

GRAPHONOMICS AND YOUR BRAIN ON ART, CREATIVITY, AND INNOVATION

EDITED BY: Jose Luis Contreras-Vidal, Jose M. Azorin, Réjean Plamondon,
Claudio De Stefano and Surjo R. Soekadar
PUBLISHED IN: Frontiers in Human Neuroscience





frontiers

Frontiers eBook Copyright Statement

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

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

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence.

The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

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

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

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

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

ISSN 1664-8714

ISBN 978-2-88974-547-0

DOI 10.3389/978-2-88974-547-0

About Frontiers

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

Frontiers Journal Series

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

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

GRAPHONOMICS AND YOUR BRAIN ON ART, CREATIVITY, AND INNOVATION

Topic Editors:

Jose Luis Contreras-Vidal, University of Houston, United States

Jose M. Azorin, Miguel Hernández University of Elche, Spain

Réjean Plamondon, Polytechnique Montréal, Canada

Claudio De Stefano, University of Cassino, Italy

Surjo R. Soekadar, Charité Universitätsmedizin Berlin, Germany

Citation: Contreras-Vidal, J. L., Azorin, J. M., Plamondon, R., De Stefano, C., Soekadar, S. R., eds. (2022). Graphonomics and Your Brain on Art, Creativity, and Innovation. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88974-547-0

Table of Contents

- 04 Central and Peripheral Shoulder Fatigue Pre-screening Using the Sigma–Lognormal Model: A Proof of Concept**
Anaïs Laurent, Réjean Plamondon and Mickael Begon
- 20 A Theoretical Framework for How We Learn Aesthetic Values**
Hassan Aleem, Ivan Correa-Herran and Norberto M. Grzywacz
- 38 The Relationships Between Trait Creativity and Resting-State EEG Microstates Were Modulated by Self-Esteem**
Xin Wu, Jiajia Guo, Yufeng Wang, Feng Zou, Peifang Guo, Jieyu Lv and Meng Zhang
- 50 The Attentive Cursor Dataset**
Luis A. Leiva and Ioannis Arapakis
- 57 Think Hard or Think Smart: Network Reconfigurations After Divergent Thinking Associate With Creativity Performance**
Hong-Yi Wu, Bo-Cheng Kuo, Chih-Mao Huang, Pei-Jung Tsai, Ai-Ling Hsu, Li-Ming Hsu, Chi-Yun Liu, Jyh-Horng Chen and Changwei W. Wu
- 68 Characterization of the Stages of Creative Writing With Mobile EEG Using Generalized Partial Directed Coherence**
Jesus G. Cruz-Garza, Akshay Sujatha Ravindran, Anastasiya E. Kopteva, Cristina Rivera Garza and Jose L. Contreras-Vidal
- 76 Feasibility and Safety of Bilateral Hybrid EEG/EOG Brain/Neural–Machine Interaction**
Marius Nann, Niels Peekhaus, Cornelius Angerhöfer and Surjo R. Soekadar
- 89 Using Posterior EEG Theta Band to Assess the Effects of Architectural Designs on Landmark Recognition in an Urban Setting**
James D. Rounds, Jesus Gabriel Cruz-Garza and Saleh Kalantari
- 104 What to Expect When the Unexpected Becomes Expected: Harmonic Surprise and Preference Over Time in Popular Music**
Scott A. Miles, David S. Rosen, Shaun Barry, David Grunberg and Norberto Grzywacz
- 114 Stochasticity, Nonlinear Value Functions, and Update Rules in Learning Aesthetic Biases**
Norberto M. Grzywacz
- 141 Imagining How Lines Were Drawn: The Appreciation of Calligraphy and the Facilitative Factor Based on the Viewer’s Rating and Heart Rate**
Kazuki Matsumoto and Takeshi Okada



Central and Peripheral Shoulder Fatigue Pre-screening Using the Sigma–Lognormal Model: A Proof of Concept

Anaïs Laurent¹, Réjean Plamondon^{2*} and Mickael Begon^{3,4}

¹ Laboratoire Scribens, Département de Génie Électrique, Programme de Génie Biomédical, Polytechnique Montréal, Montreal, QC, Canada, ² Laboratoire Scribens, Département de Génie Électrique, Polytechnique Montréal, Montreal, QC, Canada, ³ Laboratoire de Simulation et de Modélisation du Mouvement, School of Kinesiology and Physical Activity Sciences, Faculty of Medicine, Université de Montréal, Montreal, QC, Canada, ⁴ CHU Sainte-Justine, Montreal, QC, Canada

OPEN ACCESS

Edited by:

Giovanni Di Pino,
Campus Bio-Medico University, Italy

Reviewed by:

Jiri Mekyska,
Brno University of Technology,
Czechia
Antonio Parziale,
University of Salerno, Italy

*Correspondence:

Réjean Plamondon
rejean.plamondon@polymtl.ca

Specialty section:

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

Received: 15 February 2020

Accepted: 20 April 2020

Published: 19 May 2020

Citation:

Laurent A, Plamondon R and
Begon M (2020) Central
and Peripheral Shoulder Fatigue
Pre-screening Using
the Sigma–Lognormal Model: A Proof
of Concept.
Front. Hum. Neurosci. 14:171.
doi: 10.3389/fnhum.2020.00171

Background: Clinical tests for detecting central and peripheral shoulder fatigue are limited. The discrimination of these two types of fatigue is necessary to better adapt recovery intervention. The Kinematic Theory of Rapid Human Movements describes the neuromotor impulse response using lognormal functions and has many applications in pathology detection. The ideal motor control is modeled and a change in the neuromuscular system is reflected in parameters extracted according to this theory.

Objective: The objective of this study was to assess whether a shoulder neuromuscular fatigue could be detected through parameters describing the theory, if there is the possibility to discriminate central from peripheral fatigue, and which handwriting test gives the most relevant information on fatigue.

Methods: Twenty healthy participants performed two sessions of fast stroke handwriting on a tablet, before and after a shoulder fatigue. The fatigue was in internal rotation for one session and in external rotation during the other session. The drawings consisted of simple strokes, triangles, horizontal, and vertical oscillations. Parameters of these strokes were extracted according to the Sigma–Lognormal model of the Kinematic Theory. The evolution of each participant was analyzed through a *U*-Mann–Whitney test for individual comparisons. A Hotelling's T^2 -test and a *U*-Mann–Whitney test were also performed on all participants to assess the group evolution after fatigue. Moreover, a correlation among parameters was calculated through Spearman coefficients to assess intrinsic parameters properties of each handwriting test.

Results: Central and peripheral parameters were statistically different before and after fatigue with a possibility to discriminate them. Participants had various responses to fatigue. However, when considering the group, parameters related to the motor program execution showed significant increase in the handwriting tests after shoulder fatigue. The test of simple strokes permits to know more specifically where the fatigue comes from, whereas the oscillations tests were the most sensitive to fatigue.

Conclusion: The results of this study suggest that the Sigma–Lognormal model of the Kinematic Theory is an innovative approach for fatigue detection with discrimination between the central and peripheral systems. Overall, there is a possibility to implement the setting for clinics and sports personalized follow-up.

Keywords: Sigma–Lognormal model, Kinematic Theory of rapid human movement, central fatigue, peripheral fatigue, rotator cuff, handwriting, shoulder

INTRODUCTION

One hundred million workers in the European population suffer from chronic musculoskeletal disorders and pain (Bevan, 2015). Direct and indirect costs for treating them are expensive, as they accounted, respectively, for up to \$796.3 billion (which represents 5.2% of the national gross domestic product) and \$130.7 billion in the US population per year between 2009 and 2011 (U.S. Bone and Joint Initiative, 2014). Shoulder is considered to be one of the most affected joints, as it represents the third cause of clinical consultation after the lumbar and cervical regions. Disorders at the rotator cuff in the shoulder region represents 50–85% of all shoulder musculoskeletal diseases in Québec (Roy et al., 2015). Overhead and arm elevation repetition movement is an important risk factor (Hagberg and Wegman, 1987; Svendsen et al., 2004; Ebaugh et al., 2006). In fact, while performing these movements, neuromuscular fatigue generates muscular and kinematic adaptations (Ebaugh et al., 2006; Gaudet et al., 2018), which can lead to musculoskeletal disorders (Beach et al., 1992). Sports requiring this kind of motion are then more affecting its players, like in volleyball, baseball, tennis, etc. (Wang and Cochrane, 2001; Mullaney et al., 2005; Wilk et al., 2009; Joshi et al., 2011). Detecting shoulder fatigue at an early stage could be a meaningful approach to avoid shoulder injuries.

Neuromuscular fatigue corresponds to “any exercise-induced loss of ability to produce force with a muscle or muscle group” (Taylor et al., 2006). It can be decomposed into two categories: central fatigue (Gandevia, 2001) and peripheral fatigue (Enoka and Stuart, 1992). Central fatigue implies the neural system: the voluntary activation and information conduction for movement execution are dysfunctional (Sesboüé and Guinestre, 2006; Boyas and Guével, 2011). Central fatigue can come from the supraspinal and spinal areas. Peripheral fatigue involves the muscles: in that case, the muscular excitation is impaired. It can cause for example a deterioration in the action potentials propagation or in the excitation–contraction coupling responsible for contraction (Sesboüé and Guinestre, 2006). A poor metabolite substrates supply can also be a consequence of peripheral fatigue, implying also an alteration of the excitation–contraction coupling (Boyas and Guével, 2011). The output force is reduced and the contractile mechanisms are dysfunctional (Bigland-Ritchie and Woods, 1984). In all cases, fatigue is different depending on the task (duration and weight lifted) and on the type of contraction (Chaffin et al., 2006a).

Several methods for detecting fatigue already exist. Numerous scales have been developed which are fatigue and task specific

(Dittner et al., 2004). For example, the Visual Analog Scale is a reliable scale used to analyze a global fatigue and has already been used for muscular fatigue (Lee et al., 1991). However, this method is not so accurate for low intensities contractions (Leung et al., 2004). The Perceived Exertion Force is commonly used in fatigue studies, with the Borg’s scale (Borg, 1998) and in comparison to other scales, it seems to be one of the most accurate (Neely et al., 1992). However, results from this scale have to be analyzed carefully as it remains subjective (Chen et al., 2002). Objective approaches for fatigue detection also exist. One of the most frequently used is the electromyography (EMG) which was first employed by Piper (1912), according to Cifrek et al. (2009). Parameters such as the amplitude of the root mean square of the EMG signal increase with fatigue, as more motor units are recruited for the same amount of force produced (Merletti et al., 2004). The mean or median frequency of the power spectrum density decreases, as the velocity of action potentials is slowed down (Edwards and Lippold, 1956; Lindström et al., 1977; Al-Mulla et al., 2012). A complication with EMG is the quality of the signal to be assessed. It is essential to have good anatomical knowledge for electrodes placement, in order to avoid crosstalk problems as much as possible, which can lead to misinterpretations in the results analysis (Hermens et al., 2000; Merletti et al., 2001; Farina et al., 2004). EMG presents some difficulties for clinical evaluation as the electrode placement and signal treatment is time-consuming. There is also the possibility of using biomarkers as for example lactate concentrations (Tesch et al., 1978; Finsterer, 2012). Its intracellular concentration is supposed to diminish with the apparition of fatigue. Nevertheless, even if they are accurate methods, they remain invasive and hard to implement easily. Other non-invasive methods are employed for peripheral fatigue detection, such as sonomyography, near-infrared spectroscopy, mechanomyography, or acoustic myography (Mancini et al., 1994; Huang et al., 2007; Shi et al., 2007; Al-Mulla et al., 2011; Ibitoye et al., 2014). However, most of these techniques have to be synchronized with EMG to detect muscle fatigue and they cannot assess central fatigue. On the other hand, central fatigue can be evaluated either with percutaneous nerve stimulation –with an electrical nerve stimulation- or transcranial magnetic stimulation –with a nerve cells magnetic stimulation- (Gandevia, 2001; Taylor and Gandevia, 2001; Rozand et al., 2015) during maximal contractions. If the stimulation evokes an extra-force, it suggests that central fatigue is present (Merton, 1954). One more time, EMG can complement the methods to detect central and peripheral fatigue. Moreover, transcranial magnetic stimulation requires a magnetic coil to stimulate the motor cortex, which can interfere with EMG recordings (Valero-Cabré

et al., 2011). In complement, percutaneous nerve stimulation is to our knowledge, so far not applicable to all muscles and requires an experimental set up lengthy and difficult to implement (Palmieri et al., 2004). This increases the risk of experimental misinterpretations, thus the difficulty to transpose it to a clinic. It is then necessary to find a method for detecting central and peripheral shoulder fatigue, which would be usable in clinics on a daily basis.

It has been shown that the Kinematic Theory of rapid human movements describes accurately the neuromotor control (Plamondon, 1995a,b). This theory is based on the analysis of the velocity profiles of the end effector of a movement, like the finger, the wrist, the arm, the shoulder, the head, the trunk, the eye movements, etc. These movements can be modeled using lognormal functions, which depict the impulse response of the neuromuscular system of a participant (Plamondon et al., 2008). Thus, both central and peripheral information can theoretically be extracted from the movement reconstruction (Plamondon et al., 2003). In comparison, the Minimum-Jerk model (Hogan, 1984; Flash and Hogan, 1985) postulates that end effector trajectories are chosen by the central nervous system (CNS) such that the time integral of the squared magnitude of hand jerk is minimal, which is equivalent to maximizing the smoothness of the trajectory. Both approaches describe the same bell-shaped velocity using different analytical equations. Working with the Minimum-Jerk model does not give access to the command profile sent by the CNS. It is assumed that the alpha motoneuron signals, at the muscle level, correspond to a movement trajectory. This representation does not take into account the instant when the movement command is sent to the end-effector, neither the time required by the CNS to build and send the appropriate signals to the motor cortex neurons or the moment when the muscle starts to contract. The Minimum-Jerk does not give access to the central and peripheral information that we are investigating in this paper (Djioua and Plamondon, 2010). For these reasons, the Kinematic Theory was preferred to the Minimum-Jerk model for movement reconstruction. Moreover, the Kinematic Theory has been used in several pathologies for motor control studies, such as attention deficit hyperactivity disorder (Laniel et al., 2019), Parkinson's disease (Lebel et al., 2017, 2018a,b; Nadeau et al., 2018), stroke risk factors (O'Reilly and Plamondon, 2011), concussion (Faci et al., 2020b), and it requires a non-invasive, low cost and plug-in-play experimental set-up made up of a digitizing tablet connected to a laptop. Its most recent implementation is ergonomic and very easy to use (Faci et al., 2018). Since during a shoulder fatigue the kinematic parameters and the fine motor control are modified (Qin et al., 2014), we hypothesized that the Kinematic Theory of rapid human movements may be relevant to monitor and assess shoulder fatigue analysis through graphomotricity. The objective of this work is to report a feasibility study aiming at the objective detection of muscular fatigue and the discrimination of central and peripheral fatigue, in an economical and non-invasive way, with the Kinematic Theory of rapid human movements.

MATERIALS AND METHODS

Participants

Eleven males and nine females took part in the experiment. They were all healthy active adults (age: 23.2 ± 3.2 years, height: 173 ± 8.3 cm, mass: 71.7 ± 10.0 kg, 18 right-handed and 2 left-handed). All participants were free of any upper-limb musculoskeletal disorder and had no history of shoulder surgery or neurological disease in the past. The study was approved by the Research Ethics Committee of Polytechnique Montréal (CER-1819-23 v.3).

Experimental Part

Participants completed two sessions in which they performed four series of fast strokes on a tablet before and after a task of shoulder fatigue. The two sessions were similarly performed, at the exception that the fatigue task targeted the shoulder external or internal rotators (sessions 1 and 2 randomly). There were at least 3 days of rest between the two sessions to avoid the participants to be still fatigued at the beginning of the second session. The process for each session was the following (**Figure 1C**): participants first had to execute the four series of fast strokes and then to alternate between a task of fatigue and a series of fast strokes. The series consisted of drawing simple strokes, triangles, horizontal oscillations and vertical oscillations in a random sequence.

The trajectory of fast strokes was recorded on a Wacom Cintiq 13HD tablet (Faci et al., 2018, 2020a). The tablet was positioned such that the participant's fingertip touched the bottom of the tablet when the shoulder was 90° flexed. Participants had to position the stylus on the starting point of the tablet (**Figure 1A**). They started their movement as fast as possible at an audible stimulus ("bip" at 1 kHz for 500 ms) which was emitted after a random and unpredictable delay between 1 and 10 s. Depending on the task, a different guide-screen was displayed to help participants get the right movement [see further details regarding the protocol in O'Reilly et al. (2014)].

- (A) Simple strokes: participants were asked to draw 30 simple strokes from a starting point to a broad finish area. At the end of each stroke, participants had to maintain the stylus on the finish area for at least 1 s. A training period of 5–10 strokes was carried out before the recording.
- (B) Triangles: 30 triangles had to be drawn, passing through 3 points in the same clockwise or anticlockwise direction -chosen by the participant- and they had to wait with their stylus on the tablet for at least 1 s at the end. A training period was also carried out before the recording.
- (C) Horizontal and (D) vertical oscillations: 10 s of oscillations at maximal speed between two parts spaced 50 mm apart were performed between two audible stimuli. After the second signal, stylus kinematics was still recorded until the participant completely stopped and maintained the stylus on the tablet for at least 1 s. Only one trial was registered in these cases, without any training period to avoid fatiguing participants with these two maximal speed tests.

