perirhinal and postrhinal cortices. *Hippocampus* 17, 709–722. doi: 10.1002/hipo.20314

Georges-François, P., Rolls, E. T., and Robertson, R. G. (1999). Spatial view cells in the primate hippocampus: allocentric view not head direction or eye position or place. *Cereb. Cortex* 9, 197–212. doi: 10.1093/cercor/9.3.197

Goodale, M. A., and Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends Neurosci.* 15, 20–25.

Goodale, M. A., and Milner, A. D. (2018). Two visual pathways – where have they taken us and where will they lead in future? *Cortex* 98, 283–292. doi: 10.1016/j.cortex.2017.12.002

Grill-Spector, K., Kourtzi, Z., and Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vis. Res.* 41, 1409–1422. doi: 10.1016/S0042-6989(01)00073-6

Haxby, J. V., Grady, C. L., Horwitz, B., Ungerleider, L. G., Mishkin, M., Carson, R. E., et al. (1991). Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proc. Natl. Acad. Sci. U.S.A.* 88, 1621–1625. doi: 10.1073/pnas.88.5.1621

Ho, J. W., and Burwell, R. D. (2014). "Perirhinal and postrhinal functional inputs to the hippocampus," in *Space, Time And Memory In The Hippocampal Formation*, ed. K. J. Derdikman (Vienna: Springer Vienna), 55–81.

Hong, H., Yamins, D. L., Majaj, N. J., and DiCarlo, J. J. (2016). Explicit information for category-orthogonal object properties increases along the ventral stream. *Nat. Neurosci.* 19, 613–622. doi: 10.1038/nn.4247

Hubel, D. H., and Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *J. Physiol.* 195, 215–243. doi: 10.1113/jphysiol.1968.sp008455

Iaria, G., Chen, J. K., Guariglia, C., Ptito, A., and Petrides, M. (2007). Retrosplenial and hippocampal brain regions in human navigation: complementary functional contributions to the formation and use of cognitive maps. *Eur. J. Neurosci.* 25, 890–899. doi: 10.1111/j.1460-9568.2007.05371.x

Jeannerod, M., Decety, J., and Michel, F. (1994). Impairment of grasping movements following a bilateral posterior parietal lesion. *Neuropsychologia* 32, 369–380. doi: 10.1016/0028-3932(94)90084-1

Kahn, I., Andrews-Hanna, J. R., Vincent, J. L., Snyder, A. Z., and Buckner, R. L. (2008). Distinct cortical anatomy linked to subregions of the medial temporal lobe revealed by intrinsic functional connectivity. *J. Neurophysiol.* 100, 129–139. doi: 10.1152/jn.00077.2008

Kobatake, E., and Tanaka, K. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *J. Neurophysiol.* 71, 856–867. doi: 10.1152/jn.1994.71.3.856

Kobayashi, Y., and Amaral, D. G. (2003). Macaque monkey retrosplenial cortex: II. Cortical afferents. *J. Comp. Neurol.* 466, 48–79. doi: 10.1002/cne.10883

Kobayashi, Y., and Amaral, D. G. (2007). Macaque monkey retrosplenial cortex: III. Cortical efferents. *J. Comp. Neurol.* 502, 810–833. doi: 10.1002/cne.21346

Kondo, H., Saleem, K. S., and Price, J. L. (2005). Differential connections of the perirhinal and parahippocampal cortex with the orbital and medial prefrontal networks in macaque monkeys. *J. Comp. Neurol.* 493, 479–509. doi: 10.1002/cne.20796

Kornblith, S., Cheng, X., Ohayon, S., and Tsao, D. Y. (2013). A network for scene processing in the macaque temporal lobe. *Neuron* 79, 766–781. doi: 10.1016/j.neuron.2013.06.015

Kravitz, D. J., Saleem, K. S., Baker, C. I., and Mishkin, M. (2011). A new neural framework for visuospatial processing. *Nat. Rev. Neurosci.* 12, 217–230. doi: 10.1038/nrn3008

Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G., and Mishkin, M. (2013). The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends Cogn. Sci.* 17, 26–49. doi: 10.1016/j.tics.2012.10.011

Lavenex, P., and Amaral, D. G. (2000). Hippocampal-neocortical interaction: a hierarchy of associativity. *Hippocampus* 10, 420–430.

Lavenex, P., Suzuki, W. A., and Amaral, D. G. (2004). Perirhinal and parahippocampal cortices of the macaque monkey: intrinsic projections

and interconnections. *J. Comp. Neurol.* 472, 371–394. doi: 10.1002/cne.20079

Lee, S. M., Jin, S. W., Park, S. B., Park, E. H., Lee, C. H., Lee, H. W., et al. (2021). Goal-directed interaction of stimulus and task demand in the parahippocampal region. *Hippocampus* 31, 717–736. doi: 10.1002/hipo.23295

Lehky, S. R., Peng, X., McAdams, C. J., and Sereno, A. B. (2008). Spatial modulation of primate inferotemporal responses by eye position. *PLoS One* 3:e3492. doi: 10.1371/journal.pone.0003492

Libby, L. A., Ekstrom, A. D., Ragland, J. D., and Ranganath, C. (2012). Differential connectivity of perirhinal and parahippocampal cortices within human hippocampal subregions revealed by high-resolution functional imaging. *J. Neurosci.* 32, 6550–6560. doi: 10.1523/JNEUROSCI.3711-11.2012

Libby, L. A., Hannula, D. E., and Ranganath, C. (2014). Medial temporal lobe coding of item and spatial information during relational binding in working memory. *J. Neurosci.* 34, 14233–14242. doi: 10.1523/JNEUROSCI.0655-14.2014

Margulies, D. S., Vincent, J. L., Kelly, C., Lohmann, G., Uddin, L. Q., Biswal, B. B., et al. (2009). Precuneus shares intrinsic functional architecture in humans and monkeys. *Proc. Natl. Acad. Sci. U.S.A.* 106, 20069–20074. doi: 10.1073/pnas.0905314106

Mishkin, M., and Ungerleider, L. G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behav. Brain Res.* 6, 57–77. doi: 10.1016/0166-4328(82)90081-x

Miyashita, Y. (1993). Inferior temporal cortex: where visual perception meets memory. *Annu. Rev. Neurosci.* 16, 245–263.

Miyashita, Y. (2019). Perirhinal circuits for memory processing. *Nat. Rev. Neurosci.* 20, 577–592. doi: 10.1038/s41583-019-0213-6

Miyashita, Y., and Chang, H. S. (1988). Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature* 331, 68–70. doi: 10.1038/331068a0

Morris, R., Pandya, D. N., and Petrides, M. (1999). Fiber system linking the mid-dorsolateral frontal cortex with the retrosplenial/presubicular region in the rhesus monkey. *J. Comp. Neurol.* 407, 183–192.

Naya, Y., Yoshida, M., and Miyashita, Y. (2001). Backward spreading of memory-retrieval signal in the primate temporal cortex. *Science* 291, 661–664. doi: 10.1126/science.291.5504.661

Naya, Y., Yoshida, M., and Miyashita, Y. (2003). Forward processing of long-term associative memory in monkey inferotemporal cortex. *J. Neurosci.* 23, 2861–2871.

Nowicka, A., and Ringo, J. L. (2000). Eye position-sensitive units in hippocampal formation and in inferotemporal cortex of the macaque monkey. *Eur. J. Neurosci.* 12, 751–759. doi: 10.1046/j.1460-9568.2000.00943.x

Olson, C. R., Musil, S. Y., and Goldberg, M. E. (1996). Single neurons in posterior cingulate cortex of behaving macaque: eye movement signals. *J. Neurophysiol.* 76, 3285–3300. doi: 10.1152/jn.1996.76.5.3285

Ranganath, C., and Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nat. Rev. Neurosci.* 13, 713–726. doi: 10.1038/nrn3338

Roe, A. W., Chelazzi, L., Connor, C. E., Conway, B. R., Fujita, I., Gallant, J. L., et al. (2012). Toward a unified theory of visual area V4. *Neuron* 74, 12–29. doi: 10.1016/j.neuron.2012.03.011

Rolls, E. T. (1999). Spatial view cells and the representation of place in the primate hippocampus. *Hippocampus*, 9, 467–480. doi: 10.1002/(SICI)1098-1063(1999)9:4<467::AID-HIPO13>3.0.CO;2-F

Rozzi, S., Calzavara, R., Belmalih, A., Borra, E., Gregoriou, G. G., Matelli, M., et al. (2006). Cortical connections of the inferior parietal cortical convexity of the macaque monkey. *Cereb. Cortex* 16, 1389–1417. doi: 10.1093/cercor/bhj076

Rushworth, M. F. S., Behrens, T. E. J., and Johansen-Berg, H. (2006). Connection patterns distinguish 3 regions of human parietal cortex. *Cereb. Cortex* 16, 1418–1430. doi: 10.1093/cercor/bhj079

Sakai, K., and Miyashita, Y. (1991). Neural organization for the long-term memory of paired associates. *Nature* 354, 152–155. doi: 10.1038/354152a0

Saleem, K. S., Tanaka, K., and Rockland, K. S. (1993). Specific and columnar projection from area TEO to TE in the macaque inferotemporal cortex. *Cereb. Cortex* 3, 454–464. doi: 10.1093/cercor/3.5.454

Schwartz, E. L., Desimone, R., Albright, T. D., and Gross, C. G. (1983). Shape recognition and inferior temporal neurons. *Proc. Natl. Acad. Sci. U.S.A.* 80, 5776–5778. doi: 10.1073/pnas.80.18.5776

Sheinberg, D. L., and Logothetis, N. K. (1997). The role of temporal cortical areas in perceptual organization. *Proc. Natl. Acad. Sci. U.S.A.* 94, 3408–3413. doi: 10.1073/pnas.94.7.3408

Sobotka, S., Zuo, W., and Ringo, J. L. (2002). Is the functional connectivity within temporal lobe influenced by saccadic eye movements? *J. Neurophysiol.* 88, 1675–1684. doi: 10.1152/jn.2002.88.4.1675

Squire, L. R., Stark, C. E., and Clark, R. E. (2004). The medial temporal lobe. *Annu. Rev. Neurosci.* 27, 279–306. doi: 10.1146/annurev.neuro.27.070203.144130

Suzuki, W. A. (2009). Perception and the medial temporal lobe: evaluating the current evidence. *Neuron* 61, 657–666. doi: 10.1016/j.neuron.2009.03.002

Suzuki, W. A., and Amaral, D. G. (1994a). Perirhinal and parahippocampal cortices of the macaque monkey: cortical afferents. *J. Comp. Neurol.* 350, 497–533. doi: 10.1002/cne.903500402

Suzuki, W. A., and Amaral, D. G. (1994b). Topographic organization of the reciprocal connections between the monkey entorhinal cortex and the perirhinal and parahippocampal cortices. *J. Neurosci.* 14(3 Pt 2), 1856–1877. doi: 10.1523/jneurosci.14-03-01856.1994

Suzuki, W. A., and Naya, Y. (2014). The perirhinal cortex. *Annu. Rev. Neurosci.* 37, 39–53. doi: 10.1146/annurev-neuro-071013-014207

Takeda, M., Koyano, K. W., Hirabayashi, T., Adachi, Y., and Miyashita, Y. (2015). Top-down regulation of laminar circuit via inter-area signal for successful object memory recall in monkey temporal cortex. *Neuron* 86, 840–852. doi: 10.1016/j.neuron.2015.03.047

Ungerleider, L. G., Galkin, T. W., Desimone, R., and Gattass, R. (2008). Cortical connections of area V4 in the macaque. *Cereb. Cortex* 18, 477–499. doi: 10.1093/cercor/bhm061

Van Hoesen, G. W., and Pandya, D. N. (1975). Some connections of the entorhinal (area 28) and perirhinal (area 35) cortices of the rhesus monkey. III. Efferent connections. *Brain Res.* 95, 39–59. doi: 10.1016/0006-8993(75)90206-1

Vaziri, S., Carlson, E. T., Wang, Z., and Connor, C. E. (2014). A channel for 3D environmental shape in anterior inferotemporal cortex. *Neuron* 84, 55–62. doi: 10.1016/j.neuron.2014.08.043

Vincent, J. L., Kahn, I., Van Essen, D. C., and Buckner, R. L. (2010). Functional connectivity of the macaque posterior parahippocampal cortex. *J. Neurophysiol.* 103, 793–800. doi: 10.1152/jn.00546.2009

Vogt, B. A., and Pandya, D. N. (1987). Cingulate cortex of the rhesus monkey: II. Cortical afferents. *J. Comp. Neurol.* 262, 271–289. doi: 10.1002/cne.902620208

Vogt, B. A., Finch, D. M., and Olson, C. R. (1992). Functional heterogeneity in cingulate cortex: the anterior executive and posterior evaluative regions. *Cereb. Cortex* 2, 435–443. doi: 10.1093/cercor/2.6.435-a

Witter, M. P., Naber, P. A., van Haeften, T., Machielsen, W. C. M., Rombouts, S. A. R. B., Barkhof, F., et al. (2000). Cortico-hippocampal communication by way of parallel parahippocampal-subicular pathways. *Hippocampus* 10, 398–410.

Zhang, B., and Naya, Y. (2020). Medial prefrontal cortex represents the object-based cognitive map when remembering an egocentric target location. *Cereb. Cortex* 30, 5356–5371. doi: 10.1093/cercor/bhaa117

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.

frontiers in Human Neuroscience

ORIGINAL RESEARCH
published: 26 January 2022
doi: 10.3389/fnhum.2021.798870

An fNIRS Study of Brain Lateralization During Observation and Execution of a Fine Motor Task

Kosar Khaksari[1]†, Elizabeth G. Smith[2†], Helga O. Miguel[1], Selin Zeytinoglu[3], Nathan Fox[3] and Amir H. Gandjbakhche[1]*

[1]National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD, United States,
[2]Department of Behavioral Medicine and Clinical Psychology, Cincinnati Children's Hospital, Cincinnati, OH, United States,
[3]Department of Human Development and Quantitative Methodology, University of Maryland, College Park, MD, United States

Brain activity in the action observation network (AON) is lateralized during action execution, with greater activation in the contralateral hemisphere to the side of the body used to perform the task. However, it is unknown whether the AON is also lateralized when watching another person perform an action. In this study, we use fNIRS to measure brain activity over the left and right cortex while participants completed actions with their left and right hands and watched an actor complete action with their left and right hands. We show that while activation is lateralized when the participants themselves are moving, brain lateralization is not affected by the side of the body when the participant is observing another person's action. In addition, we demonstrate that individual differences in hand preference and dexterity between the right and left hands are related to brain lateralization patterns.