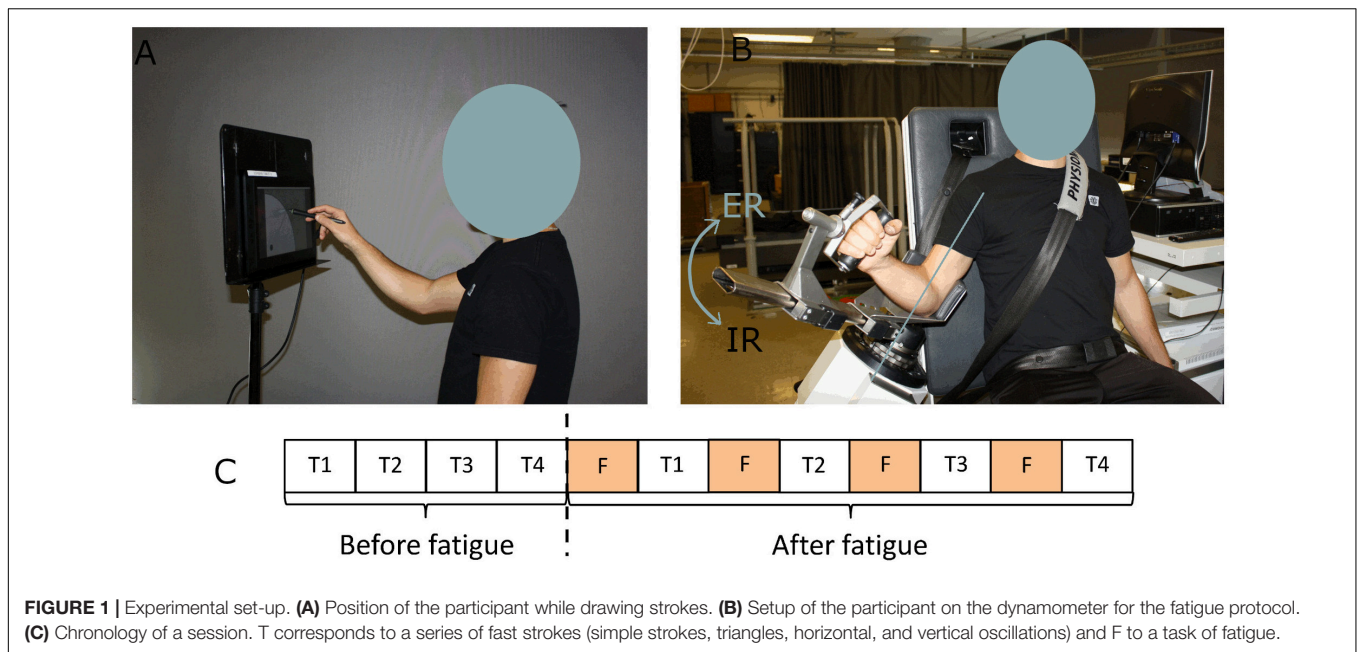


FIGURE 1 | Experimental set-up. **(A)** Position of the participant while drawing strokes. **(B)** Setup of the participant on the dynamometer for the fatigue protocol. **(C)** Chronology of a session. T corresponds to a series of fast strokes (simple strokes, triangles, horizontal, and vertical oscillations) and F to a task of fatigue.

The fatigue task consisted in repetitive submaximal dynamic contractions (concentric – continuous passive mode) at $90^\circ/\text{s}$ (70° of amplitude) in internal or external rotation on an isokinetic dynamometer (CON-TREX[®] MJ; Schnaittach, Germany). Participants were securely fastened using a belt so as not to move their back. Their arm was positioned at 30° of elevation (**Figure 1B**). A training period was allocated to familiarize the participant with the isokinetic effort and to warmup. To determine a target zone of $50 \pm 7.5\%$ of their maximum voluntary contraction, participants performed first a maximum voluntary isokinetic contraction in external rotation (or internal rotation during the other session). At each external rotation (or internal rotation) the participant was instructed to reach this target zone and to rest during the internal rotation (or external rotation). The Borg CR10 Scale (Borg, 1998) which varies from 0 (no effort at all) to 10 (the hardest exercise ever made) was asked every minute, to monitor perceived exertion. Stopping criteria of the fatigue trials were similar to those defined in Yang et al. (2018): (i) Borg number reached 9/10, (ii) three consecutive fails in reaching the target zone, (iii) after 30, 20, 15, and 10 min for the first, second, third, and fourth exercise of fatigue, respectively. The participants were not aware of these criteria. Verbal encouragements were provided as soon as the performance was outside the target zone. The number of Borg has been recorded for 19 participants in external rotation and 18 participants in internal rotation.

Sigma-Lognormal Model

Data captured using the tablet were modeled according to the Kinematic Theory paradigm (Plamondon, 1995a,b). This theory describes the velocity profile of an end effector as the synergetic impulse response of neuromuscular systems. Each of these systems is made of an infinite of subsystems, which are linked with a proportionality relationship between their cumulative time

delays. From this postulate it is then predicted, according to the Central Limit Theorem (Plamondon et al., 2003) that the impulse response of a neuromuscular system tends toward a lognormal shape.

$$\vec{v}_i(t - t_0) = \vec{D}_i \Lambda_i(t; t_{0i}, \mu_i, \sigma_i^2) \quad (1)$$

where i represents one lognormal, shifted with a time t_0 with a command amplitude D ; μ and σ representing timing properties of each lognormal such that:

$$\Lambda_i(t; t_{0i}, \mu_i, \sigma_i^2) = \frac{1}{\sigma_i \sqrt{2\pi}(t - t_{0i})} \exp \left\{ -\frac{1}{2\sigma_i^2} [\ln(t - t_{0i}) - \mu_i]^2 \right\} \quad (2)$$

In the case of a simple pointing task, the movement is seen as a synergy of two neuromuscular systems: an *agonist* and an *antagonist*. The *agonist* one is made up of muscles generating the desired action, whereas the *antagonist* system is made up of muscles working in the opposite direction of the desired movement. To that extent, *agonist* and *antagonist* lognormals can be distinguished based on the starting angle θ_{si} (see Equations 4 and 5). If the starting angle of the lognormal points toward the movement direction, the lognormal is *agonist*. If it points toward the opposite direction, the lognormal is *antagonist*. In that case, the resulting velocity can be expressed as the velocity of the *agonist* minus the *antagonist* lognormals. For more complex planar movements, the velocity can be described using a vector summation of lognormals. In that case, trajectories to reconstruct the movement are circle arcs which connect virtual targets defining an action plan. This means that the number of lognormals describing a movement corresponds to the number of virtual targets representing its trajectory (Plamondon and Djioua,

2006; O'Reilly and Plamondon, 2009).

$$\vec{v}(t) = \sum_{i=1}^N \vec{v}_i(t; t_{0i}, \mu_i, \sigma_i^2) \quad (3)$$

$$= \sum_{i=1}^N D_i \begin{bmatrix} \cos(\theta_i(t)) \\ \sin(\theta_i(t)) \end{bmatrix} \Lambda_i(t; t_{0i}, \mu_i, \sigma_i^2); N \geq 2 \quad (4)$$

$$\theta_i(t) = \theta_{si} + \frac{(\theta_{ei} - \theta_{si})}{2} \left[1 + \operatorname{erf} \left(\frac{\ln(t - t_{0i}) - \mu_i}{\sigma_i \sqrt{2}} \right) \right] \quad (5)$$

These lognormal profiles have been observed and confirmed time and again in the last 15 years [see Plamondon (2020) for an extended survey (O'Reilly et al., 2013; Plamondon et al., 2013a)], which led to postulating and formalizing the guiding principle subtending the present research program: the Lognormality Principle (Plamondon et al., 2013b; Plamondon, 2020). According to this paradigm, the emergent lognormality of the neuromuscular impulse response of a given human motor system is a basic global feature reflecting the behavior of individuals who are in perfect control of their movements. The production of complex movements is accomplished by time superimposing and, summing up lognormal vectors, with the goal of minimizing their number in a given task, to produce efficient and fluent gestures and optimize the energy required to generate them. In this context, it is expected that neuromuscular fatigue will affect the lognormal parameters extracted from reconstructing a given set of gestures produced by a subject.

The main parameters describing a lognormal were extracted using an in-house software referred to as Script Studio (O'Reilly and Plamondon, 2009) and were splitted into four categories, which are resumed in the **Supplementary Table S1** (Plamondon, 1995b; Plamondon et al., 2003). Five parameters are regulated from the input level and they describe the central system command: (i) the time that takes the brain to perceive the stimulus and emit the command to the musculoskeletal system: t_0 (s). It has to be differentiated to the stimulus onset, which is $T = 0$ s (O'Reilly et al., 2013) and the reaction time (RT) measured by the instant of movement onset. In other words, t_0 refers to the moment when a population of neurons sends a motor command, it occurs after the audible stimulus is perceived and the motor command is prepared; (ii) $\Delta(t_0)$ (s), which reflects the rhythmicity of an input command. It represents the time elapsed between two successive t_0 and is used in the oscillations only; (iii) the amplitude of the lognormal command: D (mm), which corresponds to the distance covered by the resulting lognormal; (iv) the starting and (v) ending angles of the lognormal: θ_s and θ_e (rad). They describe the action plan made up of the lognormals.

Two parameters describe the timing properties of the neuromuscular system, in other terms the peripheral system of a participant: (vi) the long-time delay or the time taken to reach half of the distance movement on a logarithmic scale: μ [$\ln(s)$]. It corresponds to the rapidity of a reaction to a command by a system; (vii) the log response time or the time taken from the neuromuscular system to respond to a command

on a logarithmic scale: σ [$\ln(s)$]. It is also linked to the movement duration and is a measure of the asymmetry of the lognormal.

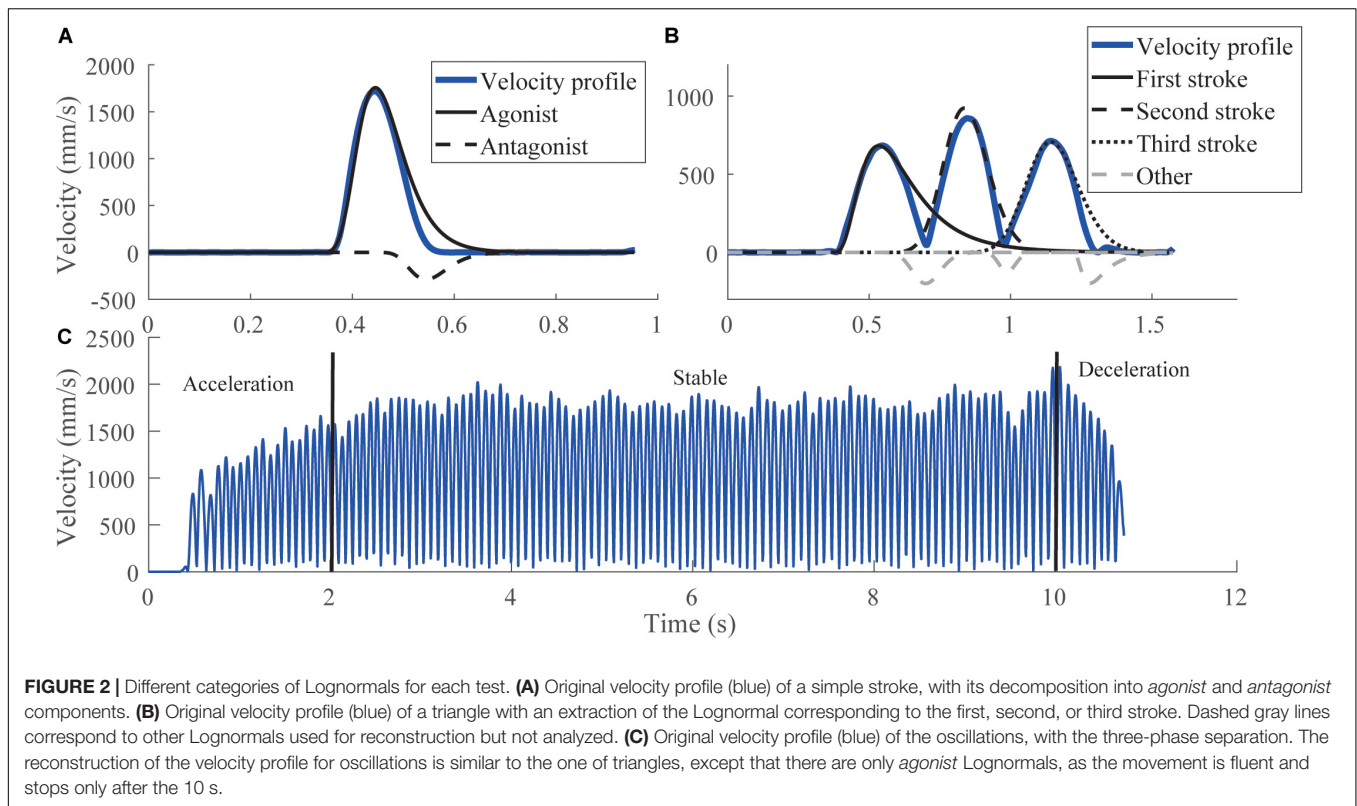
The last two main parameters describe the global state of the neuromotor system: (viii) the number of lognormals required to reconstruct the velocity profile of the movement: $Nb\log$; and (ix) the measure of the quality of the movement reconstruction, Signal-to-Noise Ratio: SNR (dB). They are completed with one derived parameter, (x) the $SNR/Nb\log$ (dB), that is used as a performance criterion and represents the motor control fluency of a gesture. The lognormality principle predicts that the ideal movement converges toward a lognormal profile. When the $SNR/Nb\log$ increases, the movement is closer to the ideal one, as postulated by the lognormal behavior (Plamondon et al., 2013b).

For our study, five derived parameters were also calculated for each type of strokes, representing the motor program execution (see equations in **Supplementary Table S1**). They give information about the velocity at which someone will react or execute a command, and the quality of its response: (xi) the *mode* (s), that is the time at which the maximum value of the lognormal impulse response is reached; (xii) the *median* (s), that is the time at which the half value of the integral under the lognormal curve (50% of the covered distance) is reached; (xiii) the *time delay* (s) which represents the rapidity of a neuromuscular system to respond to a command; (xiv) the *response time* (s) which is a measure of the spread of the impulse response; (xv) the *asymmetry* which characterizes the shape of the lognormal.

A last parameter, not from the theory, was also extracted: (xvi) the *reaction time* (RT) (s) that is the time needed to start the movement after a stimulus. In the present study, it was computed as the time required to reach 10% of the maximal velocity during the test. From this parameter, we calculated (xvii) $RT - t_0$ (s) which is the duration of the command propagation (Woch and Plamondon, 2001).

Data Formatting

The lognormals extracted from each test were split into components as follows. For the simple strokes, two lognormals that defined the largest *agonist* and largest *antagonist* components were analyzed (**Figure 2A**). Strokes composed by only one lognormal were classified as *agonists* (Laurent et al., 2019). For the triangles, strokes were decomposed into the three largest lognormals explaining stroke 1, stroke 2, and stroke 3 (**Figure 2B**). It was manually checked that triangles were properly reconstructed. Those whose lognormals did not describe their correct trajectory were rejected. It is noticed when the starting angle of the reconstruction did not point toward the stroke direction. For oscillations, strokes were split into three phases (**Figure 2C**): acceleration (0–2 s), stable (2–10 s), and deceleration (10 s and more) phases. Lognormals whose amplitude was lower than 50 mm were considered as artifacts and rejected. The remaining lognormals were then classified according to the gesture performed. For the horizontal oscillations (vertical oscillations in the other case), if the cosine (sine in the other case) of the starting angle was positive, the lognormals were considered as an external rotation movement, otherwise they



were considered as an internal rotation movement. For each type of strokes, and each participant, lognormals having at least one parameter outside the mean $\pm 3SD$ were rejected. The proportion of lognormals retained by tests is reported in Section “Results.”

Statistical Analyses

To assess the evolution of each participant after fatigue, individual comparisons using a paired *U*-Mann–Whitney test (non-parametric paired *t*-test) were completed. This test was performed using all the Lognormals of the 30 strokes and depending on the type of fatigue (ER or IR). For the simple strokes and the triangles, 16 parameters were compared, the statistical significance level was thus set at $p < 0.00031$ (i.e., 0.05/16) after Bonferroni correction. For the oscillations, *SNR*, *SNR/Nblog*, and *Nblog* were not analyzed as there was only one value with the oscillations, the significance level was set at $p < 0.0042$ (i.e., 0.05/12).

For group comparisons ($n = 20$) t_0 , D , μ , σ , $|\cos(\theta_s)|$, $|\cos(\theta_e)|$, *mode*, *median*, *time delay*, *response time*, *Nblog*, *SNR*, and *SNR/Nblog* were chosen for the simple strokes and triangles. For the oscillations, $\Delta(t_0)$, D , μ , σ , *mode*, *median*, *time delay*, *response time*, *Nblog*, *SNR*, and *SNR(dB)/Nblog* were selected. Parameters of the oscillations were extracted from the stable phase. Only the *SNR* and *SNR/Nblog* were calculated from the whole signal. Due to signals recording problems, the data of four participants were rejected for the analysis of the *Nblog*, *SNR*, and *SNR/Nblog* for the vertical

oscillations during an internal rotation fatigue and of one participant, for the horizontal oscillations during an internal rotation fatigue.

A non-parametric paired Hotelling’s T^2 -test on each series of fast strokes was first performed including all the parameters. This multivariate test assessed whether there are statistical differences between the two conditions (without and with fatigue) considering all the parameters. When the test was statistically significant ($p < 0.05$), the non-parametric paired *U*-Mann–Whitney test was performed on each parameter separately. The statistical significance was set at $p < 0.00385$ (0.05/13) for the simple strokes and the triangles and at $p < 0.0042$ (0.05/12) for the oscillations. Comparisons were performed on all lognormals, considering separately *agonist* and *antagonist* components for the simple strokes, except for the *Nblog*, *SNR*, and *SNR/Nblog*, as the whole signal was considered. No such distinctions between lognormals were made for the triangles and the oscillations since no supplementary information could be assessed. The Cohen’s *d* effect size was also calculated to estimate the importance of the parameters evolution after fatigue. As referred in Sawilowsky (2009) the description for magnitude is the following: $d(0.01)$ = very small, $d(0.20)$ = small, $d(0.50)$ = medium, $d(0.80)$ = large, $d(1.20)$ = very large, and $d(2.0)$ = huge.