Keywords: mirror-neuron system, lateralization, imitation, handedness, fNIRS, action observation network (AON)

OPEN ACCESS

Edited by:
Kuniyoshi L. Sakai,
The University of Tokyo, Japan

Reviewed by:
Xiaolong Liu,
Sichuan Normal University, China
Hendrik Santosa,
University of Pittsburgh, United States

****Correspondence:***
Kosar Khaksari
kosar.khaksari@nih.gov

†These authors have contributed equally to this work

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

Received: *20 October 2021*
Accepted: *28 December 2021*
Published: *26 January 2022*

Citation:
Khaksari K, Smith EG, Miguel HO, Zeytinoglu S, Fox N and Gandjbakhche AH (2022) An fNIRS Study of Brain Lateralization During Observation and Execution of a Fine Motor Task. Front. Hum. Neurosci. 15:798870. doi: 10.3389/fnhum.2021.798870

INTRODUCTION

The action observation network (AON) is comprised of brain regions that are active when watching another person execute an action (Lepage and Théoret, 2006; Cross et al., 2009; Condy et al., 2021). Self-initiated motor actions using one side of the body lead to lateralized activation patterns in the brain, such that they primarily activate contralateral motor cortex with more limited activation in ipsilateral cortical regions (Colebatch et al., 1991; Rao et al., 1993; Pulvermüller et al., 1995). However, lateralization patterns in the AON when watching another person act and their relation to lateralization for self-executed actions, are unknown. Lateralization patterns of brain activation during observation of others' actions and self-executed actions are important because they provide clues regarding the development and function of the AON. Specifically, brain lateralization patterns while watching another person's actions may clarify neural mechanisms of imitation, an important developmental task that influences play, social development, and learning throughout the lifespan (Acharya and Shukla, 2012; Kilner and Lemon, 2013).

Lateralization patterns in the AON during action observation could vary in the brain for imitation depending on the individual's perception. Specifically, imitation of others' motor behaviors can be characterized as either "mirror" imitation or "anatomical" imitation, depending on how the imitator maps their own body to the person they are imitating (Franz et al., 2007). Mirror imitation occurs when the imitator matches the side of the shared space with the other, as if they were looking in a mirror (e.g., mapping the

other's right hand to their left, see **Figure 1**). On the other hand, anatomical imitations occur when the imitator matches the side of their body as if they were mapping their body onto the other's (e.g., mapping their right hand with the other's right hand).

Likewise, activity in the AON can be either "mirror" oriented or "anatomically" oriented. Specifically, if the observed actor, who is sitting across from the observer, moves their right hand, it would lead to greater activation in the right relative to the left motor cortex for mirror representation (i.e., as if the observer were using their left hand). For anatomical representation in the brain, the same observed movement would lead to greater left relative to right motor cortex activation. Although electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) are the primary methods used to study the AON, brain lateralization of the AON is challenging to study using EEG due to limited spatial resolution or using fMRI due to effects of motion artifact on the MR signal (Condy et al., 2021).

Functional near infrared spectroscopy (fNIRS) is an ideal methodology for studying lateralization of the AON for several reasons (Fukuda and Mikuni, 2012). First, compared to other neuroimaging methods, it is relatively unperturbed by motion artifacts that are intrinsic to action-observation paradigms. That is, removal of trials with motion artifact and removal of motion artifact are less likely to influence results in fNIRS (Cooper et al., 2012; Brigadoi et al., 2014; Brigadoi and Cooper, 2015). In addition, while fNIRS does not have the spatial resolution of fMRI or MEG, it relies on localizable changes in oxygenated and deoxygenated hemoglobin on the cortical surface, and thus has the resolution to detect lateralization differences and to probe the primary motor cortex. Finally, fNIRS is highly tolerated, portable, and poses minimal risk, making it ideal for use in developmental science, including studies with children in imitation or within the AON. fNIRS uses near infrared (NIR) light to measure the diffusion of photons in human tissue or cortical tissue in the case of brain imaging. Hemodynamic fluctuations in the cortex are calculated as the difference in light absorbance of oxygenated hemoglobin (HbO) and deoxygenated hemoglobin (HbR) (Sassaroli et al., 2006; Yu et al., 2006). Thus, while fMRI is the gold standard for *in vivo* imaging of the human brain, fNIRS has exceptional portability and robustness to noise along with a higher temporal resolution (Strangman et al., 2002; Cutini et al., 2012; Gagnon et al., 2012; Wilcox and Biondi, 2015; Herold et al., 2017; Pinti et al., 2018).

Although there have been several fNIRS studies of the AON, most measured activation only in the contralateral hemisphere (Condy et al., 2021) and thus could not be used for the investigation of lateralization patterns. Those fNIRS studies that did measure activation across both hemispheres either did not have action and observation conditions across both the left and right hand (Bhat et al., 2017; Crivelli et al., 2018), did not use a live person as the "actor" (Holper et al., 2010), focused on atypical populations (Kajiume et al., 2013), or used a non-lateralized action (i.e., walking; Zhang et al., 2019).

The present study therefore uses fNIRS to measure lateralization patterns in brain activity while participants: (1) complete both left- and right-handed motor actions themselves; and (2) observe both right- and left-handed motor actions of an individual sitting across from them. Observation and self-action were completed across both the left and right hand while changes in blood oxygenation levels were measured across the left and right motor cortices. In addition, we also measured hand preference and manual dexterity. We hypothesized that: (1) for the self-action condition, participants would show stronger activation in the contralateral vs. ipsilateral cortex for use of both the left and right hands; (2) laterality of brain activations for observing the actor using their left vs. right hands would vary. Specifically, activation of left and right motor cortex would show opposite lateralization patterns reflecting either mirror representation (e.g., greater left vs. right activation when the actor sitting across from them used the left hand) or anatomical representation (i.e., greater right vs. left activation when actor sitting across from them used the left hand); and (3) we expected the strength of these patterns to be positively correlated with degree of handedness. Our alternative hypothesis was that activation while observing an actor move would either: (1) activate both left and right motor cortices similarly (indicating non-lateralization of the AON for this specific task) or (2) activate the brain as if the dominant hand were being used, independent of whether the actor was using their right or left hand (indicating a handedness or experience-driven lateralization of the AON).

METHODS

This study was approved by the Institutional Review Board at the University of Maryland. All participants provided written informed consent before start of procedures.

Participants and Measures. Participants were 41 undergraduate students (66% female) who were recruited through undergraduate psychology courses *via* SONA (20 ± 1.6 yrs.). Six subjects were Hispanic or Latino, 11 African American, 10 white, 11 Asian, 1 white/Asian, two more than one race, and two subjects did not report their race. Thirty-eight subjects were self-reported right-handers and three were left-handed. Left-handers were excluded from these analyses since there were not enough in the left-handed group for group-based analyses. Instead, continuous measures of hand preference and differential dexterity were used to determine effects of handedness. Handedness was evaluated through use of the Edinburgh handedness inventory (Oldfield, 1971; Jin et al., 2020) while manual dexterity was evaluated with the Purdue pegboard task (Tiffin and Asher, 1948; Hannanu et al., 2020). The EHI is a dimensional hand preference score—scores range from 0 to 50, with 0 indicating the highest possible preference for the left hand and 50 indicating the highest preference for the right hand. The Purdue Pegboard (Buddenberg and Davis, 2000) task involves placing as many pegs as possible within a standardized pegboard using either the right or left hand within 30 s. Two trials of 30 s were completed per hand across two different orders to reduce practice effects across the sample.

Self-action Task. Participants completed simple fine motor tasks across "left" and "right" conditions while seated at a

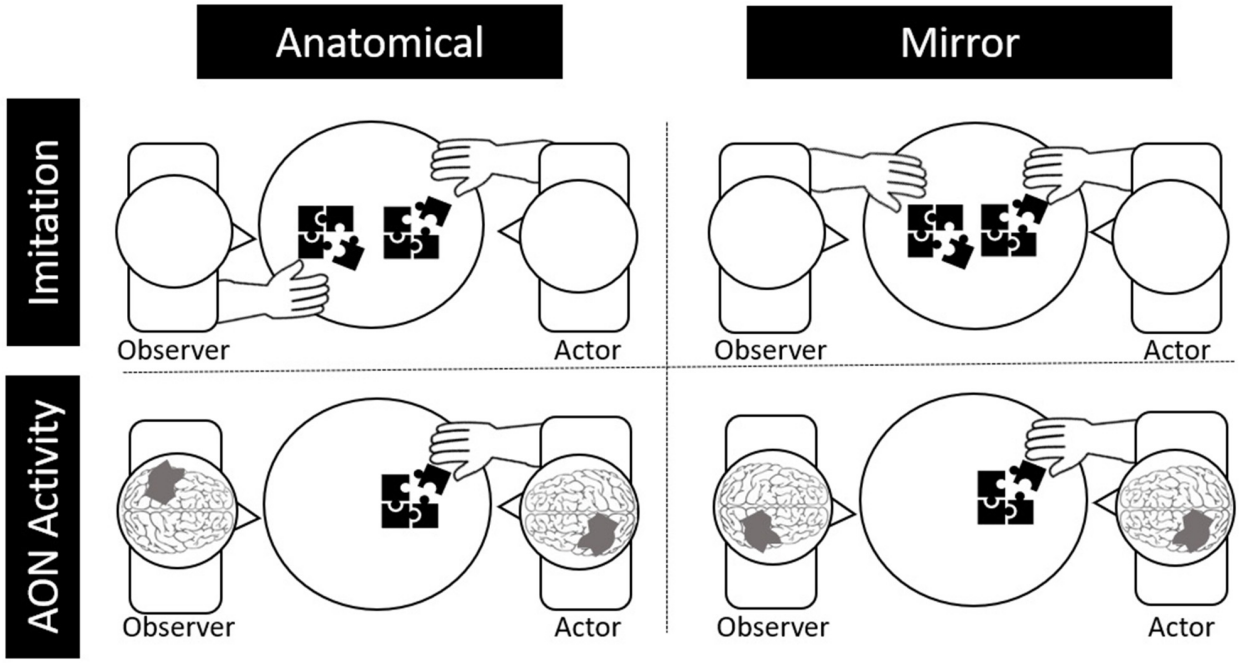

FIGURE 1 | Depictions of anatomical and behavioral imitation (top) and anatomical and behavioral neural activity during observation (bottom).

small table with a research assistant sitting opposite from them. These tasks were designed for downward expansion with young children and are similar to basic fine motor tasks that are challenging but achievable in the preschool years. The five fine motor tasks included: (1) putting plastic pennies in a bank; (2) moving small pom poms from one bowl to another with large plastic tweezers; (3) putting plastic pegs in a pegboard; (4) putting large wooden beads on a rod; and (5) turning large wooden bolts on a wooden screw. Each of these five tasks was completed across left and right hands (10 times total) for 15 s each time. Five tasks were used to maintain interest in extensions to younger age groups and to increase ecological validity. Tasks were randomly distributed across five blocks, and prior to each block the task objects (e.g., pennies and bank) were placed on the table in front of the participant and they were told whether they would be using their right or left hand. After an auditory trigger from the fNIRS interface subjects began performing the fine motor activity continuously for 15 s per trial. This resulted in five blocks of eight trials for both right and left hand.

Observation Task. Participants observed a trained research assistant who was sitting across from them complete the same five fine motor tasks described above with either their left or right hand. Participants were told to count the number of repetitions of each act to ensure their attention to the actor. Whereas for the Self-action task, participants were told which hand to use prior to the start of the block; for the Observation Task, the participant did not know which hand the research assistant would use but did see the task objects placed on the table in front of the research assistant. As in the self-action task, there were five trials in which the research assistant used their right hand and five trials in which they used their left, determined ahead of time through five semi-random orders. Self-action and observation trials were interleaved, as were 20 trials across two other conditions (observe next-to left and observe next-to right) which were not analyzed here. Between trials, there was a random jitter of 12–15 s of rest presented.

fNIRS System and Probe. Changes in oxyhemoglobin were measured using the Hitachi ETG 4000 fNIRS system with 24 channels (10 sources, eight detectors) distributed bilaterally (12 channels per side), with the most medial and anterior channels centered at CZ based on the international 10–20 electrode placement system (see **Figure 2**). The sampling rate was 10 Hz and source-detector separation distance was 3 cm for all the channels.

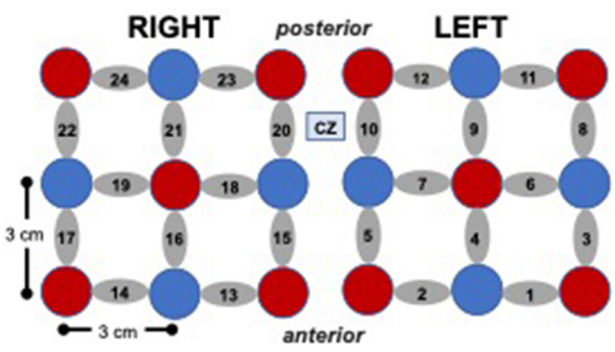

FIGURE 2 | Cap array, with red circles representing light sources, blue circles represent detectors, and gray ellipses representing channels.