Correlation matrices were finally calculated to assess the relationships between parameters. The correlation between the *reaction time* and t_0 was assessed to determine the importance of using t_0 for central system analyses and the

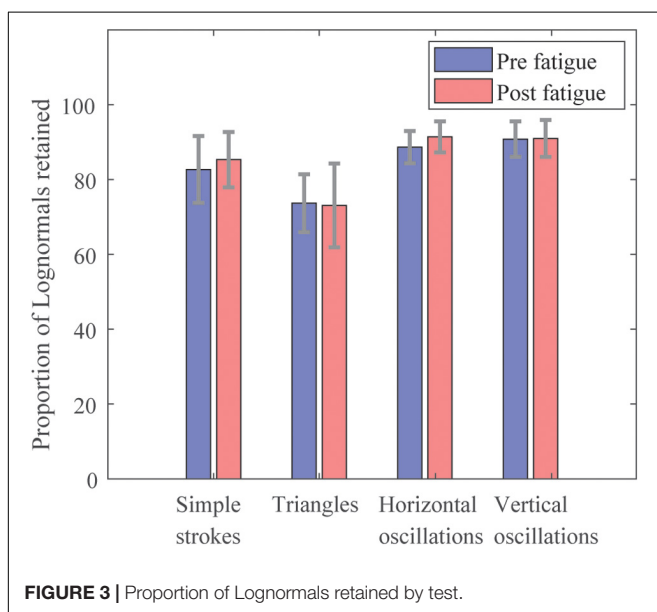
correlation of t_0 with μ to evaluate the independence of parameters related to the central and peripheral systems. To do so, Spearman coefficients were evaluated on the mean of each parameter by test. Statistical significance was set at $p < 0.05$.

RESULTS

The filtering of data led to retain lognormals with properties verifying the conditions mentioned in Section “Data Formatting.” The proportion of the lognormals retained out of the entire set of strokes drawn by test, is illustrated in **Figure 3**. The proportions are similar pre- and post-fatigue. It is observed that horizontal and vertical oscillations have the highest amount of lognormals retained (between $88.7 \pm 0.04\%$ and $91.4 \pm 0.04\%$), whereas triangles count the lowest numbers of them ($73.7 \pm 0.08\%$ pre-fatigue versus $73.1 \pm 0.11\%$ post-fatigue). Simple strokes have around 84% of lognormals retained.

Effects of Fatigue by Participant Torques and Borg Number

As depicted in **Figure 4**, the participants experienced internal rotation (IR) or external rotation (ER) fatigue differently. The ER fatigue trials lasted longer than the IR ones (15 ± 9.4 min versus 3.9 ± 1.6 min). We observed also that the time necessary to fatigue after each series of strokes decreased for the ER fatigue (respectively, 15, 8.9, 6.2, and 5.1 min) whereas it stayed quite stable for the IR fatigue (3.9, 3.2, 3, and 2.7 min). However, for both fatigues, the number on the Borg Scale was similar (7.7 ± 1.4 for the ER fatigue and 7.9 ± 1.1 for the IR fatigue). Despite the shorter time to fatigue, participants perceived intense effort in IR.



Kinematic Parameters

The proportion of participants affected in their parameters by fatigue is presented in **Table 1**. Both central and peripheral systems were affected by fatigue in each stroke of each test, as preliminary reported in Laurent et al. (2019), for both ER or IR fatigue. Parameters affected by fatigue were not necessarily the same among all participants and neither was their evolution. This inter-subject variability is illustrated in **Figure 5**, where four velocity profiles are drawn pre- and post-fatigue, characterizing different participant's behavior. For example, after fatigue, the velocity profile was either displaced to the right (**Figure 5A**), to the left (**Figure 5C**), or not evolving (**Figures 5B,D**).

For the simple strokes, 90% of the participants had almost one parameter describing their central system significantly different post ER fatigue, either in the *agonist* or *antagonist* component (**Table 1**). It was changed in 95% of the population after an IR fatigue. For the parameters reflecting the peripheral system, more differences were noticed in the *agonist* parameters (40%) than in the *antagonist* parameters (15%) after an ER fatigue. After an IR fatigue, no such distinction between *agonist* and *antagonist* parameters was found for the peripheral system (20% of statistical changes in both cases). The conduction time was affected in only 5% of the population after an ER fatigue and 10% after an IR fatigue.

For the triangles, the global state of the neuromotor system was impacted in 10% of the population after an ER fatigue and in no participant at all after IR fatigue. The motor program execution showed numerous differences pre- and post-fatigue for the triangles, as in average more than 85% of the population presented statistical differences (85% after an IR fatigue and 95% after an ER fatigue). The conduction time $|t_0 - RT|$ was affected in 15% of the population after an ER fatigue, and in no participants after an IR fatigue.

The oscillations were the tests in which the most significant differences were observed pre- and post-fatigue. All participants had statistical changes in the parameters related to the central system and the motor program execution, both after ER or IR fatigue. Moreover, after an ER fatigue, peripheral parameters changed in 75% of the population in horizontal oscillations versus 65% in the vertical oscillations. After an IR fatigue, they changed in 85% both in horizontal and vertical oscillations.

Group Effect of Fatigue Parameters Evolution

Hotelling's T^2 -tests were all statistically significant ($p < 0.05$), except for the triangles after an IR fatigue. As a matter of fact, they validated in those cases the use of the *U*-Mann-Whitney test to assess each parameter evolution. For the *agonist* component of simple strokes, σ and the *time delay* were significantly higher after fatigue ($p = 0.0001$), with a medium and large effect size ($d = 0.66$ and 0.93) (**Table 2**). In the *antagonist* components t_0 and the *response time* were significantly higher after fatigue ($p = 0.0001$), with a medium effect size ($d = 0.51$ and 0.56 , respectively). The *SNR/Nblog* significantly decreased after fatigue ($p = 0.0002$). After

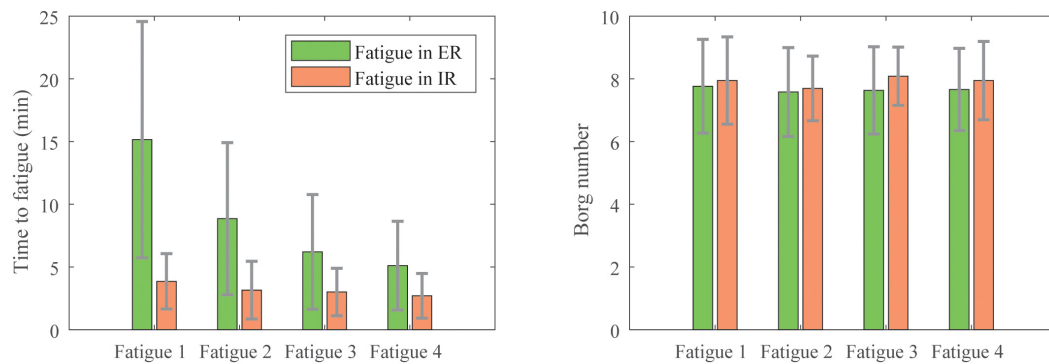


FIGURE 4 | Means and standard deviations of the duration of fatigue (**Left**) and the perceived exertion on the Borg scale (**Right**) of the participants during each of the four fatigue sessions.

TABLE 1 | Percentage of participants ($N = 20$) with significant differences for each component of their tests.

Test	Stroke	Central system	Peripheral system	Both systems	Motor program execution	Global state of the neuromotor system	RT	$ t_0 - RT $ or $ \Delta(t_0) - RT $
Fatigue in external rotation (ER)								
Simple strokes	<i>Agonist</i>	70	40	30	40	50	45	5
	<i>Antagonist</i>	60	15	10	15	45		
	Total	90	40	35	45	55		
Triangles	Stroke 1	45	25	10	55	10	20	15
	Stroke 2	45	20	5	65			
	Stroke 3	25	5	5	55			
	Total	75	30	20	95			
Horizontal oscillations	External rotation	95	80	75	80	x	x	70
	Internal rotation	100	70	70	85			
	Total	100	75	85	95			
Vertical oscillations	External rotation	90	55	55	65	x	x	95
	Internal rotation	90	35	35	75			
	Total	95	65	60	80			
Fatigue in internal rotation (IR)								
Simple strokes	<i>Agonist</i>	95	20	20	25	50	30	10
	<i>Antagonist</i>	30	20	15	25	45		
	Total	95	35	30	45	55		
Triangles	Stroke 1	30	15	5	50	0	10	0
	Stroke 2	20	25	10	65			
	Stroke 3	45	20	20	70			
	Total	75	40	25	85			
Horizontal oscillations	External rotation	95	80	80	90	x	x	85
	Internal rotation	100	75	75	90			
	Total	100	85	85	95			
Vertical oscillations	External rotation	90	75	65	80	x	x	85
	Internal rotation	100	65	65	90			
	Total	100	85	85	90			

We could not perform *t*-test for the parameters reflecting the global state of the neuromotor system and the reaction time (RT) for the oscillations tests as we have only one value per subject. For the oscillations $|\Delta(t_0) - RT|$ was reported instead of $|t_0 - RT|$.

an IR fatigue, t_0 was significantly higher after fatigue only for the *agonist* components (*agonist*, $p = 0.0002$; *antagonist*, $p = 0.7$). The *mode*, *median* and *time delay* increased for both components ($p \leq 0.0002$), with medium effect size, ranging from 0.53 to

0.79. In the triangles, D , the *mode*, the *median* and the *time delay* were significantly higher after an ER fatigue ($p = 0.0001$) (Table 3). For the horizontal oscillations (Table 4), $\Delta(t_0)$ was significantly higher after an ER fatigue with a large effect size

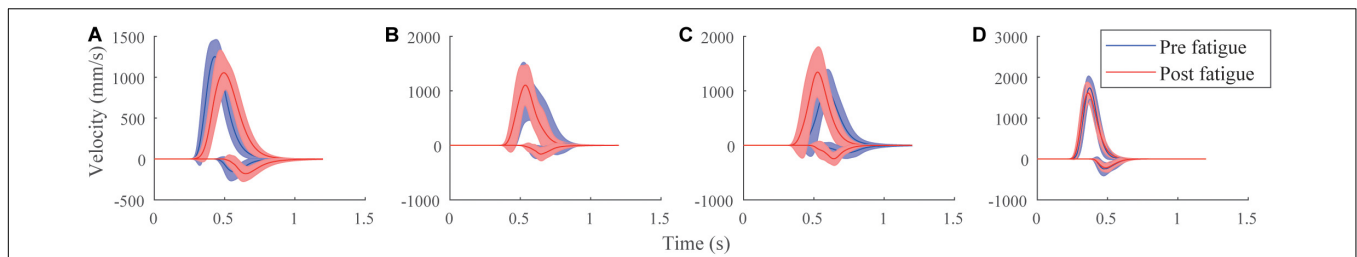


FIGURE 5 | Velocity profiles of four participants with the mean \pm SD of the simple strokes, before (blue) and after (red) fatigue for the *agonist* (positive) and *antagonist* (negative) components. Panel (A) had a significant increase after fatigue in μ , *mode*, *median*, *time delay*, *RT* for the *agonist* parameters; and t_0 , σ , *response time*, *asymmetry*, and *RT* for the *antagonist* parameters. Panel (B) had no statistical changes. Panel (C) had a significant increase in *D* for the *agonist* parameters. It had a significant decrease in $|\cos(\theta_3)|$, μ , *mode*, *median*, *time delay*, *Nblog* in the *agonist* parameters and *SNR* and *TR* in both components. Panel (D) had a significant increase after fatigue of t_0 , σ and $|t_0-RT|$ and a significant decrease in μ and *RT* in the *agonist* parameters.

($d = 0.80$). Regarding the peripheral system, μ increased after both an ER and IR fatigue ($p = 0.0001$), with a large effect size ($d = 0.82$ and 1.21 , respectively). The *mode*, *median*, *time delay* and *response time* were significantly higher ($p = 0.0001$) with a large effect size after ER and IR fatigues ($d = 0.80$ – 1.18). The *SNR/Nblog* was significantly higher after an IR fatigue, with a medium effect size ($d = 0.61$). In the vertical oscillations, *D* was significantly higher ($p = 0.0001$), after an ER fatigue. After an IR fatigue the *mode*, *median*, *time delay*, and *response time* were significantly higher ($p = 0.0001$), with a medium effect size ($d = 0.58$ – 0.61).

Correlation Between Parameters

The reaction time (*RT*) was correlated with t_0 of the *agonist* and *antagonist* component of the simple strokes (Table 5). The correlation with the triangles existed only with the t_0 of the first stroke. These correlations were present pre- and post-fatigue (ρ from 0.78 to 0.89 for the triangles and from 0.72 to 0.93 for the simple strokes). Parameters were more correlated for the *agonist* component than the *antagonist*. For example, before ER fatigue, the correlation was set at $\rho = 0.88$ for the *agonist* component and at $\rho = 0.76$ for the *antagonist* component. Regarding t_0 and μ , there is no evidence of high correlation for the simple strokes (ρ between -0.51 and -0.08 depending on the test) but it appears for the triangles (ρ from -0.86 to -0.47).

DISCUSSION

This study aimed to settle an innovative, economical and non-invasive method to detect shoulder muscular fatigue and discriminate between central and peripheral fatigue.

Distinction of the Type of Fatigue

In the Kinematic Theory of rapid human movements, the distinction between central and peripheral fatigue is possible through the intrinsic properties of the parameters extracted from each stroke (O'Reilly and Plamondon, 2013). In addition to these parameters, we proposed a series of derived parameters, which translate a more global approach of motor control analysis. For the oscillations, using $\Delta(t_0)$ instead of t_0 seems

relevant for the central system analysis. As t_0 represents the timing emission for a command (Plamondon, 1995b), $\Delta(t_0)$ describes the frequency at which the emission command is sent. A statistical change means that the brain rapidity for generating command signals is impaired due to fatigue. On the other hand, the peripheral system is reflected through μ and σ , which are the temporal properties of the neuromuscular system. A significant difference in one of these parameters theoretically means that the peripheral system of the participant was impaired by fatigue. As exposed in Table 1, all the tests performed could reflect those changes. In practice, the correlation between the *reaction time* and t_0 (Table 5) consolidated our position of using t_0 for the central nervous system analysis. In fact, the *reaction time* is a commonly used parameter for cognitive studies (Tanaka et al., 2009; Sant'Ana et al., 2017). This correlation was higher for the first stroke of the simple strokes and the triangles. As the *antagonist* component appears after the *agonist* one, there is a delay, so a lower correlation. The same remark can be made for the triangles: strokes 2 and 3 appear later implying an absence of correlation between *RT* and their t_0 . In addition, the calculation of the conduction time $|t_0-RT|$ enables to locate more precisely the origin of the central fatigue. In our study, it changed for a small population ($\leq 15\%$), whether simple strokes or triangles. This means that the time taken from the brain to propagate the information to the end effector does not change for most of the participants. Moreover, for clinical purposes, it would be of interest to differentiate the fatigued muscle, whether the *infraspinatus* (ER fatigue) or the *subscapularis* (IR fatigue), through the evolution of the Kinematic parameters. Machine learning algorithms, such as support vector machines, have already shown interesting results in discriminating the kinematic parameters in attention deficit hyperactivity disorder and control group children (Faci et al., 2020c). The use of these algorithms could be of interest for a differentiation of the type of neuromuscular fatigue (ER or IR fatigue).

Intra-Participant Follow-Up

Our study showed that an individual monitoring of fatigue is possible using a tablet. It is clinically relevant as pre- and post-fatigue variations are different regarding the type of fatigue, the participant and the test performed. This can be explained by the

TABLE 2 | Parameters evolution in the simple strokes after a shoulder fatigue in external or internal rotation.

	Agonist				Antagonist			
	Pre-fatigue	Post-fatigue	P-value	Effect size	Pre-fatigue	Post-fatigue	P-value	Effect size
External rotation fatigue (N = 20)								
Central system								
t_0	0.23 ± 0.08	0.25 ± 0.07	0.0246	0.32	0.40 ± 0.12	0.44 ± 0.12	0.0002*	0.51 ^a
D	214 ± 19.8	217 ± 21.6	0.005	0.20	30.6 ± 6.94	33.0 ± 8.33	0.0001*	0.33
$ \cos(\theta_s) $	0.81 ± 0.08	0.81 ± 0.10	0.0001*	0.15	0.95 ± 0.03	0.94 ± 0.03	0.0001*	0.31
$ \cos(\theta_e) $	0.96 ± 0.03	0.96 ± 0.03	0.0342	0.33	0.91 ± 0.11	0.94 ± 0.05	0.0001*	0.36
Peripheral system								
μ	-1.42 ± 0.18	-1.41 ± 0.15	0.4648	0.02	-1.78 ± 0.18	-1.82 ± 0.20	0.3042	0.23
σ	0.27 ± 0.06	0.30 ± 0.06	0.0001*	0.66 ^a	0.36 ± 0.11	0.39 ± 0.12	0.0004*	0.47
Motor program execution								
Mode	0.47 ± 0.08	0.48 ± 0.08	0.0002*	0.28	0.59 ± 0.10	0.62 ± 0.10	0.0001*	0.41
Median	0.49 ± 0.08	0.51 ± 0.08	0.0001*	0.33	0.61 ± 0.11	0.64 ± 0.11	0.0001*	0.45
Time delay	0.50 ± 0.09	0.51 ± 0.09	0.0001*	0.36	0.62 ± 0.11	0.65 ± 0.11	0.0001*	0.47
Response time	0.07 ± 0.02	0.08 ± 0.02	0.0001*	0.93 ^b	0.06 ± 0.03	0.07 ± 0.03	0.0001*	0.56 ^a
Global state of the neuromotor system – whole stroke								
Nblog	2.18 ± 0.26	2.29 ± 0.30	0.0002*	0.48				
SNR	30.3 ± 1.14	29.7 ± 1.29	0.0001*	0.44				
SNR/Nblog	14.7 ± 1.53	13.8 ± 1.90	0.0002*	0.52 ^a				
Internal rotation fatigue (N = 20)								
Central system								
t_0	0.21 ± 0.08	0.24 ± 0.08	0.0002*	0.52 ^a	0.40 ± 0.11	0.41 ± 0.11	0.7018	0.20
D	208 ± 25.0	208 ± 23.7	0.713	0.00	29.0 ± 6.66	30.4 ± 7.53	0.2932	0.23
$ \cos(\theta_s) $	0.81 ± 0.09	0.79 ± 0.09	0.0001*	0.65 ^a	0.93 ± 0.06	0.94 ± 0.04	0.7014	0.25
$ \cos(\theta_e) $	0.95 ± 0.03	0.96 ± 0.02	0.0001*	0.44	0.91 ± 0.09	0.93 ± 0.03	0.0004*	0.47
Peripheral system								
μ	-1.43 ± 0.18	-1.46 ± 0.18	0.0154	0.27	-1.86 ± 0.22	-1.80 ± 0.16	0.0454	0.24
σ	0.27 ± 0.05	0.29 ± 0.05	0.0001*	0.40	0.36 ± 0.08	0.36 ± 0.10	0.2972	0.06
Motor program execution								
Mode	0.45 ± 0.08	0.47 ± 0.09	0.0002*	0.53 ^a	0.57 ± 0.10	0.59 ± 0.10	0.0002*	0.79 ^a
Median	0.47 ± 0.09	0.49 ± 0.09	0.0001*	0.58 ^a	0.59 ± 0.10	0.61 ± 0.11	0.0001*	0.73 ^a
Time delay	0.48 ± 0.09	0.50 ± 0.10	0.0001*	0.58 ^a	0.60 ± 0.11	0.63 ± 0.12	0.0001*	0.67 ^a
Response time	0.07 ± 0.02	0.07 ± 0.02	0.0001*	0.40	0.06 ± 0.02	0.06 ± 0.03	0.0492	0.17
Global state of the neuromotor system – whole stroke								
Nblog	2.20 ± 0.24	2.20 ± 0.25	0.9086	0.02				
SNR	30.1 ± 1.12	30.0 ± 1.37	0.7352	0.09				
SNR/Nblog	14.5 ± 1.52	14.5 ± 1.77	0.9802	0.00				

The simple strokes are separated into their agonist or antagonist components.

Values are expressed as mean ± standard deviation.

*Represents statistical significance ($p < 0.05/13 = 0.0038$).

^{a,b}Represent medium and large effect sizes, respectively.

task dependency of fatigue (Enoka and Stuart, 1992) and the uniqueness of each participant (Chaffin et al., 2006b). In fact, during submaximal muscle contraction, fatigue development depends on the type of fibers activated and on the duration of the contraction (Enoka and Stuart, 1992; Chaffin et al., 2006a). Large variability in duration may come from the inter-subject difference, but also from their ability to generate a maximal force (Edwards, 1981). Different strategies were taken by the participants to counteract the effects of fatigue, which is reflected in the kinematic parameters. This can be observed as

well in **Figure 5**, where behavior differences are illustrated for four velocity profiles. Some participants have a slower general response (**Figure 5A**), faster (**Figure 5C**), or not evolving due to fatigue (**Figures 5B,D**). The participant in **Figure 5B** was a former high-level swimming athlete, and therefore could be accustomed to shoulder fatigue. The participant in **Figure 5D** presents many statistical differences in parameters shown by the *U*-Mann-Whitney test. As μ – the longtime delay – significantly decreased and σ – the log response time – significantly increased for this participant, the resulting velocity profile showed little

TABLE 3 | Parameters evolution in the triangles, after a shoulder fatigue in external rotation.

	Triangles – whole stroke (<i>N</i> = 20)			
	Pre-fatigue	Post-fatigue	<i>P</i> -value	Effect size
Central system				
<i>t</i> ₀	0.23 ± 0.08	0.23 ± 0.06	0.488	0.12
<i>D</i>	154.9 ± 9.29	158.0 ± 10.2	0.0001*	0.53 ^a
Peripheral system				
<i>μ</i>	−0.80 ± 0.21	−0.77 ± 0.18	0.0028*	0.31
<i>σ</i>	0.19 ± 0.03	0.19 ± 0.03	0.0226	0.25
Motor program execution				
Mode	0.74 ± 0.08	0.75 ± 0.08	0.0001*	0.40
Median	0.75 ± 0.09	0.77 ± 0.08	0.0001*	0.38
Time delay	0.76 ± 0.09	0.77 ± 0.08	0.0001*	0.36
Response time	0.08 ± 0.01	0.08 ± 0.01	0.6214	0.01
Global state of the neuromotor system				
Nblog	5.18 ± 0.61	5.21 ± 0.38	0.6678	0.06
SNR	27.0 ± 0.30	26.8 ± 0.47	0.0034*	0.46
SNR/Nblog	5.37 ± 0.69	5.37 ± 0.45	0.7124	0.01

Parameters come from combined strokes.

Values are expressed as mean ± standard deviation.

*Represents statistical significance ($p < 0.05/11 = 0.0045$).

^aRepresents medium effect size.

visual differences after fatigue as compared to before. The counterbalanced parameter changes masked the fatigue effect on the velocity profile, pointing out the interest of analyzing the lognormal parameters. Moreover, some participants have a higher μ (Figures 5A,C), meaning a diminution of the neuromuscular system to respond rapidly to a command, whereas some others have a lower value (Figure 5D). To compensate for a lower μ , participants can use a higher σ or t_0 . This, respectively, means that the participant will take more time to make the entire movement and that the brain will send the response command later. However, sometimes t_0 significantly decreased after fatigue. The participant reacted faster to the stimulus, as physical exercise can improve someone's cognitive function (Hillman et al., 2008). Different action plans for drawing strokes were also made: some participants made for example shorter strokes (smaller *D*, Figure 5A) because they had difficulties in executing them, whereas some others made larger strokes (higher *D*, Figure 5C) because they had, for instance, difficulties in stopping them. This is probably due to motor variability as movement is reorganized to prevent the apparition of disorders. In that sense, spatiotemporal muscular recruitment is variable after fatigue, which is assessed here (Falla and Farina, 2007; Srinivasan and Mathiassen, 2012; Yang et al., 2018). This method enables to study the evolution of each parameters and compensations made by participants for a case-by-case study, which is essential for example in personalized top-level athletes training.

Group Effect of Fatigue

A group effect was noticed from the analyses, signifying that a general pattern is highlighted after a neuromuscular fatigue. This analysis is a first step in the process of using the method in clinics.

More studies would be needed to ensure that the parameters evolution highlighted in this study are specific to shoulder fatigue. In the simple strokes, the peripheral system was more impacted after an ER than an IR fatigue. As the time to fatigue was longer in ER, additional mechanisms of fatigue may have been present, such as at the level of the excitation-contraction coupling (Baker et al., 1993), which is then observed in the Kinematic parameters related to the peripheral system. On the contrary, the IR fatigue was perceived harder and may have impacted more the parameters related to the central system. In fact, the action plan of the agonist components of simple strokes is changed, with for example an increase of the time to send the motor command. Moreover, it was noticed that the motor program execution was the most impaired system for most of the tests (Tables 2–4). In fact, the *mode*, *median*, *time delay* were significantly higher after fatigue, meaning a decline in the command velocity (Laniel et al., 2019). However, parameters describing the global state of the neuromotor system, such as the *Nblog* and *SNR/Nblog* have a general trend to increase and decrease, respectively, for the simple strokes and horizontal oscillations after an external rotation fatigue. The evolution of these two parameters reflects a worsening of the motor control quality. In fact, as in Cortes et al. (2014), fatigue impacts the smoothness and motor control of a person. On the other hand, the *SNR* does not seem to change in many cases, only in the simple strokes and triangles after an ER fatigue. As explained in Laniel et al. (2019), the reconstruction of the velocity profiles stops when a 25 dB *SNR* is reached, and adds lognormals until that condition is met. Studying the *SNR* of simple strokes after a Delta-Lognormal extraction might be more appropriate (Plamondon, 1995a, 1998; Woch et al., 2011), as it is expected to reconstruct the kinematics with only two lognormals, the *agonist* and *antagonist*. The evolution of the parameters, and

TABLE 4 | Parameters evolution in the horizontal and vertical oscillations, after a shoulder fatigue in external or internal rotation.

Horizontal oscillations (<i>N</i> = 20)								
	External rotation fatigue				Internal rotation fatigue			
	Pre-fatigue	Post-fatigue	<i>P</i> -value	Effect size	Pre-fatigue	Post-fatigue	<i>P</i> -value	Effect size
Central system								
$\Delta(t_0)$	0.08 ± 0.01	0.09 ± 0.01	0.0012*	0.80 ^b	0.09 ± 0.01	0.09 ± 0.01	0.0150	0.96
<i>D</i>	124.8 ± 23.7	126.2 ± 26.2	0.0008*	0.10	124.2 ± 24.5	125.2 ± 28.8	0.2290	0.04
Peripheral system								
μ	−0.81 ± 0.07	−0.75 ± 0.10	0.0001*	0.82 ^b	−0.81 ± 0.11	−0.75 ± 0.10	0.0001*	1.21 ^c
σ	0.06 ± 0.00	0.06 ± 0.00	0.8870	0.09	0.06 ± 0.00	0.06 ± 0.00	0.0568	0.32
Motor program execution								
Mode	0.53 ± 0.04	0.56 ± 0.06	0.0001*	0.80 ^b	0.53 ± 0.06	0.57 ± 0.06	0.0001*	1.18 ^b
Median	0.53 ± 0.04	0.57 ± 0.06	0.0001*	0.80 ^b	0.53 ± 0.06	0.57 ± 0.06	0.0001*	1.18 ^b
Time delay	0.53 ± 0.04	0.57 ± 0.06	0.0001*	0.80 ^b	0.53 ± 0.06	0.57 ± 0.06	0.0001*	1.18 ^b
Response time	0.03 ± 0.02	0.03 ± 0.00	0.0001*	0.81 ^b	0.03 ± 0.00	0.03 ± 0.00	0.0001*	1.17 ^b
Global state of the neuromotor system								
Nblog	94.7 ± 6.78	89.1 ± 8.79	0.0002*	0.76 ^a	94.7 ± 10.4	89.4 ± 9.21	0.0002*	0.66 ^a
SNR	28.4 ± 1.40	27.6 ± 1.26	0.0868	0.38	28.4 ± 1.12	28.2 ± 0.99	0.3824	0.10
SNR/Nblog	0.21 ± 0.02	0.23 ± 0.03	0.0068	0.56 ^a	0.21 ± 0.03	0.23 ± 0.02	0.0002*	0.61 ^a
Vertical oscillations (<i>N</i> = 20)								
Central system								
$\Delta(t_0)$	0.09 ± 0.01	0.09 ± 0.01	0.7646	0.09	0.09 ± 0.02	0.10 ± 0.01	0.0256	0.56
<i>D</i>	118.4 ± 19.2	123.9 ± 19.4	0.0001*	0.57 ^a	119.4 ± 26.0	119.8 ± 21.6	0.4200	0.03
Peripheral system								
μ	−0.73 ± 0.11	−0.73 ± 0.14	0.9816	0.05	−0.74 ± 0.17	−0.69 ± 0.16	0.0001*	0.64 ^a
σ	0.06 ± 0.00	0.06 ± 0.00	0.6778	0.04	0.06 ± 0.00	0.06 ± 0.00	0.1486	0.23
Motor program execution								
Mode	0.57 ± 0.07	0.58 ± 0.09	0.4820	0.11	0.58 ± 0.10	0.61 ± 0.09	0.0001*	0.61 ^a
Median	0.58 ± 0.07	0.58 ± 0.09	0.4854	0.11	0.59 ± 0.10	0.61 ± 0.09	0.0001*	0.61 ^a
Time delay	0.58 ± 0.07	0.58 ± 0.09	0.4906	0.11	0.59 ± 0.10	0.61 ± 0.09	0.0001*	0.61 ^a
Response time	0.03 ± 0.00	0.028 ± 0.004	0.1838	0.11	0.03 ± 0.01	0.03 ± 0.004	0.0001*	0.58 ^a
Global state of the neuromotor system								
Nblog	88.0 ± 9.51	88.1 ± 12.1	0.9334	0.03	88.4 ± 14.9	84.8 ± 12.1	0.0314	0.40
SNR	27.6 ± 1.81	27.9 ± 1.46	0.5268	0.15	28.0 ± 1.64	27.4 ± 1.39	0.3432	0.27
SNR/Nblog	0.22 ± 0.02	0.23 ± 0.04	0.3886	0.20	0.23 ± 0.04	0.24 ± 0.04	0.6482	0.18

Parameters come from combined strokes.

Values are expressed as mean ± standard deviation.

For the global state of the neuromotor system, *N* = 19 after an ER fatigue and *N* = 16 after an IR fatigue.

*Represents statistical significance (*p* < 0.05/11 = 0.0045).

^{a,b,c}Represent medium, large and very large effect sizes, respectively.

especially the ones reflecting the motor program execution, is similar between participants, which is interesting for using the tablet as a clinical tool for fatigue detection.

Performance of the Tests

As a general overview, simple strokes reveal information about *agonist/antagonist* systems. According to Turpin et al. (2011), muscular activity changes after fatigue but the coordination between muscles does not. The same muscles will create the *agonist/antagonist* synergy. For this purpose, analyzing and discriminating the two categories of muscles is appropriate. In case of complex tasks (i.e., triangles or oscillations), the distinction between those two systems is meaningless

since there is no stop at the intermediate points, only at the end. The use of the speed/accuracy tradeoff tests could provide more information, as it can express further relationships between *agonist* and *antagonist* components and their evolution with fatigue (O'Reilly and Plamondon, 2013). Moreover, central and peripheral system parameters do not show a correlation in simple strokes (t_0 and μ , Table 5), whereas this correlation exists when movements get longer. Participants anticipate them by targeting virtual points (Plamondon et al., 2003; Plamondon and Djoua, 2006). With the independency of parameters in the simple strokes, this test can be specifically used to differentiate a central from a peripheral fatigue.

TABLE 5 | Spearman correlation coefficients between parameters, with rho-values of correlation pre- and post-fatigue.

Parameters	Test	Type of fatigue	Type of stroke	Rho pre-fatigue	Rho post-fatigue
t_0 and RT	Simple strokes	ER	Agonist	0.88*	0.93*
			Antagonist	0.76*	0.72*
		IR	Agonist	0.91*	0.92*
			Antagonist	0.86*	0.74*
	Triangles	ER	Stroke 1	0.89*	0.78*
			Stroke 2	0.07	0.20
			Stroke 3	-0.02	0.15
		IR	Stroke 1	0.78*	0.81*
			Stroke 2	0.23	-0.10
			Stroke 3	0.10	0.24
t_0 and μ	Simple strokes	ER	Agonist	-0.37	-0.19
			Antagonist	-0.42	-0.51*
		IR	Agonist	-0.08	-0.21
			Antagonist	-0.14	-0.18
	Triangles	ER	Stroke 1	-0.74*	-0.61*
			Stroke 2	-0.86*	-0.83*
			Stroke 3	-0.74*	-0.85*
		IR	Stroke 1	-0.58*	-0.47*
			Stroke 2	-0.83*	-0.82*
			Stroke 3	-0.76*	-0.52*

*Represents significant correlations ($p < 0.05$).

On the other hand, triangles seem to be more difficult to perform than simple strokes and oscillations. The inter-participant variability may be higher and therefore significant differences can be harder to notice in group studies. However, as depicted in **Table 1**, individual changes are detectable on triangles, and can therefore be used for extensive studies. A more specific method for extracting triangles may be more adapted, as they have the lowest number of lognormals retained compared to other tests (**Figure 3**). As a matter of fact, it also depicts the importance of performing more repetitions.

In addition, studying larger movements, such as the oscillations seems more efficient to detect fatigue, as they depict a more biomechanical movement. Nevertheless, a compensatory effect between participants is noticed for the vertical oscillations. As participants adopted different postures to execute the movements, the individual kinematic might have been affected (Fuller et al., 2009). It would have been interesting to record the overall kinematic of the upper-body. In fact, a test performance is often the same pre- and post-fatigue, but strategies to perform the tests are different (Côté et al., 2002; Emery and Côté, 2012).

Opening

In a wider context, the Sigma-Lognormal model seems appropriate to study fatigue at different levels of the body, whether it is the upper limb or lower limb. In fact, fatigue results in deficiency in motor control and motion changes due to a modification at different biological levels of the human body physiology (Enoka and Stuart, 1992; Gandevia, 2001;

Cortes et al., 2014). That is why, it is expected that any other impairment in the body, due to fatigue, could be detectable by a similar method. Also in Cowley and Gates (2017), it has been noticed that finger or shoulder fatigue affect movement coordination in different manners. In this way, we think that it would be possible to discriminate fatigue from different parts of the body and parameters from the theory could reflect those changes. By performing wider movements with the use of a white board (Fischer et al., 2014), or by registering them in 3D (Schindler et al., 2018), it would then probably be easier to discriminate them. The use of a board seems interesting, as the system would remain easy to use. As the Kinematic Theory describes fine motor control and is suitable for many end effectors [such as fingertips, head (Lebel et al., 2018b), eye movements (Plamondon, 1995a) etc.], the use of markers directly on the studied articulation would be interesting to complete analyses.

CONCLUSION

This study highlights that shoulder neuromuscular fatigue is detectable in healthy active adults with the use of a digitizing tablet and the Kinematic Theory. The type of fatigue (central or peripheral) and the location of central fatigue (preparation or conduction time) are distinguishable through the parameters extracted from handwriting. An individual monitoring is relevant to determine the compensatory reactions made by each participant to counteract the effects of fatigue. Overall, common

patterns in the parameters evolution are noticeable and are significant for clinical studies. Parameters having a more global approach, such as the *mode*, *median*, *time delay* tend to increase after fatigue, whereas the *SNR/Nblog* tends to decrease. We also observed that all handwriting tests were sensitive to fatigue. Nevertheless, the simple strokes test could discriminate between the central and peripheral systems independently and between the *agonist/antagonist* systems, and the oscillations test is the most effective to detect shoulder fatigue.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, to any qualified researcher interested in collaborating with our team.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Research Ethics Committee of Polytechnique Montréal (CER-1819-23 v.3). The patients/participants provided their written informed consent to participate in this study.