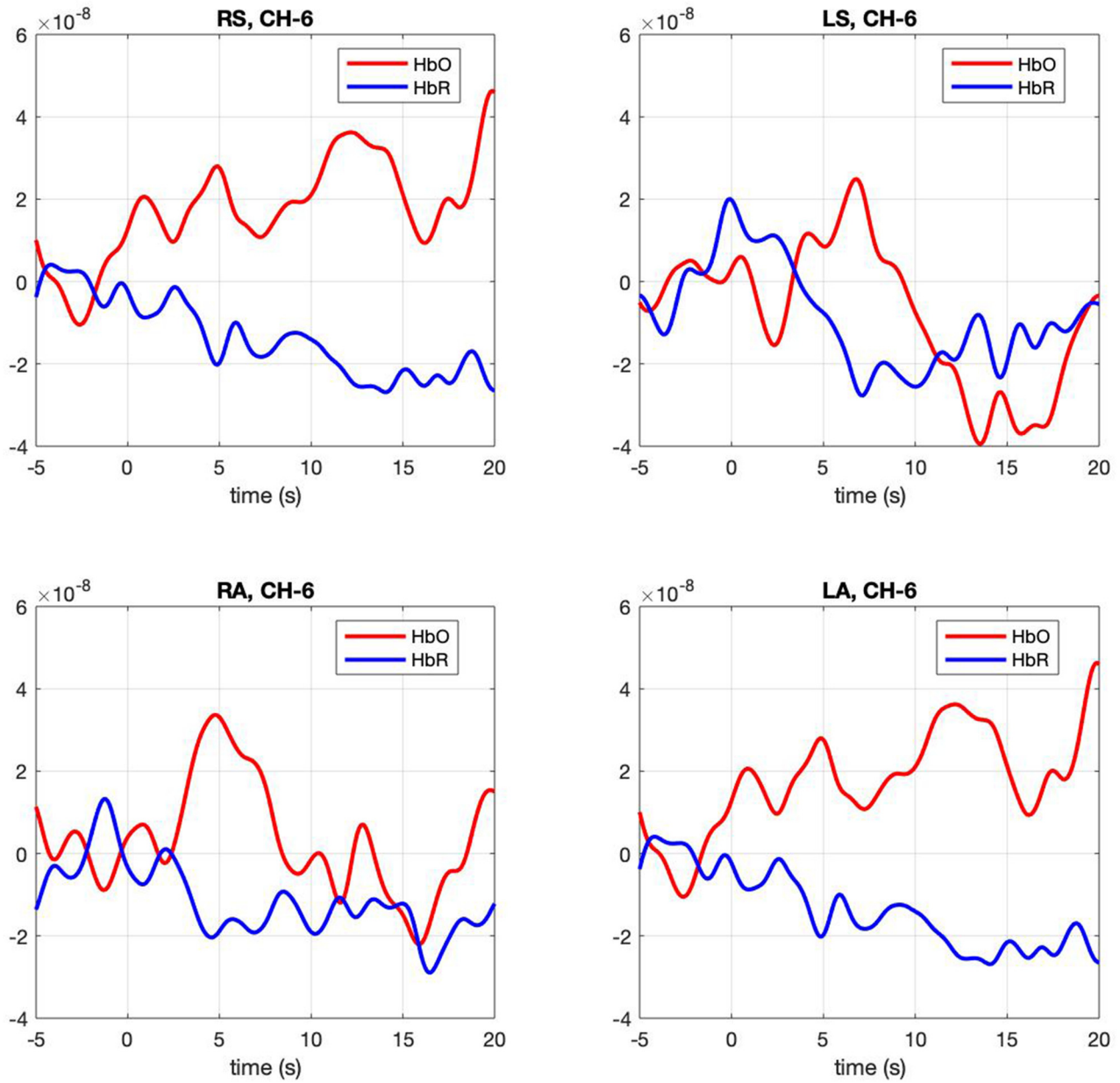

FIGURE 3 | Hemodynamic response, HbO (Red) and HbR (Blue), across four conditions (RS, right self; LS, left self; RA, right across; LA, left across) across the significant channel 6.

fNIRS data Processing. Raw Hitachi fNIRS data were processed using HOMER2 (Huppert et al., 2009) using the intensity data collected at two wavelengths (695 and 830 nm). Noisy channels (i.e., those from optodes with little to no scalp contact) were detected and removed using the HOMER2 Prune Channel function. Data were then transformed from intensity to optical density. Motion artifacts were removed using a wavelet motion correction filter of order 6. Biological (e.g., heartrate) and technical (e.g., motion) artifacts were removed using a low pass filter with a cutoff frequency of 0.1 Hz, while linear and nonlinear trends in the signals were removed by fitting a low order (an order of 3) polynomial to the fNIRS signals and subtracting it from the original signal (Huppert et al., 2009; Cooper et al., 2012; Dashtestani et al., 2019).

The processed optical density signals were then converted to oxyhemoglobin, deoxyhemoglobin, and total hemoglobin using the modified Beer-Lambert Law (mBLL). A single differential pathlength factor (DPF) of 6 was used for both wavelengths of 695 nm and 830 nm in the analysis based on the value found from previous NIRS studies on human brain (Pierro et al., 2012). We examined the NIRS signal with respect to the baseline for each subject to minimize

the effects of extra-cerebral layer contamination. The source-detector distance (3 cm) was sufficient to ensure that the cerebral cortex was sampled. Additionally, previous NIRS studies have shown that the task-related effects on extra-cerebral layer hemoglobin concentration is negligible (Brigadoi and Cooper, 2015). Traces were segmented into 20-s epochs around the trigger stimulus for each trial with each epoch starting 5 s prior to each stimulus. Baseline correction corresponded to the mean HbO/HbR values from −5 to 0 s. The hemodynamic response function was then generated for each channel during each condition for each participant by averaging the response curves from all trials within a condition into a single hemodynamic curve. For each channel, the maximum change in HbO (increase in chromophore concentration) and HbR (decrease in chromophore concentration) between 5 and 20 s in response to each experimental condition (observation and execution) were computed to be used as the dependent variable in subsequent analyses. Due to a greater signal-to-noise ratio, and consistent with previous fNIRS studies, we only used the HbO signal in the remaining analyses (Yamamoto and Kato, 2002; Rahimpour et al., 2020).

Statistical Analysis. Statistical analysis was performed using SAS (Statistical Analysis Software) 9.4v. For each channel, the maximum change in HbO was first assessed relative to the baseline using paired *t*-tests. Reported *p* values are Bonferroni adjusted for this analysis. Following this initial analysis, a mixed model was computed using channels that showed an increase in HbO hemodynamic activity relative to baseline to examine the contrasts between conditions. Furthermore, paired *t*-test were conducted between the average hemodynamic response of channels in left hemisphere (channels 1–12) and channels on the right hemisphere (channels 12–24), to determine lateralized responses per condition. To test the relation between hemodynamic response and handedness, we computed Pearson correlation coefficients between hemodynamic response ratio and handedness scores. +. Handedness scores included the Edinburgh score (0–50, with 50 indicating the highest preference for right-handed activity; Ransil and Schachter, 1994) as well as the differential dexterity ratio as determined by (Mean number of pegs placed with right hand/ Number of pegs placed with left hand; **Figure 4**). Values above 1 indicate better dexterity with the right hand, while values below one indicates better dexterity with the left hand.

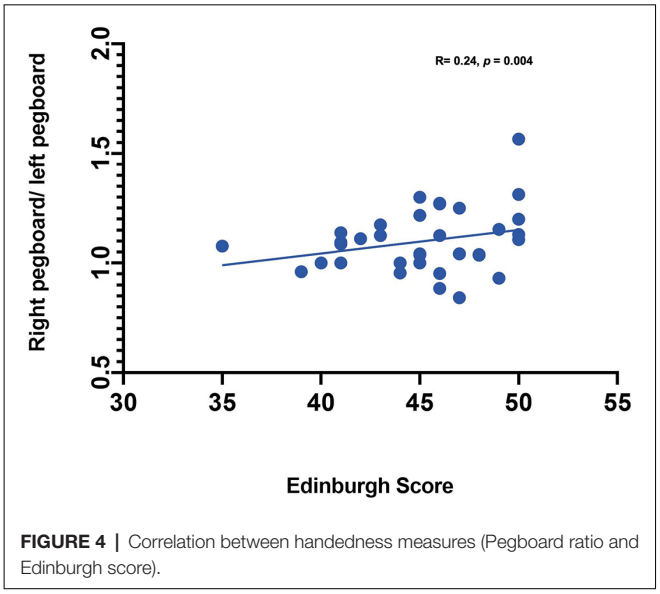

FIGURE 4 | Correlation between handedness measures (Pegboard ratio and Edinburgh score).

RESULTS

Contrasts between conditions. The mixed model examining contrasts between conditions revealed a significant effect for channels 5, $F_{(3,91)} = 2.67$, $p = 0.005$ and 6 $F_{(3,91)} = 2.67$, $p = 0.005$. In both channels right-self resulted in greater hemodynamic response ($M = 0.016$, SE = 0.003 and $M = 0.011$, SE = 0.001) than left-self ($M = 0.01$, SE = 0.003 and $M = 0.005$, SE = 0.0012), right across ($M = 0.01$ SE = 0.025 and $M = 0.006$ SE = 0.001) and left across ($M = 0.013$ SE = 0.003 and $M = 0.006$ SE = 0.001), respectively. Channel 6 hemodynamic response is shown in **Figure 3**. The statistical significance of maximum oxyhemoglobin activation was determined against baseline using a t-test.

Lateralization of brain activity. Paired t-tests revealed a greater hemodynamic response in the left hemisphere for the right-self condition ($t_{(32)} = 2.40$, $p = 0.022$). No significant results were found for left-self, right-across, and left-across.

Effect of handedness. The results from Pearson correlations indicated that there was a significant negative correlation between the L/R ratio for the left-self condition and handedness. Specifically, the correlation was significant when handedness was measured using Purdue pegboard ($rs_{(X)} = -0.484$, $p = 0.0040$; **Figure 5**), and Edinburgh handedness inventory ($rs_{(X)} = -0.388$, $p = 0.019$; **Figure 6**). No significant results were found for right-self, right-across and left-across.

Role of Handedness vs. Hand Used in Brain Lateralization

GLM with hand used predicting lateralization ratio was significant ($t_{(35)} = 2.28$, $p = 0.029$). Specifically, in the self-condition, using the right hand was associated with a higher left, over right ratio than using the left hand. However, in the observer condition, the hand used by the other did not predict brain lateralization ($t_{(35)} = 0.861$, $p = 0.395$).

We also used GLM to determine if hand preference predicts brain lateralization in the self-condition above and beyond what is predicted by whether the participant uses their right or left hand. Hand preference did not predict brain lateralization above and beyond the side used in the self-condition ($t_{(35)} = -1.16$, $p = 0.25$). Differential dexterity also did not predict brain lateralization above and beyond side in self condition ($t_{(35)} = -0.031$, $p = 0.976$). Hand preference also does not predict brain lateralization in the observe condition ($t_{(35)} = -0.691$, $p = 0.494$). However, differential dexterity did predict brain lateralization in the other condition

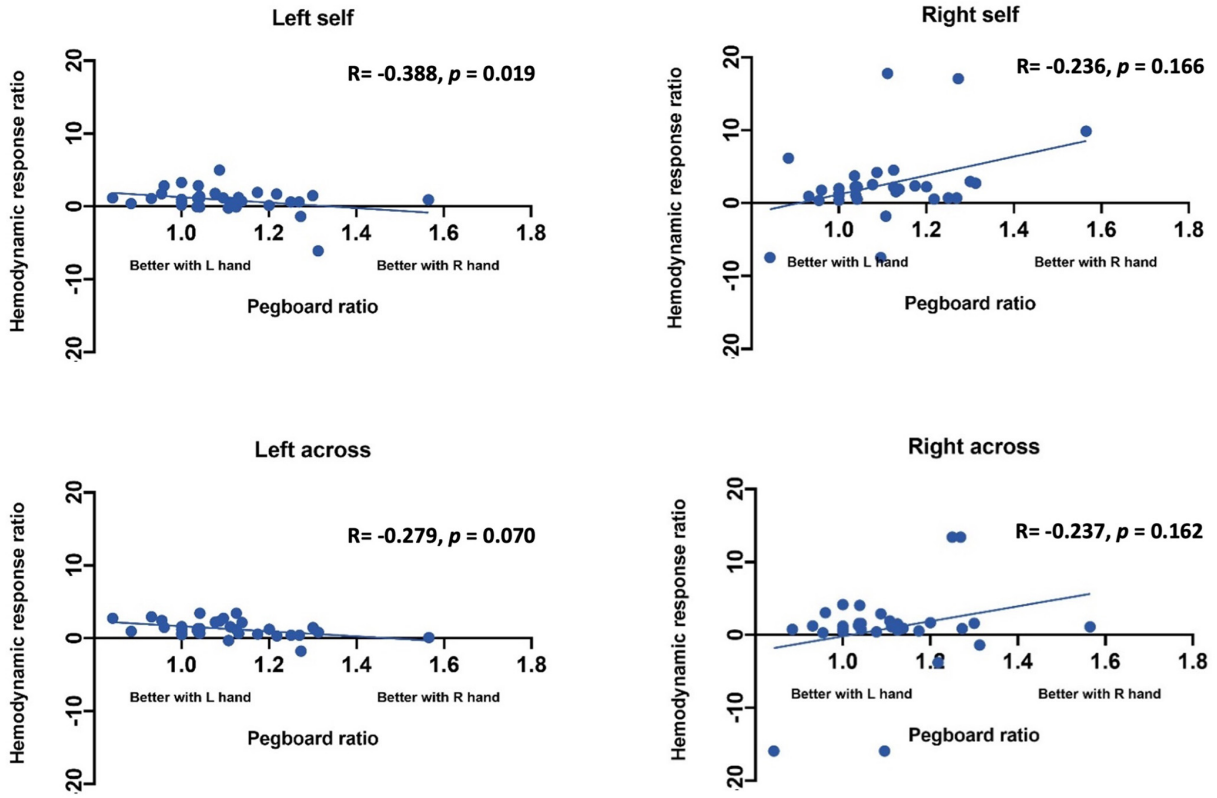

FIGURE 5 | Correlation between ratio of hemodynamic response function and handedness measured by the Purdue Pegboard test.

($t_{(35)} = -2.167$, $p = 0.0375$). Specifically, individuals who were more dexterous with their right hand showed less left lateralized brain activity when watching the actor use both their left and right hands.

DISCUSSION

In this article, we showed that lateralization of activity in the action observation network varies for self-initiated actions and for observed actions. Specifically, using your right vs. left hand led to increased leftward lateralization of brain activity in the self-condition, whereas lateralization of brain activity did not vary when observing the other use their right vs. left hands. We demonstrated that channels placed in the left motor cortex (approximately) are specifically activated by use of the right hand during self-actions (vs. all other conditions). We also show that hand preference and differential dexterity are related to lateralization of the AON across both self and observe conditions. Specifically, we show that while hand used by the person being observed does not drive lateralization of brain activity in the observer, brain lateralization patterns are related to the patterns of hand preference and dexterity. Specifically, preferring the right hand and being better with the right hand were both associated with decreased leftward lateralization when acting with the left hand. Also, having higher dexterity with the right vs. left hand was associated with decreased leftward lateralization during the observe condition independent of which hand the actor was using.

Although mirror vs. anatomical representation have not been directly studied in the AON, multiple studies have shown differential lateralization patterns in the brain for action execution vs. observation. One pattern that has been seen across multiple studies is that when using right-handed subjects, there is greater left hemisphere activation in motor regions regardless of whether the observed action is with the left or right hand (Koski et al., 2002). This is very much in line with our findings, and suggests that once handedness has emerged, brain lateralization patterns associated with handedness are an important determinant of neural mirroring patterns. Extrapolated to mapping others' actions, this supports the possibility that observers activate brain regions "as if" they were doing the task with their dominant hand. It is unclear how these patterns would present in younger children with less lateralized motor behaviors and completing this research in young children would clarify how handedness may drive development of neural mirroring and/or imitation abilities.

Importantly, individuals' handedness is also known to affect brain lateralization. Specifically, the ratio of contralateral to