REFERENCES

- Al-Mulla, M. R., Sepulveda, F., and Colley, M. (2011). A review of non-invasive techniques to detect and predict localised muscle fatigue. *Sensors* 11, 3545–3594. doi: 10.3390/s110403545
- Al-Mulla, M. R., Sepulveda, F., and Colley, M. (2012). “sEMG techniques to detect and predict localised muscle fatigue,” in *EMG Methods for Evaluating Muscle and Nerve Function*, ed. M. Mr. Schwartz (Rijeka: InTech), doi: 10.5772/25678
- Baker, A. J., Kostov, K. G., Miller, R. G., and Weiner, M. W. (1993). Slow force recovery after long-duration exercise: metabolic and activation factors in muscle fatigue. *J. Appl. Physiol.* 74, 2294–2300. doi: 10.1152/jappl.1993.74.5.2294
- Beach, M. L., Whitney, S. L., and Dickoff-Hoffman, S. (1992). Relationship of shoulder flexibility, strength, and endurance to shoulder pain in competitive swimmers. *J. Orthop. Sports Phys. Ther.* 16, 262–268. doi: 10.2519/jospt.1992.16.6.262
- Bevan, S. (2015). Economic impact of musculoskeletal disorders (MSDs) on work in Europe. *Best Pract. Res. Clin. Rheumatol.* 29, 356–373. doi: 10.1016/j.berh.2015.08.002
- Bigland-Ritchie, B., and Woods, J. J. (1984). Changes in muscle contractile properties and neural control during human muscular fatigue. *Muscle Nerve* 7, 691–699. doi: 10.1002/mus.880070902
- Borg, G. (1998). *Borg's Perceived Exertion and Pain Scales*. Champaign, IL: Human kinetics.
- Boyas, S., and Guével, A. (2011). Neuromuscular fatigue in healthy muscle: underlying factors and adaptation mechanisms. *Ann. Phys. Rehabil. Med.* 54, 88–108. doi: 10.1016/j.rehab.2011.01.001
- Chaffin, D. B., Andersson, G., and Martin, B. J. (2006a). *Occupational Biomechanics*. New York, NY: Wiley.
- Chaffin, D. B., Andersson, G. B. J., and Martin, B. J. (2006b). *The Structure and Function of the Musculoskeletal System and (Chapter 2) Occupational Biomechanics*, 4th Edn, Hoboken, NJ: Wiley-Interscience, 11–36.
- Chen, M. J., Fan, X., and Moe, S. T. (2002). Criterion-related validity of the Borg ratings of perceived exertion scale in healthy individuals: a meta-analysis. *J. Sports Sci.* 20, 873–899. doi: 10.1080/026404102320761787

AUTHOR CONTRIBUTIONS

AL, RP, and MB contributed conception and design of the study. AL recruited the participants and performed data collection, data formatting, and statistical analyses. AL, RP, and MB interpreted the results. AL wrote the first draft of the manuscript. RP and MB critically reviewed it. The final version was approved by all authors.

FUNDING

This research was conducted as part of the TransMedTech Institute's activities and AL was funded in part by the Canada First Research Excellence Fund. This work was also partly supported by NSERC CANADA Discovery Grants RGPIN-2015-06409 to RP and RGPIN-2019-04978 to MB.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnhum.2020.00171/full#supplementary-material>

- Cifrek, M., Medved, M., Tonković, S., and Ostojić, S. (2009). Surface EMG based muscle fatigue evaluation in biomechanics. *Clin. Biomech.* 24, 327–340. doi: 10.1016/j.clinbiomech.2009.01.010
- Cortes, N., Onate, J., and Morrison, S. (2014). Differential effects of fatigue on movement variability. *Gait Post.* 39, 888–893. doi: 10.1016/j.gaitpost.2013.11.020
- Côté, J. N., Mathieu, P. A., Levin, M. F., and Feldman, A. G. (2002). Movement reorganization to compensate for fatigue during sawing. *Exp. Brain Res.* 146, 394–398. doi: 10.1007/s00221-002-1186-6
- Cowley, J. C., and Gates, D. H. (2017). Proximal and distal muscle fatigue differentially affect movement coordination. *PLoS One* 12:e0172835. doi: 10.1371/journal.pone.0172835
- Dittner, A. J., Wessely, S. C., and Brown, R. G. (2004). The assessment of fatigue: a practical guide for clinicians and researchers. *J. Psychosom. Res.* 56, 157–170. doi: 10.1016/S0022-3999(03)00371-4
- Djioua, M., and Plamondon, R. (2010). The limit profile of a rapid movement velocity. *Hum. Mov. Sci.* 29, 48–61. doi: 10.1016/j.humov.2009.02.007
- Ebaugh, D. D., McClure, P. W., and Karduna, A. R. (2006). Effects of shoulder muscle fatigue caused by repetitive overhead activities on scapulothoracic and glenohumeral kinematics. *J. Electromyogr. Kinesiol.* 16, 224–235. doi: 10.1016/j.jelekin.2005.06.015
- Edwards, R. G., and Lippold, O. C. J. (1956). The relation between force and integrated electrical activity in fatigued muscle. *J. Physiol.* 132, 677–681. doi: 10.1113/jphysiol.1956.sp005558
- Edwards, R. H. (1981). “Human muscle function and fatigue. *Ciba Found. Symp.* 82, 1–18.
- Emery, K., and Côté, J. N. (2012). Repetitive arm motion-induced fatigue affects shoulder but not endpoint position sense. *Exp. Brain Res.* 216, 553–564. doi: 10.1007/s00221-011-2959-6
- Enoka, R. M., and Stuart, D. G. (1992). Neurobiology of muscle fatigue. *J. Appl. Physiol.* 72, 1631–1648. doi: 10.1152/jappl.1992.72.5.1631
- Faci, N., Boyogueno Bidias, S. P., Plamondon, R., and Bergeron, N. (2018). “A new experimental set-up to run neuromuscular tests,” in *Proceedings of the International Conference on Pattern Recognition and Artificial Intelligence*, Montreal, 753–757.

- Faci, N., Boyogueno Bidas, S. P., Plamondon, R., and Bergeron, N. (2020a). "An interactive tablet-based system to run neuromuscular tests," in *The Lognormality Principle: Applications for e-Security, e-Health and e-Learning*. (New York, NY: World Scientific Publishing).
- Faci, N., Désiré, N., Beauchamp, M. H., Gagnon, I., and Plamondon, R. (2020b). "Analysing the evolution of children neuromotor system lognormality after mild traumatic brain injury," in *The Lognormality Principle: Applications for e-Security, e-Health and e-Learning*. (New York, NY: World Scientific Publishing).
- Faci, N., Nguyen, H. T., Laniel, P., Gauthier, B., Beauchamp, M. H., Nakagawa, M., et al. (2020c). "Classifying the kinematics of fast pen strokes in children with ADHD using different machine learning models," in *The Lognormality Principle: Applications for e-Security, e-Health and e-Learning*, (New York, NY: World Scientific Publishing).
- Falla, D., and Farina, D. (2007). Periodic increases in force during sustained contraction reduce fatigue and facilitate spatial redistribution of trapezius muscle activity. *Exp. Brain Res.* 182, 99–107. doi: 10.1007/s00221-007-0974-4
- Farina, D., Merletti, R., Indino, B., and Graven-Nielsen, T. (2004). Surface EMG crosstalk evaluated from experimental recordings and simulated signals. *Methods Inf. Med.* 43, 30–35.
- Finsterer, J. (2012). Biomarkers of peripheral muscle fatigue during exercise. *BMC Musculoskelet. Disord.* 13:218. doi: 10.1186/1471-2474-13-218
- Fischer, A., Plamondon, R., O'Reilly, C., and Savaria, Y. (2014). "Neuromuscular representation and synthetic generation of handwritten whiteboard notes," in *Proceedings of International Conference on Frontiers in Handwriting Recognition, ICFHR*, Heraklion.
- Flash, T., and Hogan, N. (1985). The coordination of arm movements: an experimentally confirmed mathematical model. *J. Neurosci.* 5, 1688–1703. doi: 10.1523/JNEUROSCI.05-07-01688.1985
- Fuller, J. R., Lomond, K. V., Fung, J., and Côté, J. N. (2009). Posture-movement changes following repetitive motion-induced shoulder muscle fatigue. *J. Electromyogr. Kinesiol.* 19, 1043–1052. doi: 10.1016/j.jelekin.2008.10.009
- Gandevia, S. C. (2001). Spinal and supraspinal factors in human muscle fatigue. *Physiol. Rev.* 81, 1725–1789. doi: 10.1152/physrev.2001.81.4.1725
- Gaudet, S., Tremblay, J., and Begon, M. (2018). Muscle recruitment patterns of the subscapularis, serratus anterior and other shoulder girdle muscles during isokinetic internal and external rotations. *J. Sports Sci.* 36, 985–993. doi: 10.1080/02640414.2017.1347697
- Hagberg, M., and Wegman, D. H. (1987). Prevalence rates and odds ratios of shoulder-neck diseases in different occupational groups. *Br. J. Ind. Med.* 44, 602–610. doi: 10.1136/oem.44.9.602
- Hermens, H. J., Freriks, B., Disselhorst-Klug, C., and Rau, G. (2000). Development of recommendations for SEMG sensors and sensor placement procedures. *J. Electromyogr. Kinesiol.* 10, 361–374. doi: 10.1016/s1050-6411(00)00027-4
- Hillman, C. H., Erickson, K. I., and Kramer, A. F. (2008). Be smart, exercise your heart: exercise effects on brain and cognition. *Nat. Rev. Neurosci.* 9, 58–65. doi: 10.1038/nrn2298
- Hogan, N. (1984). An organizing principle for a class of voluntary movements. *J. Neurosci.* 4, 2745–2754. doi: 10.1523/JNEUROSCI.04-11-02745.1984
- Huang, Q. H., Zheng, Y. P., Chena, X., He, J. F., and Shi, J. (2007). A System for the synchronized recording of sonomyography, electromyography and joint angle. *Open Biomed. Eng. J.* 1, 77–84. doi: 10.2174/1874120700701010077
- Ibitoye, M. O., Hamzaid, N. A., Zuniga, J. M., and Abdul Wahab, A. K. (2014). Mechanomyography and muscle function assessment: a review of current state and prospects. *Clin. Biomech.* 29, 691–704. doi: 10.1016/j.clinbiomech.2014.04.003
- Joshi, M., Thigpen, C. A., Bunn, K., Karas, S. G., and Padua, D. A. (2011). Shoulder external rotation fatigue and scapular muscle activation and kinematics in overhead athletes. *J. Athl. Train.* 46, 349–357. doi: 10.4085/1062-6050-46.4.349
- Laniel, P., Faci, N., Plamondon, R., Beauchamp, M. H., and Gauthier, B. (2019). Kinematic analysis of fast pen strokes in children with ADHD. *Appl. Neuropsychol. Child* 9, 125–140. doi: 10.1080/21622965.2018.1550402
- Laurent, A., Plamondon, R., and Begon, M. (2019). "Pre-screening for central or peripheral shoulder fatigue using the sigma-lognormal model," in *Proceedings of the International Graphonomics Society - Your Brain on Art*, Cancun.
- Lebel, K., Nguyen, H., Duval, C., Plamondon, R., and Boissy, P. (2017). Capturing the cranio-caudal signature of a turn with inertial measurement systems: methods, parameters robustness and reliability. *Front. Bioeng. Biotechnol.* 5:51. doi: 10.3389/fbioe.2017.00051
- Lebel, K., Nguyen, H., Duval, C., Plamondon, R., and Boissy, P. (2018a). Cranio-caudal kinematic turn signature assessed with inertial systems as a marker of mobility deficits in Parkinson's disease. *Front. Neurol.* 9:22. doi: 10.3389/fneur.2018.00022
- Lebel, K., Nguyen, H., Duval, C., Plamondon, R., and Boissy, P. (2018b). Turn cranio-caudal signature assessment from inertial systems for mobility deficit identification in Parkinson's disease patients. *Parkinsonism Relat. Disord.* 46, e24–e25. doi: 10.1016/j.parkreldis.2017.11.081
- Lee, K. A., Hicks, G., and Nino-Murcia, G. (1991). Validity and reliability of a scale to assess fatigue. *Psychiatry Res.* 36, 291–298. doi: 10.1016/0165-1781(91)90027-M
- Leung, A. W. S., Chan, C. C. H., Lee, A. H. S., and Lam, K. W. H. (2004). Visual analogue scale correlates of musculoskeletal fatigue. *Percept. Mot. Skills* 99, 235–246. doi: 10.2466/pms.99.1.235-246
- Lindström, L., Kadefors, R., and Petersén, I. (1977). An electromyographic index for localized muscle fatigue. *J. Appl. Physiol.* 43, 750–754. doi: 10.1152/jappl.1977.43.4.750
- Mancini, D. M., Bolinger, L., Li, H., Kendrick, K., Chance, B., and Wilson, J. R. (1994). Validation of near-infrared spectroscopy in humans. *J. Appl. Physiol.* 77, 2740–2747. doi: 10.1152/jappl.1994.77.6.2740
- Merletti, R., Rainoldi, A., and Farina, D. (2001). Surface electromyography for noninvasive characterization of muscle. *Exerc. Sport Sci. Rev.* 29, 20–25. doi: 10.1097/00003677-200101000-00005
- Merletti, R., Rainoldi, A., and Farina, D. (2004). "Myoelectric manifestations of muscle fatigue," in *Electromyography - Physiology, Engineering and Noninvasive Applications*, 1st Edn, eds R. Merletti, and P. Parker (Hoboken, NJ: John Wiley & Sons), 233–258.
- Merton, P. A. (1954). Voluntary strength and fatigue. *J. Physiol.* 123, 553–564.
- Mullaney, M. J., McHugh, M. P., Donofrio, T. M., and Nicholas, S. J. (2005). Upper and lower extremity muscle fatigue after a baseball pitching performance. *Am. J. Sports Med.* 33, 108–113. doi: 10.1177/0363546504266071
- Nadeau, A., Lungu, O., Boré, A., Plamondon, R., Duchesne, C., Robillard, M.-E., et al. (2018). A 12-week cycling training regimen improves upper limb functions in people with Parkinson's disease. *Front. Hum. Neurosci.* 12:351. doi: 10.3389/fnhum.2018.00351
- Neely, G., Ljunggren, G., Sylvén, C., and Borg, G. (1992). Comparison between the Visual Analogue Scale (VAS) and the category ratio scale (CR-10) for the evaluation of leg exertion. *Int. J. Sports Med.* 13, 133–136. doi: 10.1055/s-2007-1021244
- O'Reilly, C., and Plamondon, R. (2009). Development of a Sigma-Lognormal representation for on-line signatures. *Pattern Recogn.* 42, 3324–3337. doi: 10.1016/j.patcog.2008.10.017
- O'Reilly, C., and Plamondon, R. (2011). Impact of the principal stroke risk factors on human movements. *Hum. Mov. Sci.* 30, 792–806. doi: 10.1016/j.humov.2010.07.010
- O'Reilly, C., and Plamondon, R. (2013). Agonistic and antagonistic interaction in speed/accuracy tradeoff: a delta-lognormal perspective. *Hum. Mov. Sci.* 32, 1040–1055. doi: 10.1016/j.humov.2012.07.005
- O'Reilly, C., Plamondon, R., Landou, M. K., and Stemmer, B. (2013). Using kinematic analysis of movement to predict the time occurrence of an evoked potential associated with a motor command. *Eur. J. Neurosci.* 37, 173–180. doi: 10.1111/ejn.12039
- O'Reilly, C., Plamondon, R., and Lebrun, L.-H. (2014). Linking brain stroke risk factors to human movement features for the development of preventive tools. *Front. Aging Neurosci.* 6:150. doi: 10.3389/fnagi.2014.00150
- Palmieri, R. M., Ingersoll, C. D., and Hoffman, M. A. (2004). The hoffmann reflex: methodologic considerations and applications for use in sports medicine and athletic training research. *J. Athl. Train.* 39, 268–277.
- Piper, H. (1912). *Elektrophysiologie Menschlicher Muskeln*. Berlin: Springer.
- Plamondon, R. (1995a). A kinematic theory of rapid human movements. Part I. Movement representation and generation. *Biol. Cybern.* 72, 295–307. doi: 10.1007/bf00202785
- Plamondon, R. (1995b). A kinematic theory of rapid human movements. Part II. Movement time and control. *Biol. Cybern.* 72, 309–320. doi: 10.1007/bf00202786

- Plamondon, R. (1998). A kinematic theory of rapid human movements: part III. Kinetic outcomes. *Biol. Cybern.* 78, 133–145. doi: 10.1007/s004220050420
- Plamondon, R. (2020). “The lognormality principle: a personalized survey,” in *The Lognormality Principle and its Applications in e-Security, e-Health and e-Learning*, (Singapore: E.W.S. Publishing), 44.
- Plamondon, R., and Djiova, M. (2006). A multi-level representation paradigm for handwriting stroke generation. *Hum. Mov. Sci.* 25, 586–607. doi: 10.1016/j.humov.2006.07.004
- Plamondon, R., Djiova, M., and Mathieu, P. A. (2013a). Time-dependence between upper arm muscles activity during rapid movements: observation of the proportional effects predicted by the kinematic theory. *Hum. Mov. Sci.* 32, 1026–1039. doi: 10.1016/j.humov.2012.07.006
- Plamondon, R., O'Reilly, C., Rémi, C., and Duval, T. (2013b). The lognormal handwriter: learning, performing, and declining. *Front. Psychol.* 4:945. doi: 10.3389/fpsyg.2013.00945
- Plamondon, R., Feng, C., and Djiova, M. (2008). *The Convergence of a Neuromuscular Impulse Response Towards a Lognormal, from Theory to Practice*. Technical Report No. EPM-RT-2008-08. Montréal: Polytechnique Montréal.
- Plamondon, R., Feng, C., and Woch, A. (2003). A kinematic theory of rapid human movement. Part IV: a formal mathematical proof and new insights. *Biol. Cybern.* 89, 126–138. doi: 10.1007/s00422-003-0407-9
- Qin, J., Lin, J.-H., Faber, G. S., Buchholz, B., and Xu, X. (2014). Upper extremity kinematic and kinetic adaptations during a fatiguing repetitive task. *J. Electromyogr. Kinesiol.* 24, 404–411. doi: 10.1016/j.jelekin.2014.02.001
- Roy, J.-S., Desmeules, F., Frémont, P., Dionne, C. E., and MacDermid, J. C. (2015). *L'ÉVALUATION CLINIQUE, les Traitements et le Retour en Emploi de Travailleurs Souffrant d'atteintes de la Coiffe des Rotateurs - Bilan des Connaissances*. Montréal: IRSST.
- Rozand, V., Grosprêtre, S., Stapley, P. J., and Lepers, R. (2015). Assessment of neuromuscular function using percutaneous electrical nerve stimulation. *J. Vis. Exp.* 103:e52974. doi: 10.3791/52974
- Sant'Ana, J., Franchini, E., da Silva, V., and Diefenthaler, F. (2017). Effect of fatigue on reaction time, response time, performance time, and kick impact in taekwondo roundhouse kick. *Sports Biomech.* 16, 201–209. doi: 10.1080/14763141.2016.1217347
- Sawilowsky, S. S. (2009). New effect size rules of thumb. *J. Mod. Appl. Stat. Methods* 8:26. doi: 10.22237/jmasm/1257035100
- Schindler, R., Bouillon, M., Plamondon, R., and Fischer, A. (2018). “Extending the sigma-lognormal model of the kinematic theory to three dimensions,” in *Proceedings of the International Conference on Pattern Recognition and Artificial Intelligence*, Beijing.
- Sesboüé, B., and Guinestre, J.-Y. (2006). La fatigue musculaire. *Ann. Réadapt. Méd. Phys.* 49, 257–264. doi: 10.1016/j.annrmp.2006.04.021
- Shi, J., Zheng, Y. P., Chen, X., and Huang, Q. H. (2007). Assessment of muscle fatigue using sonomyography: muscle thickness change detected from ultrasound images. *Med. Eng. Phys.* 29, 472–479. doi: 10.1016/j.medengphys.2006.07.004
- Srinivasan, D., and Mathiassen, S. E. (2012). Motor variability in occupational health and performance. *Clin. Biomech.* 27, 979–993. doi: 10.1016/j.clinbiomech.2012.08.007
- Svendsen, S. W., Bonde, J. P., Mathiassen, S. E., Stengaard-Pedersen, K., and Frich, L. H. (2004). Work related shoulder disorders: quantitative exposure-response relations with reference to arm posture. *Occup. Environ. Med.* 61, 844–853. doi: 10.1136/oem.2003.010637
- Tanaka, M., Mizuno, K., Tajima, S., Sasabe, T., and Watanabe, Y. (2009). Central nervous system fatigue alters autonomic nerve activity. *Life Sci.* 84, 235–239. doi: 10.1016/j.lfs.2008.12.004
- Taylor, J. L., and Gandevia, S. C. (2001). Transcranial magnetic stimulation and human muscle fatigue. *Muscle Nerve* 24, 18–29. doi: 10.1002/1097-4598(200101)24:1<aid-mus2>3.0.co;2-d
- Taylor, J. L., Todd, G., and Gandevia, S. C. (2006). Evidence for a supraspinal contribution to human muscle fatigue. *Clin. Exp. Pharmacol. Physiol.* 33, 400–405. doi: 10.1111/j.1440-1681.2006.04363.x
- Tesch, P., Sjödin, B., Thorstensson, A., and Karlsson, J. (1978). Muscle fatigue and its relation to lactate accumulation and LDH activity in man. *Acta Physiol. Scand.* 103, 413–420. doi: 10.1111/j.1748-1716.1978.tb06235.x
- Turpin, N. A., Guevel, A., Durand, S., and Hug, F. (2011). Fatigue-related adaptations in muscle coordination during a cyclic exercise in humans. *J. Exp. Biol.* 214, 3305–3314. doi: 10.1242/jeb.057133
- U.S. Bone and Joint Initiative (2014). *The Burden of Musculoskeletal Diseases in the United States (BMUS)*. Rosemont, IL: OrgtoPaedic Research Society.
- Valero-Cabré, A., Pascual-Leone, A., and Coubard, O. A. (2011). La stimulation magnétique transcrânienne (SMT) dans la recherche fondamentale et clinique en neuroscience. *Rev. Neurol.* 167, 291–316. doi: 10.1016/j.neurol.2010.10.013
- Wang, H. K., and Cochrane, T. (2001). A descriptive epidemiological study of shoulder injury in top level english male volleyball players. *Int. J. Sports Med.* 22, 159–163. doi: 10.1055/s-2001-11346
- Wilk, K. E., Obma, P., Simpson, C. D., Cain, E. L., Dugas, J., and Andrews, J. R. (2009). Shoulder injuries in the overhead athlete. *J. Orthop. Sports Phys. Ther.* 39, 38–54. doi: 10.2519/jospt.2009.2929
- Woch, A., and Plamondon, R. (2001). “Rapid movement analysis with the $\Delta\Delta$ model: towards a better understanding of movement generation,” in *Proceedings of the International Graphonomics Society*, Nijmegen.
- Woch, A., Plamondon, R., and O'Reilly, C. (2011). Kinematic characteristics of bidirectional delta-lognormal primitives in young and older subjects. *Hum. Mov. Sci.* 30, 1–17. doi: 10.1016/j.humov.2009.10.006
- Yang, C., Bouffard, J., Srinivasan, D., Ghayourmanesh, S., Cantú, H., Begon, M., et al. (2018). Changes in movement variability and task performance during a fatiguing repetitive pointing task. *J. Biomech.* 76, 212–219. doi: 10.1016/j.jbiomech.2018.05.025

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

Copyright © 2020 Laurent, Plamondon and Begon. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



A Theoretical Framework for How We Learn Aesthetic Values

Hassan Aleem¹, Ivan Correa-Herran^{2,3} and Norberto M. Grzywacz^{4,5*}

¹ Interdisciplinary Program in Neuroscience, Georgetown University, Washington, DC, United States, ² Department of Neuroscience, Georgetown University, Washington, DC, United States, ³ Facultad de Artes, Universidad Nacional de Colombia, Bogotá, Colombia, ⁴ Department of Psychology, Loyola University Chicago, Chicago, IL, United States, ⁵ Department of Molecular Pharmacology and Neuroscience, Loyola University Chicago, Chicago, IL, United States

How do we come to like the things that we do? Each one of us starts from a relatively similar state at birth, yet we end up with vastly different sets of aesthetic preferences. These preferences go on to define us both as individuals and as members of our cultures. Therefore, it is important to understand how aesthetic preferences form over our lifetimes. This poses a challenging problem: to understand this process, one must account for the many factors at play in the formation of aesthetic values and how these factors influence each other over time. A general framework based on basic neuroscientific principles that can also account for this process is needed. Here, we present such a framework and illustrate it through a model that accounts for the trajectories of aesthetic values over time. Our framework is inspired by meta-analytic data of neuroimaging studies of aesthetic appraisal. This framework incorporates effects of sensory inputs, rewards, and motivational states. Crucially, each one of these effects is probabilistic. We model their interactions under a reinforcement-learning circuitry. Simulations of this model and mathematical analysis of the framework lead to three main findings. First, different people may develop distinct weighing of aesthetic variables because of individual variability in motivation. Second, individuals from different cultures and environments may develop different aesthetic values because of unique sensory inputs and social rewards. Third, because learning is stochastic, stemming from probabilistic sensory inputs, motivations, and rewards, aesthetic values vary in time. These three theoretical findings account for different lines of empirical research. Through our study, we hope to provide a general and unifying framework for understanding the various aspects involved in the formation of aesthetic values over time.

Keywords: aesthetics (as scholarly discipline), reinforcement learning, art, motivation, preference, computational modeling

OPEN ACCESS

Edited by:

Claudio De Stefano,
University of Cassino, Italy

Reviewed by:

Giuseppe Boccignone,
University of Milan, Italy
Tsung-Ren (Tren) Huang,
National Taiwan University, Taiwan

*Correspondence:

Norberto M. Grzywacz
norberto@luc.edu

Specialty section:

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

Received: 25 May 2020

Accepted: 03 August 2020

Published: 11 September 2020

Citation:

Aleem H, Correa-Herran I and
Grzywacz NM (2020) A Theoretical
Framework for How We Learn
Aesthetic Values.
Front. Hum. Neurosci. 14:345.
doi: 10.3389/fnhum.2020.00345

INTRODUCTION

Our aesthetic preferences are an important part of our lives because they shape our decision making and consequently our personality (Skov, 2010). We define ourselves both as individuals and as parts of larger groups through our likes and dislikes (Brown and Dissanayake, 2009). How exactly do these individual preferences come about? Currently, little is known about how preferences form early on and what happens to them throughout our lives. Understanding this process of preference formation has important implications not just for aesthetics, but for

philosophy, psychology, neuroscience, marketing, and many other fields. Indeed, philosophy was likely the first discipline to ponder this question (Sartwell, 2012). Philosophers have long wondered whether beauty is shared (universal) or in the eye of the beholder (individual)? We discuss this philosophical question in greater detail elsewhere but touch on it briefly here to frame our work. In our earlier publication, we argue that one can think of universal aspects of preference as innate, formed due to evolutionary pressures (Aleem et al., 2019). Examples include preferences for round contours, symmetry, and contrast. Such preferences are likely present at birth across all populations (Göksun et al., 2014). Here, we are instead more interested in the learned aspects of aesthetic preferences. Specifically, how do individualized aesthetic preferences form under constraints from our environment and experience? How do existing universal aspects of these preferences undergo individual and context-specific changes?

Aesthetic preferences form early on in life, such that by preschool age, children already show idiosyncrasies of their cultures (Senzaki et al., 2014). Furthermore, these preferences continue to evolve over our lifetimes (Park and Huang, 2010). What mechanisms underlie this lifetime evolution? Many of the existing frameworks of aesthetics do not explicitly consider the time-course of preferences. However, several frameworks stress its importance implicitly by focusing on time sensitive variables such as learning and exposure. For example, Leder et al. stress the importance of familiarity, which has been shown to influence preference over time (Leder et al., 2004). However, the framework proposed by Leder et al. is primarily concerned with understanding the aesthetic experience as it plays out, not how preferences form. A closer perspective comes from Vessel and colleagues, who develop an associative theory of aesthetics (Biederman and Vessel, 2006; Vessel and Rubin, 2010). In their view, aesthetic preferences are shaped by associative experiences over our lifetimes. However, their theory focuses primarily on reasons for shared versus individual tastes, not the dynamics of preferences over time *per se*. Several other theories and empirical findings have implications for temporal aspects of aesthetics (see section “Discussion”), but a framework specifically dedicated to understanding this aspect is so far missing.

What should a framework that aims to understand the dynamics of preferences over-time look like? We have constrained our search for such a framework to be within the general principles of neuroscience. Moreover, we have avoided frameworks in which aesthetic values are formed by specialized mechanisms, but rather have focused on known and existing circuitry. A relevant meta-analysis of neuroimaging studies supports this viewpoint for our framework (Brown et al., 2011). These authors analyzed commonalities of aesthetic appraisal across multiple sensory modalities. The results show generalized mechanisms for appraisal centered around a reward-based learning circuit. The central importance of reward is further bolstered by many other imaging studies of aesthetics and appraisal (Lacey et al., 2011; Vartanian and Skov, 2014; Wang et al., 2015). These studies suggest that a reward-based learning mechanism, likely, reinforcement learning, is fundamental to any framework for understanding how aesthetic preferences form.

However, these studies and the results from the meta-analysis by Brown et al. (2011), suggest that many factors influence this process of reward-based learning. For example, these factors include interoceptive inputs such as motivations and exteroceptive inputs such as the statistics of sensory stimuli (Brown et al., 2011).

How can a reasonable mechanism of reinforcement learning account for aesthetic individuality? Our individual motivations can greatly influence how we interact with the environment and what decisions we make, thereby having a direct effect on our preferences (Nelson and Morrison, 2005). Motivations can influence reward, for example, activation in reward related regions in the brain in response to certain foods is greatly modulated by food specific satiation (Howard and Kahnt, 2017). Similarly, Brown et al., consider internal drives, or motivations as a key factor in aesthetic appraisal. Since such motivations are individual, they can help account for individuality (Silvia et al., 2009). Therefore, this suggests that motivation may modulate learning of aesthetic preferences. Next, we consider an important factor in aesthetic learning not explicitly addressed by Brown et al., that is, the statistical nature of inputs. Evidence suggests that our perception of incoming inputs is statistical in nature (Pouget et al., 2013). Sensory inputs show many statistical properties that convey useful information which can influence aesthetic preferences. For example, preference for facial symmetry has been shown to be modulated by the presence of pathogen cues (Little et al., 2011). This statistical nature also accounts for differences in preferences amongst cultures, as they impose contingencies on rewards through value systems (Park and Huang, 2010). Finally, internal states such as motivation are also statistical, because we act according to states that vary across time (for example, hunger, tiredness, and sex drive) (Craig, 2009). Sensory inputs, rewards, and motivation do not form a comprehensive list, since a theoretical framework for the learning of aesthetic values is not complete without accounting for semantics, expertise, and much more. However, our framework may capture some of the essential components of aesthetic learning and thus, helps us focus on a simpler model that raises testable predictions. By keeping the model general, we leave ample room for further modifications and increase in complexity.

Following the guidelines listed above, we developed a theoretical framework and a related computational model to investigate the formation and dynamics of aesthetic preferences. The model focuses on visual aesthetic preferences, but our interests go beyond vision or art *per se*. Instead, we are interested in a theoretical framework that is general to the many different domains of preferences. A detailed description of our framework is presented in section “Theoretical Framework.” In turn, section “Materials and Methods” develops the model and described methods for computer simulations of this model. We used these simulations and mathematical analyses to test the following questions: First, we investigated whether aesthetic values would show a dynamic time course, possibly with multiple stages. Moreover, we considered whether these values would be stochastic due to the probabilistic nature of the inputs. Second, we explored how the contingent probabilities of the different variables could lead to a segregation of different

trajectories, possibly mimicking different cultures. Third, we investigated whether the individual specific motivation variable would lead to further partitioning of learning trajectories, leading to individuality.

THEORETICAL FRAMEWORK

We have split the description of the theoretical framework into two subsections, general and mathematical. The general section has a description of the ideas without any equations. Our goal here is to help the reader understand the elements of the theoretical framework at an intuitive level. In turn, the mathematical section lays out the equations used to specify the framework precisely. The general description section may allow some readers to skip the equations and go directly to the Results.

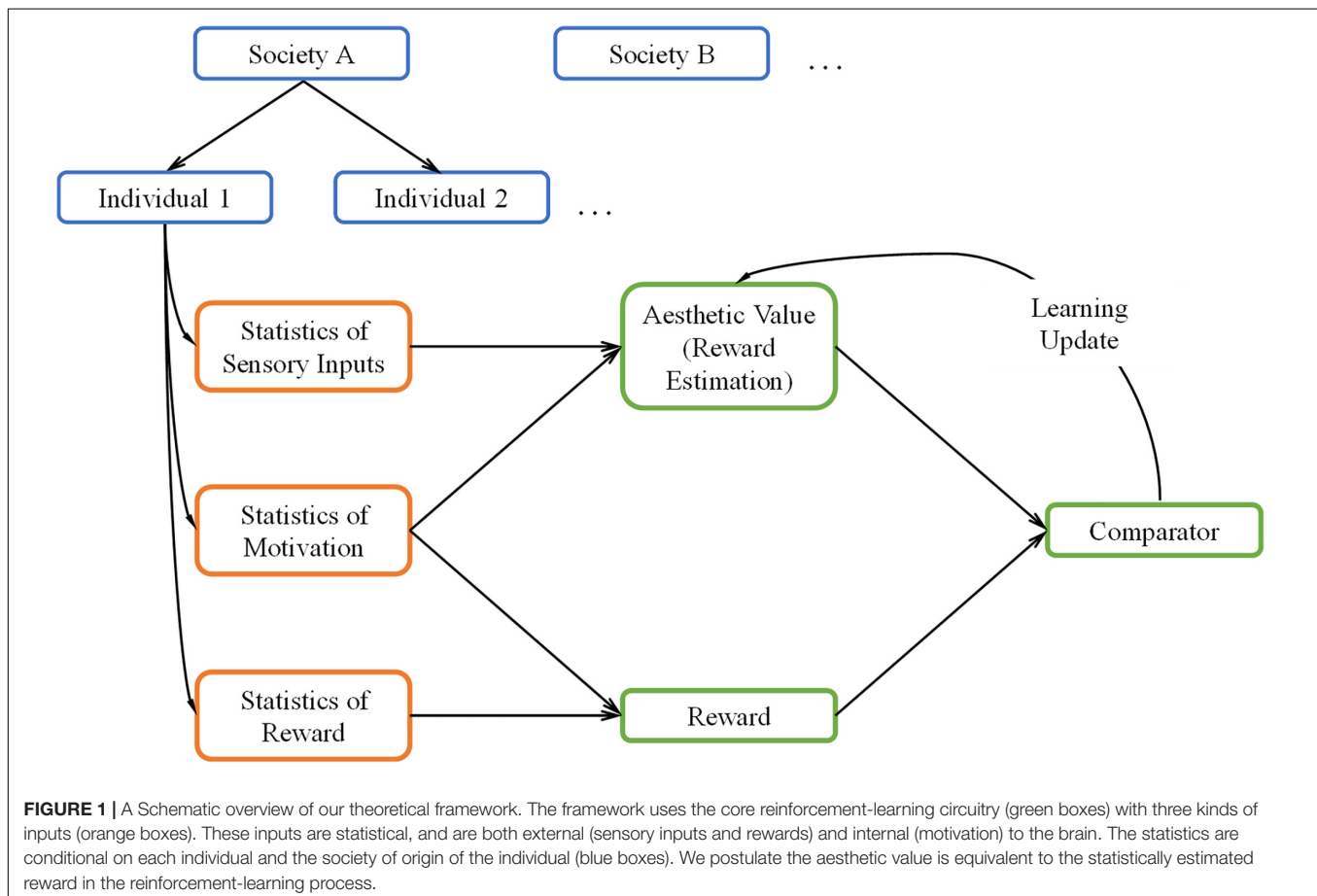
General Description of the Theoretical Framework

A general overview of our theoretical framework can be seen in **Figure 1**. The green boxes in this figure illustrate the core reinforcement-learning system. We discuss it briefly here but see Sutton and Barto (2018) for an extensive overview. In typical reinforcement learning, the system first receives inputs from the external world and from the body (orange boxes). The system

then uses these inputs to form an internal model to estimate the reward when taking some action (“Reward Estimation” box), commonly referred to as value. When rewards arrive (“Reward” box), they are compared with the estimated reward (“Comparator” box). If there is a mismatch, the system “learns” by updating the parameters of the internal model. This update allows the system to achieve its goal of producing better reward predictions in the future.