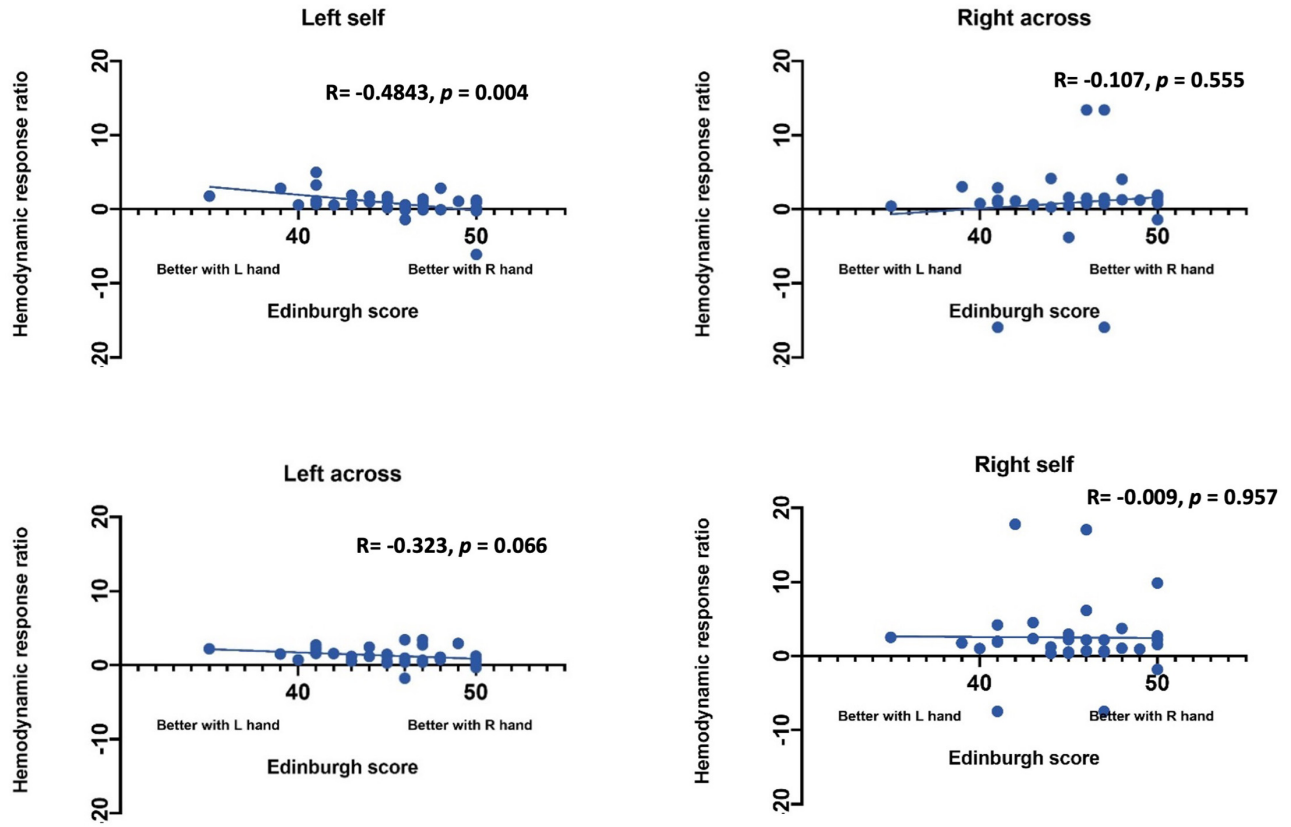

FIGURE 6 | Correlation between ratio of hemodynamic response function and handedness measured by the Edinburgh handedness inventory.

ipsilateral activation when using the dominant hand has shown a linear association with degree of handedness (Dassonville et al., 1998). The present study replicates that pattern and extends it beyond hand preference to differential dexterity as well. In addition, we show that handedness is related not only to brain lateralization for the individual during their own actions but also to brain lateralization when observing others. This provides further support to the idea that activation in the action observation network is tied to an individual's own experiences and expertise (Condy et al., 2021).

While this study is the first to examine brain lateralization across both self and other conditions and with the left and right hand, several limitations should be noted. First, we did not include left handers in this study due to low numbers recruited. While only 10% of humans are left-handed, their brain lateralization in the AON is still meaningful and needs to be studied. For example, we found here that individual handedness was related to AON activity even when observing others, but it is possible that this would not be true in left handers, which could alter the interpretation of these results. Second, while we placed the optode array in relation to the 10–20 international classification system, individuals' heads can vary and therefore we cannot clearly determine exact neuroanatomical locations of hemodynamic activity. Future studies should include digitizing methods to determine the precise location of each channel would allow for determination of activity in primary motor regions vs. somatosensory cortex. In addition, we used a newly designed motor task that was designed to be interesting enough for toddlers and preschoolers to tolerate. Since infant tasks and adult tasks typically rely on very repetitive actions, we included five different tasks that were of sufficient difficulty to be interesting to a toddler or preschooler. While our task clearly elicited lateralized activity during the self-action condition, it is possible that lack of lateralization during observation was task specific, which indicates the need for replication. Finally, the observe condition for the present study involved having the participant sit across from the person that they were observing. However, it is possible that lateralization of brain response (and thus, mirror vs. anatomical activation) would vary if the participant were sitting next to the actor. Therefore, future studies should clarify the role of actor orientation in lateralization of AON activity.

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 Institutional Review Board for the University of Maryland. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the article as follows. Study conception and design: ES, NF, and AG. Data collection: ES. Analysis and interpretation of results: KK, HM, ES, SZ, NF, and AG. Draft manuscript preparation: KK, ES, HM, and SZ. All authors reviewed the results and approved the final version of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The present study was supported by the Intramural Research Program (IRP) of the National Institute of Child Health and Human Development (Grant number: HD008882-13) and the National Institute of Health's Bench-to-Bedside Program. This research was also supported by Award F32MH127869 granted to SZ by the National Institute of Mental Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

ACKNOWLEDGMENTS

We would like to sincerely thank all our participants for taking the time to participate in this study.

REFERENCES

Acharya, S., and Shukla, S. (2012). Mirror neurons: enigma of the metaphysical modular brain. *J. Nat. Sci. Biol. Med.* 3, 118–124. doi: 10.4103/0976-9668.101878

Bhat, A. N., Hoffman, M. D., Susanna, L. T., McKenzie, L. C., Jeffrey, E., Tsuzuki, D., et al. (2017). Cortical activation during action observation, action execution and interpersonal synchrony in adults: a functional near-infrared spectroscopy (fNIRS) study. *Front. Hum. Neurosci.* 11:431. doi: 10.3389/fnhum.2017.00431

Brigadoi, S., and Cooper, R. J. (2015). How short is short? optimum source-detector distance for short-separation channels in functional near-infrared spectroscopy. *Neurophotonics* 2:025005. doi: 10.1117/1.NPh.2.2.025005

Brigadoi, S., Lisa, C., Simone, C., Fabio, S., Pietro, S., Juliette, S., et al. (2014). Motion artifacts in functional near-infrared spectroscopy: a comparison of motion correction techniques applied to real cognitive data. *Neuroimage* 85, 181–191. doi: 10.1016/j.neuroimage.2013.04.082

Buddenberg, L. A., and Davis, C. (2000). Test-retest reliability of the purdue pegboard test. *Am. J. Occup. Ther.* 54, 555–558. doi: 10.5014/ajot.54.5.555

Colebatch, J. G., Deiber, M. P., Passingham, R. E., Friston, K. J., and Frackowiak, R. S. J. (1991). Regional cerebral blood flow during voluntary arm and hand movements in human subjects. *J. Neurophysiol.* 65, 1392–1401. doi: 10.1152/jn.1991.65.6.1392

Condy, E. E., Helga, O. M., John, M., Doug, H., Kosar, K., Nathan, F., et al. (2021). Characterizing the action-observation network through functional near-infrared spectroscopy: a review. *Front. Hum. Neurosci.* 15:627983. doi: 10.3389/fnhum.2021.627983

Cooper, R., Juliette, S., Louis, G., Dorte, P., Henrik, S., Helle, I., et al. (2012). A systematic comparison of motion artifact correction techniques for functional near-infrared spectroscopy. *Front. Neurosci.* 6:147. doi: 10.3389/fnins.2012.00147

Crivelli, D., Rueda, M. D. S., and Balconi, M. (2018). Linguistic and motor representations of everyday complex actions: an fNIRS investigation. *Brain Struct. Funct.* 223, 2989–2997. doi: 10.1007/s00429-018-1646-9

Cross, E. S., David, J. M. K., Antonia, F., Hamilton, D. C., Kelley, W. M., and Grafton, S. T. (2009). Sensitivity of the action observation network to physical and observational learning. *Cereb. Cortex* 19, 315–326. doi: 10.1093/cercor/bhn083

Cutini, S., Moro, S. B., and Bisconti, S. (2012). Review: functional near infrared optical imaging in cognitive neuroscience: an introductory review. *J. Near Inf. Spectrosc.* 20, 75–92. doi: 10.1255/jnirs.969

Dashtestani, H., Zaragoza, R., Pirsiavash, H., Knutson, K. M., Kermanian, R., Cui, J, et al. (2019). Canonical correlation analysis of brain prefrontal activity measured by functional near infra-red spectroscopy (fNIRS) during a moral judgment task. *Behav. Brain Res.* 59, 73–80. doi: 10.1016/j.bbr.2018.10.022

Dassonville, P., Zhu, X.-H., Ugurbil, K., Kim, S.-G., and Ashe, J. (1998). Functional activation in motor cortex reflects the direction and the degree of handedness. *Proc. Nat. Acad. Sci. U S A* 94, 14015–14018.