While we based our aesthetic-learning theoretical framework largely on reinforcement learning, our framework has four notable extensions, which make it noteworthy:

First, we propose that the estimate of reward is equivalent to aesthetic value. To help understand this proposal, consider the following example. Imagine a person looking at an apple and smelling it, trying to decide whether to eat it. From the information that the sensory systems collect from the apple, the person makes a prediction about the rewards gained by eating the apple, for example, how sweet and nutritive it is. Then, if the person eats it, their brain will compare actual rewards and its predictions, in updating its model of apples if necessary. Hence, their brain learns that certain statistical properties of apples, for example, shape or color can inform the prediction on their value, which guides the preference for them. Now imagine that the same individual gazes at a painting of an apple. Since some of the same statistical signals may be present in the painting,



a similar prediction of reward, or value, is still generated. We propose that the previously learned value will still influence the experience of viewing the painting. If the statistics were previously rewarding, the painting will also elicit high value and generally elicit preference to a similar degree. Therefore, the previous learned value is converted into an aesthetic value.

Second, we incorporate the concept of motivation within the reinforcement-learning circuitry. Motivation, by our definition, is somewhat akin to *policy* (Averbeck and Costa, 2017). It refers to the internal drive of an individual, representing their likelihood to act given an input. For example, if an individual is not hungry, this person will not have the motivation to try a certain food and therefore not learn about its value. However, motivation is not limited to acts of consumption. Experimental interventions can increase the “need” for abstract concepts such as complexity or cognitive closure (Tinio and Leder, 2009; Steciuch et al., 2019). Furthermore, one can choose whether to engage mentally with a work of art if they are motivated. In our framework, motivation is probabilistic, varying from one moment to another, a behavior generally reflecting findings of interoceptive states in humans (Craig, 2009). How does motivation ultimately affect learning? In the simplest manner, motivation controls the rate of learning by slowing it down or accelerating it when the motivation is low or high, respectively. More generally, motivation may affect the learning of certain aesthetic values. For example, an individual who rarely eats fruits may not learn a high aesthetic value for inputs related to fruit.

Third, both interoceptive and exteroceptive inputs to the theoretical framework are statistical (orange/blue boxes in **Figure 1**). The statistical distributions of these inputs should have significant effects on the learning of values. In detail, statistical interoceptive inputs reflect the variation of motivations across individuals and over time in a single individual (blue boxes). On the other hand, the exteroceptive inputs correspond to both rewards and sensory signals (orange boxes). The statistical nature of these signals reflects the variations of the external world and how different individuals experience it (blue boxes). Continuing our example above, an apple’s sensory signals would be about smell, shape, color, and taste, the reward would be about calories and vitamins, while the interoceptive signals would be the appetite to eat it. The statistical nature of our theoretical framework has three important implications: One, it allows us to generalize over variations of individuals and objects. Two, it ensures that the interoceptive and exteroceptive signals vary stochastically over time, which better approximates real-world conditions. Three, in real life, the statistics of rewards and sensory signals are often correlated, for example, the color and shape of an apple is an indicator of its ripeness. In turn, these attributes influence internal states, for example, a ripe looking apple will be more appetizing. We can model these relationships by using probabilistic distributions for these variables.

Fourth, the inputs to our theoretical framework (orange boxes) depend not only on individuals but also across societies (blue boxes). Therefore, we propose the existence of a parameter space whose values are different across groups (nations, cultures, societies). This means that the distributions of individual statistics are largely conditional on these external parameters.

These parameters may specify ecological differences, for example, differences in climates, genetic predispositions, or exposure to diseases (Little et al., 2007; Sorokowski et al., 2014). The social parameters may also specify cultures values, for example, different rewards for certain colors or styles (Masuda et al., 2008; Park and Huang, 2010). By setting the model in the context of social and environmental backgrounds, we can approximate how different societies and cultures form distinct aesthetic values.

In sum, our model begins from a basic circuitry of reward-based learning. Inspired by empirical findings, we expand on this circuitry to include probabilistic inputs, internal drives, and other external contingencies. The combination of these factors allows our model to account for a range of phenomena from the societal all the way to the individual level.

Mathematical Description of the Theoretical Framework

Let the sensory inputs be N dimensional, with the various components corresponding to variables that the brain uses to represent the external world:

$$\vec{u}(t) = [u_1(t), u_2(t), \dots, u_N(t)]$$

where the overhead arrow indicates a vector, and t indicates that sensory inputs vary (stochastically) over time.

In this paper, we assume that the model used for estimating reward is linear. Although this assumption is common in reinforcement-learning models (Dayan and Abbott, 2001), it is not necessary. We make this assumption here for the sake of simplicity, but address the consequences in the Discussion. The assumption means that a parameter vector:

$$\vec{w}(t) = [w_1(t), w_2(t), \dots, w_N(t)]$$

exists such that the estimated reward is:

$$v(t) = m(t) \vec{w}(t) \cdot \vec{u}(t) \quad (1)$$

where $0 \leq m(t) \leq 1$ is the motivation function. This equation is important, because learning occurs in the presence of actual rewards by adjusting the w ’s. The introduction of the m function is a modification of standard reinforcement-learning models, which would use Eq. 1 with $m = 1$. This modification is necessary, since people only get rewards if they act. Thus, if we interpret m as the probability of acting, then the received reward is:

$$r(t) = m(t) r^*(t) \quad (2)$$

where r^* is the reward that a fully motivated person would get. The presence of m in the reward estimate (Eq. 1) considers that reward itself varies with motivation (Eq. 2).

The typical learning in reinforcement-learning theories follows the precept of temporal difference (Dayan and Abbott, 2001)

$$\delta(t) = r(t) - v(t) \quad (3)$$

$$\frac{d\vec{w}(t)}{dt} = k\delta(t) \vec{u}(t) \quad (4)$$

where $k > 0$ is a constant. Equation 4 is a continuous version of the delta rule and thus, tends to minimize the difference between v and r . Therefore, this minimization makes the estimated reward as close as possible to the real one. An important property of this equation is that because motivation affects both v and r (Eqs. 1 and 2), it affects learning through the delta (Eq. 3). This is important, since it shows that with no motivation, learning freezes. The freezing makes sense, since for example, if a person is not hungry, then the person will not eat. Thus, the person cannot learn if the estimated reward is large or small in regards to that specific food.

To complete the theoretical framework, we need to specify the statistical properties of \vec{u} , m , and r^* . Let us begin by considering \vec{u} and r^* . Both these variables are exteroceptive signals but with different origins. While \vec{u} is sensory (for example, seeing and smelling an apple) and used for estimating reward, r^* arises from the action (for example, eating the apple). As explained in section “General Description of the Theoretical Framework,” these variables are dependent, possibly exhibiting correlation. In a sense, the model acquires this correlation, using it to predict r^* from \vec{u} . Thus, we are interested in the probability density functions:

$$P(\vec{I}_u | \vec{B}), P((\vec{u}(t), r^*(t)) | \vec{I}_u) \quad (5)$$

where \vec{B} indicates the vector of parameters characteristic of the social and environmental background under consideration and \vec{I}_u is the vector of parameters of an individual in this society. Consequently, the first probability density function in Eq. 5 is the probability of finding an individual, while the second gives the rewards and sensory inputs that this individual gets over time.

Finally, we must specify the statistical properties of m . Because it represents interoceptive signals related to motivation, m depends on each individual. However, motivation also depends on the sensory input \vec{u} . If, say, a person is hungry, but the sensory input is not food, then the individual will not have a motivation to act, that is, to eat. But if by changing the gaze, the sensory input is changed to an appetizing food, the person will be motivated to act. We thus write the probability density function of m as:

$$P(\vec{I}_m | \vec{B}), P((m(t)) | \vec{u}(t), \vec{I}_m) \quad (6)$$

where we insert \vec{B} to indicate that individual motivation may depend on environmental and social backgrounds. For example, the motivation to smoke is prevalent in some societies but not others (Dechesne et al., 2013). As for Eq. 5, the first probability density function in Eq. 6 is the probability of finding an individual, while the second gives the motivations that this individual has over time.

MATERIALS AND METHODS

We studied the implications of our theoretical framework through mathematical analyses and computer simulations. In section “Methods for Computer Simulations,” we describe the mathematical details of simulating the model, with steps to simplify the procedure. Next, in section “Illustrative Model,” we describe the properties of the illustrative model used in this

paper, listing each component and its technical rationale. We then describe the algorithm to simulate the model in section “Summary of the Simulation Procedures.” Finally, we describe the parameters used in the standard simulation in section “Standard Simulation Parameters.” For those readers who do not have a mathematical background, we suggest first reading section “Summary of the Simulation Procedures.” That section may help get an overall understanding before reading the other sections for details.

Methods for Computer Simulations

We must simulate Eqs. 1–4. Combining these equations, we get:

$$\frac{d\vec{w}(t)}{dt} = km(t)(r^*(t) - \vec{w}(t) \cdot \vec{u}(t))\vec{u}(t) \quad (7)$$

This is a stochastic differential equation, because the \vec{u} , m , and r^* come from samples of the probability distributions in Eqs. 5 and 6.

We simplify our simulations through a mean field approximation of Eq. 6:

$$\frac{d\vec{w}(t)}{dt} = k\bar{m}(\vec{u}(t) : \vec{I}_m)(r^*(t) - \vec{w}(t) \cdot \vec{u}(t))\vec{u}(t) \quad (8)$$

where $\bar{m}(\vec{u}(t) : \vec{I}_m)$ is the mean motivation as a function of the sensory input $\vec{u}(t)$ and parametric on \vec{I}_m . The advantage of the approximation in Eq. 8 is that we do not simulate the noise in the motivation states, but only their deterministic dependence on the sensory inputs. Nevertheless, the motivation will remain stochastic, because so are the sensory inputs.

To approximate a solution to Eq. 8, we must discretize time and sample \vec{u} , m , and r^* for every t . We do this discretization as follows:

$$\begin{aligned} \vec{w}(t_{k+1}) &= \vec{w}(t_k) + \epsilon \bar{m}(\vec{u}(t_{k+1}) : \vec{I}_m) \\ &\quad (r^*(t_{k+1}) - \vec{w}(t_k) \cdot \vec{u}(t_{k+1}))\vec{u}(t_{k+1}) \end{aligned} \quad (9)$$

where $\epsilon = k(t_{k+1} - t_k)$, with $t_{k+1} - t_k$ being constant (for $k = 0, 1, 2, \dots$).

Illustrative Model

We performed computer simulations using an illustrative model developed from our theoretical framework. Although this model is just illustrative, we point out in the Results outcomes of mathematical analyses showing that the most important conclusions of the model simulations are general. We also address the generality of the simulation results in the Discussion. In this section, we specify the illustrative model used in the simulations. Because this section is highly technical, we provide a summary of the model with figures in the next section (section “Summary of the Simulation Procedures”). Section “Standard Simulation Parameters” describes the standard parameter set used in the model simulations.

To specify a model, we need to provide the probability functions in Eq. 5, the $P(\vec{I}_m | \vec{B})$ function in Eq. 6, and the \bar{m} function in Eq. 8. To begin, we took five steps to simplify the model to make the simulations fast:

- A. We did not simulate social noise by implementing explicitly $P(\vec{I}_u|\vec{B})$ and $P(\vec{I}_m|\vec{B})$. Instead, we set individual parameters by hand, changing them for different individuals to study the parametric dependence of the model:
- B. We split the individual parameters \vec{I}_u into sensory related (\vec{I}_s) and reward related (\vec{I}_r):

$$\vec{I}_u = [\vec{I}_s, \vec{I}_r] \quad (10)$$

Thus, we divided the \vec{I}_u parameters into lower-dimensional ones that separately control the samplings of \vec{u} and r^* .

- C. We made \vec{u} two-dimensional. One component was visual balance (u_b) and the other was visual complexity (u_c), making:

$$\vec{u} = [u_b, u_c]$$

where $0 \leq u_b, u_c \leq 1$, as per the definitions in (Aleem et al., 2017).

While our model is amenable to a range of sensory inputs, we simplified it to two variables in the visual domain for illustrative purposes. We briefly describe these two variables here, but refer the reader to existing literature to gain a deeper understanding. The component of visual balance but can best be surmised as an equal amount of visual weight across an image, often measured by pixel intensities (Wilson and Chatterjee, 2005). Visual complexity, on the other hand, can be best described as the amount of information in an image, for example the range of pixel intensities present (Donderi, 2006). An important interaction exists between these two variables in that they are generally negatively correlated. For example, as an image becomes more balanced (organized), its complexity generally decreases (Aleem et al., 2017). We explore this relationship in our experiments to test whether the two variables compete and influence learning.

- D. We split the second term of Eq. 5 into:

$$P((\vec{u}, r^*)|\vec{I}_u) = P(u_b, u_c|\vec{I}_s) P(r^*|u_b, u_c, \vec{I}_r) \quad (11)$$

Thus, instead of sampling directly from a relatively complex three-dimensional space (two for \vec{u} and one for r^*), we split the problem. We first sampled from a simpler two-dimensional space and then used the outcome to condition the sampling of a one-dimensional space. The splitting of probabilities in these steps greatly reduces the computation time required for sampling. However, having separate sensory and reward parameters in the two probability distributions increases the degrees of freedom, potentially leading to a wider range of observable behaviors. For example, the sensory and reward functions could be varied independently.

- E. To model the various variables in our simulations, we assumed they had Gaussian distributions. While natural scene statistics and neural-reward related processes can have a multitude of probability distributions (Field, 1994), one can often approximate them with Gaussian processes (Wainwright and Simoncelli, 2000; Dabney et al., 2020).

Hence, using Gaussian distributions here allowed us to explore the theory from a parsimonious viewpoint. In addition, sampling Gaussian distributions is fast, because of the abundance of code available for this purpose. However, future iterations could benefit from employing other distributions.

We modeled the first term of the right-hand side of Eq. 11 with a truncated bivariate Gaussian distribution (Rosenbaum, 1961),

$$P(u_b, u_c|\vec{I}_s) = \text{Tr}(G_2(u_b, u_c : \mu_b, \mu_c, \Sigma)) \quad (12)$$

where G_2 is the Gaussian over the variables u_b and u_c , with means $\vec{\mu} = [\mu_b, \mu_c]$ and covariance matrix:

$$\Sigma = \begin{bmatrix} \sigma_{u_b}^2 & \rho\sigma_{u_b}\sigma_{u_c} \\ \rho\sigma_{u_b}\sigma_{u_c} & \sigma_{u_c}^2 \end{bmatrix}$$

where σ_{u_b} and σ_{u_c} are standard deviations in the u_b and u_c directions respectively, and ρ is the correlation between u_b and u_c . In turn, the truncation function $\text{Tr}(G_2)$ is:

$$\text{Tr}(G(x, y)) = \frac{1}{\int_0^1 \int_0^1 G_2(x, y) dx dy} \begin{cases} G_2(x, y) & \text{if } 0 \leq x, y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

With these definitions for the first term of the right-hand side of Eq. 11, the individual sensory-parameter vector is therefore,

$$\vec{I}_s = [\mu_b, \mu_c, \sigma_{u_b}, \sigma_{u_c}, \rho] \quad (13)$$

To model rewards associated with the sensory variables, we assumed independent contributions of rewards from balance (r_b^*) and complexity (r_c^*), and then summed these contributions, that is,

$$r^* = r_b^* + r_c^* \quad (14)$$

Hence, if we have $P(r_b^*|u_b, \vec{I}_r)$ and $P(r_c^*|u_c, \vec{I}_r)$, then we can calculate $P(r^*|u_b, u_c, \vec{I}_r)$ as:

$$P(r^*|u_b, u_c, \vec{I}_r) = \int_{-\infty}^{\infty} P(r_b^*|u_b, \vec{I}_r) P(r_c^* = r^* - r_b^*|u_c, \vec{I}_r) dr_b^*$$

Consequently, all that remains to do to specify the second right-hand term of Eq. 11 is to define $P(r_b^*|u_b, \vec{I}_r)$ and $P(r_c^*|u_c, \vec{I}_r)$.

To start with the probability density function $P(r_b^*|u_b, \vec{I}_r)$, balance was positively related to reward (Wilson and Chatterjee, 2005). In the simplest mathematical form, balance and reward would obey a linear relationship. We thus define:

$$P(r_b^*|u_b, \vec{I}_r) = G_1(r_b^* : -\alpha + 2\alpha u_b, \sigma_{r_b^*}) \quad (15)$$

where G_1 is the univariate Gaussian distribution over the variable r_b^* , and $\alpha, \sigma_{r_b^*} > 0$ are parameters. The mean of the Gaussian is $-\alpha + 2\alpha u_b$ and the standard deviation is $\sigma_{r_b^*}$. The mean is such that the integral of $-\alpha + 2\alpha u_b$ over the range of $u_b (0 \leq u_b \leq 1)$ is zero. Positive and negative rewards occur in equal amounts.

We next define $P(r_c^*|u_c, \vec{I}_r)$. Several studies have shown that the preference for complexity displays an inverted U-curve behavior, that is, people like moderate amounts of complexity

more than they do little or much complexity (Berlyne, 1971; Donderi, 2006; Güçlütürk et al., 2016). A simple form for the relationship between reward and complexity is a Gaussian shape. We thus define:

$$P(r_c^* | u_c, \vec{I}_r) = G_1 \left(r_c^* : \phi(\beta, \gamma, \theta) + \beta e^{-\frac{(u_c - \gamma)^2}{2\theta^2}}, \sigma_{r_c^*} \right) \quad (16)$$

where G_1 is now over the variable r_c^* . The parameters are β , $\sigma_{r_c^*} > 0$, $0 \leq \gamma \leq 1$, and θ , while $\phi(\beta, \gamma, \theta)$ is a function of them. The integral of $\phi + \beta \exp(-(u_c - \gamma)^2 / (2\theta^2))$ over the range of u_c ($0 \leq u_c \leq 1$) is zero. Because $\phi + \beta \exp(-(u_c - \gamma)^2 / (2\theta^2))$ is the mean of the Gaussian and $\sigma_{r_c^*}$ is the standard deviation, we have the same amount of positive and negative rewards.

With the definitions in Eqs. 15 and 16, the individual reward parameter vector is therefore,

$$\vec{I}_r = [\alpha, \sigma_{r_b^*}, \beta, \gamma, \theta, \sigma_{r_c^*}] \quad (17)$$

Finally, we defined the motivation function in Eq. 8, namely, $\vec{m}(\vec{u}(t) : \vec{I}_m)$. For the sake of simplicity and illustration, we modeled \vec{m} as independent of u_b . As for the dependence on u_c , we consider different individuals with different peak preferences in terms of complexity. We also use the Gaussian shape to model this peak:

$$\vec{m}(\vec{u}(t) : \vec{I}_m) = m_{min} + (m_{max} - m_{min}) e^{-\frac{(u_c - \mu_m)^2}{2\sigma_m^2}} \quad (18)$$

where $0 \leq m_{min}, m_{max}, \mu_m \leq 1$ and σ_m are parameters. The parameters m_{min} and m_{max} are the minimal and maximal motivations respectively. In turn, μ_m is the complexity yielding maximal motivation and σ_m controls how quickly motivation falls as u_c moves away from μ_m . With Eq. 18, the individual motivation parameter vector is:

$$\vec{I}_m = [m_{min}, m_{max}, \mu_m, \sigma_m] \quad (19)$$

Summary of the Simulation Procedures

The simulations proceed with the following algorithm:

- Suppose that at time t_k the weights are $\vec{w}(t_k)$.
- Sample sensory inputs, $\vec{u}(t_{k+1}) = [u_b(t_{k+1}), u_c(t_{k+1})]$ from Eq. 12.
- Sample reward for balance, $r_b^*(t_{k+1})$ from Eq. 15.
- Sample reward for complexity, $r_c^*(t_{k+1})$ from Eq. 16.
- Compute overall reward, $r^*(t_{k+1})$ from Eq. 14.
- Compute motivation, $\vec{m}(\vec{u}(t_{k+1}) : \vec{I}_m)$ from Eq. 18.
- Compute updated aesthetic weights, $\vec{w}(t_{k+1})$ from Eq. 9.
- Start the process again at Step a, but at time t_{k+1} .

An example of 30,000 samples of the sensory inputs from Step b in a typical simulation appears in **Figure 2**. **Figure 2A** illustrates that balance and complexity exhibit negative correlation. **Figures 2B,C** show typical examples of the distributions used for the samples in steps c and d, respectively. In our model, reward tends to increase linearly with balance, except for the probabilistic distribution of rewards (Wilson and Chatterjee, 2005). Probabilistic fluctuations also affect the

dependence of reward on complexity, but the general trend is that of an inverted U-curve behavior (**Figure 2C**; Donderi, 2006). Importantly, the distributions in **Figures 2B,C** illustrate that rewards can be both positive and negative. In these figures and in our simulations, positive and negative rewards are balanced, summing to zero. Finally, **Figure 2D** illustrates the typical shape of the motivation function in Step f. The illustration superimposes color-coded magnitudes of motivation on samples of sensory inputs as in **Figure 2A**. In our illustrative model, motivation only depends on complexity and has a peak at a particular magnitude of complexity. The peak complexity is distinct for different individuals (not shown in **Figure 2D**). One may associate individuals with motivations for higher complexity with risk-taking, because high complexity tends to present more uncertainties, at the possible benefit of more information (Furnham and Bunyan, 1988). Similarly, motivations for low complexity may be associated with risk aversion.

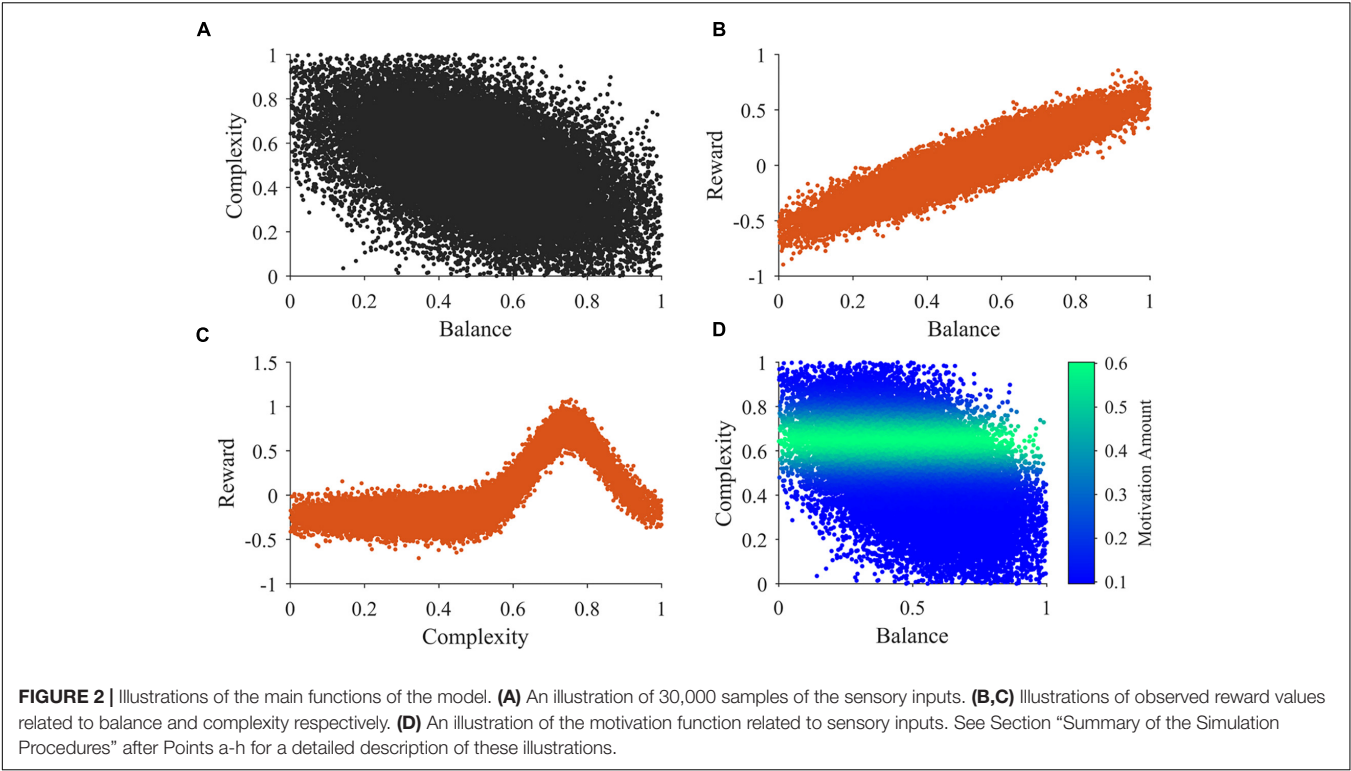
All simulations were performed with code specially written in MATLAB R2019b (MathWorks, Natick, Ma, United States). This code is available in an online repository (**Supplementary Materials, Section A**).

Standard Simulation Parameters

In this paper, we report on simulations with different parameter sets to explore the model. We have designated one of these sets as our standard set (see **Table 1**), because the corresponding results capture the data in the literature reasonably well. We also show simulations with other parameter sets to illustrate individual differences and analyze the various behaviors of the model. The table below shows the parameters of the standard simulations. Parameters for other simulations are indicated as appropriate in the Results.

RESULTS

The following sections outline the results of our simulations and the mathematical analyses. In our first experiments (sections “Learning Dynamics of Aesthetic Weights” and “Understanding the Fast and Slow Phases of Learning”), we looked at the time course of how aesthetic values form by looking at the learned weights. We were particularly interested to see if there were multiple phases (section “Learning Dynamics of Aesthetic Weights”). We found this to be the case, thus in section “Understanding the Fast and Slow Phases of Learning,” we investigated the reasons behind this and found it has to do with the shape of the function linking error between actual and predicted rewards to balance and complexity. The results of our first two experiments also showed that the weights for balance and complexity diverged, indicating an apparent competition. We explored the reasons for this apparent competition by varying different aspects of our model (section “Understanding Apparent Competition between Aesthetic Weights”). We found that motivation was a key component of this apparent competition. Therefore, in the next experiment, we further explored the role of motivation (section “The Role of Motivation and Reward on Aesthetic Individuality”). We found that differing motivation



functions can profoundly change the aesthetic weights learned. We then compared this finding to that obtained with differing social reward contingencies and saw a similar effect on aesthetic weights. Finally, we explored the landscape of the learned aesthetic values as a function of complexity and balance (section “Beauty and the Emergence of the Peak-shift Effect”). We discovered that certain regions of the sensory space had higher learned values than average. We hypothesized that this landscape might explain the value-exaggeration effect often observed in art.

Learning Dynamics of Aesthetic Weights

If aesthetic values are learned, then their corresponding aesthetic weights change over time. Ideally, their dynamics would be so that the values, i.e., the predicted rewards would approach the actual rewards as much as possible. However, weights are not free to change arbitrarily. They may exhibit interdependencies (e.g., **Figure 2A**), and have different dependences on rewards and motivations (**Figures 2B–D**). We performed multiple computer simulations to gain an understanding of the dynamics of aesthetic weights. An example with the model described in section “Illustrative Model” and the standard parameters (**Table 1**) appears in **Figure 3**.

The simulations in **Figure 3A** show an example of the dynamics of the aesthetic weights for balance and complexity. The weights start at [0,0], i.e., they reflect a hypothetical individual who knows nothing about the importance of balance and complexity at the initial point of learning (see section “Discussion”). These weights then rise quickly in an initial fast phase and then slow down in a divergent phase. In the initial phase, both balance and complexity weights rise equally

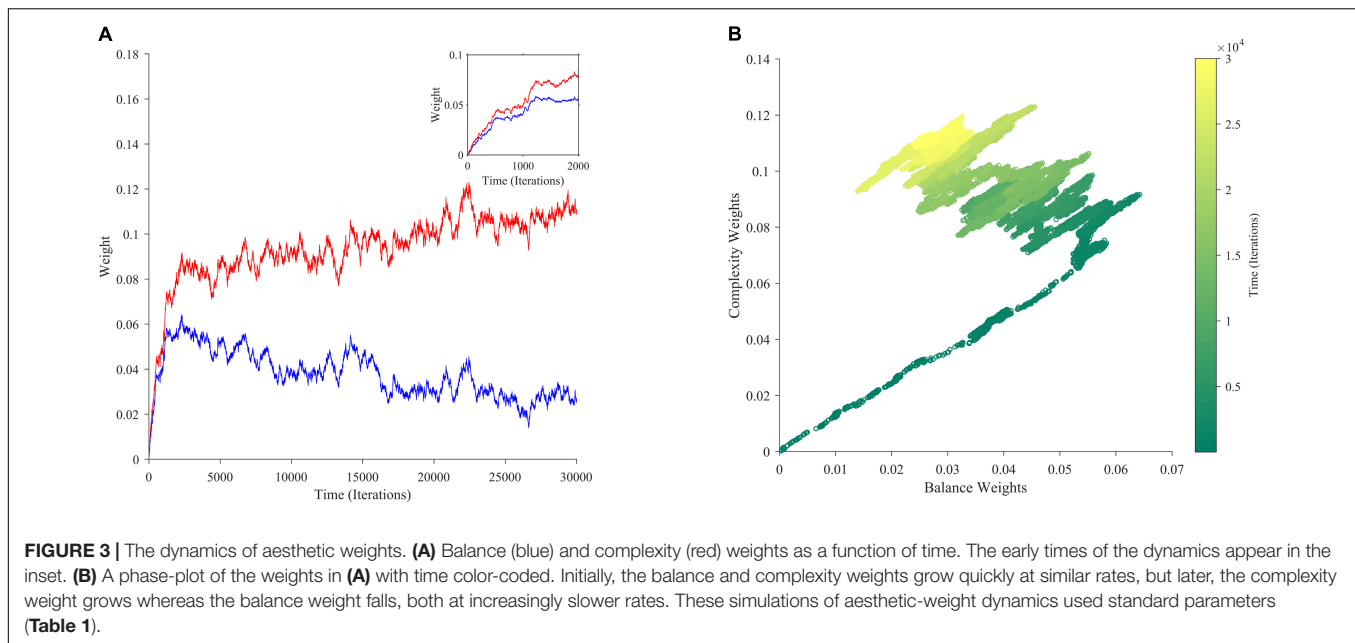
in relation to each other (**Figure 3A** inset). However, after this phase, an inflection point occurs. In the new phase, the complexity weight continues to rise while the balance weight drops, as if they are competing. Thus, these weights reach a state of slow divergence. As time increases, both weights appear to arrive to a stochastic equilibrium in relation to each other, with their separation increasing at a slow pace.

A phase-phase plot is especially helpful to visualize the learning dynamics (**Figure 3B**). Such a plot graphs the complexity weight as a function of the balance weight, color-coding for time. As the inset of **Figure 3A** shows, the rise of balance and complexity in the initial phase is tightly correlated, indicated by the linear slope in the phase plot. However, after the inflection point, a much slower drift can be seen through the formation of a cloud region. The dynamic moves slowly toward greater complexity and lower balance, eventually forming a relatively stable stochastic cloud.

Why does this stable cloud form in the phase plot? A simple hypothesis would be that the weights gravitate around a

TABLE 1 | Standard set of parameters.

Parameter(s)	Equation	Values
$\tilde{w}(t_0)$	9	[0,0]
ϵ	9	0.01
$t_{k+1} - t_k$	9	1
$\tilde{l}_s = [\mu_b, \mu_c, \sigma_{ub}, \sigma_{uc}, \rho]$	13	[0.5,0.5,0.2,0.2,-0.5]
$\tilde{l}_r = [\alpha, \sigma_{r_b}^2, \beta, \gamma, \theta, \sigma_{r_c}^2]$	17	[0.6,0.1,1,0.75,0.1,0.1]
$\tilde{l}_m = [m_{min}, m_{max}, \mu_m, \sigma_m]$	19	[0.1,0.6,0.65,0.1]



fixed point, not converging to it just because of the stochastic nature of our model. Our mathematical analyses show a more complex and interesting picture on the outcome of learning than this hypothesis suggests. The learning process leads to a gradient-descent-like optimization of the prediction of reward (**Supplementary Materials, Section B**). Specifically, value approaches reward as much as statistically possible as follows: If for every τ there is a $t > \tau$ such that $m(t) > 0$, then the learning process minimizes:

$$E(\vec{w}) = \left\langle \frac{(r(t) - v(t))^2}{m(t)} \right\rangle_t \quad (20)$$

where $\langle \rangle_t$ stands for time average. The minimization of $E(\vec{w})$ with respect to the components of \vec{w} in Eq. 20 implies that $v(t)$ tends to become statistically close to $r(t)$. The near optimization of value in terms of estimating reward as predicted by Eq. 20 is confirmed by our computer simulations (**Figure 4A**). However, $v(t)$ does not converge exactly to $r(t)$ because of two reasons: First, the theoretical framework is stochastic. If the process were not stochastic, then the value would converge exactly to the reward. Second, the optimization of Eq. 20 is modulated by the statistics of $m(t)$, \vec{u} , and r^* .

However, the mathematical analysis also shows that although value tends to gravitate around a fixed point, the weights do not necessarily do so (**Supplementary Materials, Section C**). Different sets of weights can produce the same value. To be more precise, we can define the following hyperplane in terms of weights:

$$\sum_{i=1}^N a_i(t) w_i(t) = v(t) \quad (21)$$

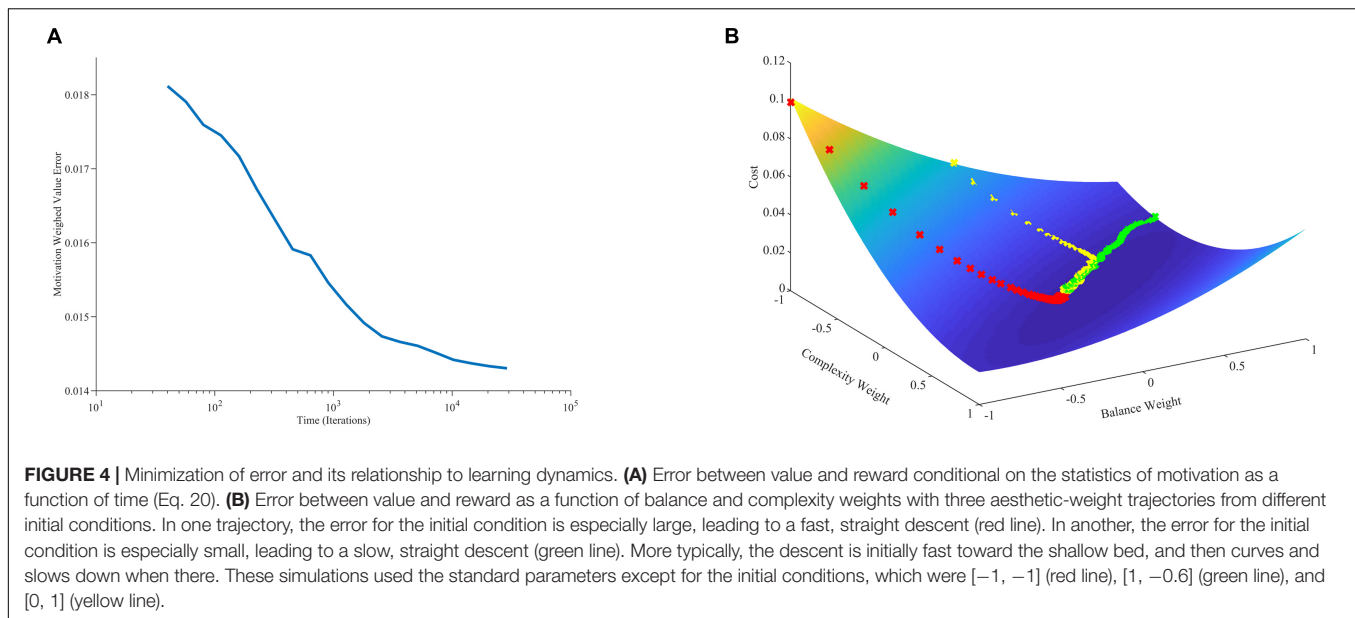
where they $a_i(t)$ are:

$$a_i(t) = m(t) u_i(t) \quad (22)$$

such any point in this hyperplane is compatible with the value $v(t)$. Because of this redundancy, the exact $\vec{w}(t)$ are not always meaningful. Is having such a redundancy in weight representation wasteful? The mathematical analysis shows that this redundancy in weights is not arbitrary, but allows the improvement of the learning rate (**Supplementary Materials, Section C**). Mathematically, the weights $\vec{w}(t)$ aim to reach the nearest point of the ideal hyperplane in a way that is dependent on their initial conditions. Consequently, because