Franz, E. A., Ford, S., and Werner, S. (2007). Brain and cognitive processes of imitation in bimanual situations: making inferences about mirror neuron systems. *Brain Res.* 1145, 138–149. doi: 10.1016/j.brainres.2007.01.136

Fukuda, M., and Mikuni, M. (2012). Clinical application of near-infrared spectroscopy (NIRS) in psychiatry: the advanced medical technology for differential diagnosis of depressive state. *Seishin Shinkeigaku Zasshi* 114, 801–806.

Gagnon, L., Meryem, A. Y., Dehaes, M., Cooper, R. J., Perdue, K. L., Selb, J., et al. (2012). Quantification of the cortical contribution to the NIRS signal over the motor cortex using concurrent NIRS-FMRI measurements. *Neuroimage* 59, 3933–3940. doi: 10.1016/j.neuroimage.2011.10.054

Hannanu, F. F., Goundous, I., Detante, O., Naegele, B., and Jaillard, A. (2020). Spatiotemporal patterns of sensorimotor FMRI activity influence hand motor recovery in subacute stroke: a longitudinal task-related FMRI study. *Cortex* 129, 80–98. doi: 10.1016/j.cortex.2020.03.024

Herold, F., Wiegel, F., Scholkmann, F., Thiers, A., Hamacher, D., and Schega, L. (2017). Functional near-infrared spectroscopy in movement science: a systematic review on cortical activity in postural and walking tasks. *Neurophotonics* 4:041403. doi: 10.1117/1.NPh.4.4.041403

Holper, L., Muehlemann, T., Scholkmann, F., Eng, K., Kiper, D., and Wolf, M. (2010). testing the potential of a virtual reality neurorehabilitation system during performance of observation, imagery and imitation of motor actions recorded by wireless functional near-infrared spectroscopy (fNIRS). *J. Neuroeng. Rehabil.* 7:57. doi: 10.1186/1743-0003-7-57

Huppert, T., Diamond, S. G., Franceschini, M. A., and Boas, D. A. (2009). HomER: a review of time-series analysis methods for nearinfrared spectroscopy of the brain. *Appl. Opt.* 48, D280–D298. doi: 10.1364/ao.48.00d280

Jin, S. H., Lee, S. H., Yang, S. T., and An, J. (2020). Hemispheric asymmetry in hand preference of right-handers for passive vibrotactile perception: an fNIRS study. *Sci. Rep.* 10, 1–10. doi: 10.1038/s41598-020-70496-y

Kajiume, A., Aoyama-Setoyama, S., Saito-Hori, Y., Ishikawa, N., and Kobayashi, M. (2013). Reduced brain activation during imitation and observation of others in children with pervasive developmental disorder: a pilot study. *Behav. Brain Funct.* 9:21. doi: 10.1186/1744-9081-9-21

Kilner, J. M., and Lemon, R. N. (2013). What we know currently about mirror neurons. *Curr. Biol.* 23, R1057–R1062. doi: 10.1016/j.cub.2013.10.051

Koski, L., Wohlschläger, A., Bekkering, H., Woods, P., Dubeau, M.-C., Mazziotta, J. C., et al. (2002). koski-2002-modulation of motor and premotor activity duringimitation af target directed actions. *Cereb. Cortex* 12, 847–855. doi: 10.1093/cercor/12.8.847

Lepage, J. F., and Théoret, H. (2006). EEG evidence for the presence of an action observation-execution matching system in children. *Eur. J. Neurosci.* 23, 2505–2510. doi: 10.1111/j.1460-9568.2006.04769.x

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

Pierro, M. L., Sassaroli, A., Bergethon, P. R., Ehrenberg, B. L., and Fantini, S. (2012). Phase-amplitude investigation of spontaneous low-frequency oscillations of cerebral hemodynamics with near-infrared spectroscopy: a sleep study in human subjects. *NeuroImage* 63, 1571–1584. doi: 10.1016/j.neuroimage.2012.07.015

Pinti, P., Aichelburg, C., Gilbert, S., Hamilton, A., Hirsch, J., Burgess, P., et al. (2018). A review on the use of wearable functional near-infrared spectroscopy in naturalistic environments. *Jpn. Psychol. Res.* 60, 347–373. doi: 10.1111/jpr.12206

Pulvermüller, F., Lutzenberger, W., Preißl, H., and Birbaumer, N. (1995). Motor programming in both hemispheres: an EEG study of the human brain. *Neurosci. Lett.* 190, 5–8. doi: 10.1016/0304-3940(95)11486-g

Rahimpour, A., Pollonini, L., Comstock, D., Balasubramaniam, R., and Bortfeld, H. (2020). Tracking differential activation of primary and supplementary motor cortex across timing tasks: an fNIRS validation study. *J. Neurosci. Methods* 341:108790. doi: 10.1016/j.jneumeth.2020.108790

Ransil, B. J., and Schachter, S. C. (1994). Test-retest reliability of the edinburgh handedness inventory and global handedness preference measurements and their correlation. *Percept. Mot. Skills* 79, 1355–1372. doi: 10.2466/pms.1994.79.3.1355

Rao, S. M., Binder, J. R., Bandettini, P. A., Hammeke, T. A., Yetkin, F. Z., Jesmanowicz, A., et al. (1993). Functional magnetic resonance imaging of complex human movements. *Neurology* 43, 2311–2318. doi: 10.1212/wnl.43.11.2311

Sassaroli, A., Frederick, B. D. B., Tong, Y., Renshaw, P. F., and Fantini, S. (2006). Spatially weighted BOLD signal for comparison of functional magnetic resonance imaging and near-infrared imaging of the brain. *Neuroimage* 33, 505–514. doi: 10.1016/j.neuroimage.2006.07.006

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

Tiffin, J., and Asher, E. J. (1948). The purdue pegboard: norms and studies of reliability and validity. *J. Appl. Psychol.* 32, 234–247. doi: 10.1037/h0061266

Wilcox, T., and Biondi, M. (2015). fNIRS in the developmental sciences. *Wiley Interdiscip. Rev. Cogn. Sci.* 6, 263–283. doi: 10.1002/wcs.1343

Yamamoto, T., and Kato, T. (2002). Paradoxical correlation between signal in functional magnetic resonance imaging and deoxygenated haemoglobin content in capillaries: a new theoretical explanation. *Phys. Med. Biol.* 47, 1121–1141. doi: 10.1088/0031-9155/47/7/309

Yu, Z., Gonciarz, M. D., Sundquist, W. I., Hill, C. P., and Jensen, J. (2006). A temporal comparison of BOLD, ASL and NIRS hemodynamic responses to motor stimuli in adult humans. *Neuroimage* 377, 364–377. doi: 10.1016/j.neuroimage.2005.08.065

Zhang, Q., Zhang, P., Song, L., Yang, Y., Yuan, S., Chen, Y., et al. (2019). Brain activation of elite race walkers in action observation, motor imagery and motor execution tasks: a pilot study. *Front. Hum. Neurosci.* 13, 1–13. doi: 10.3389/fnhum.2019.00080

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.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership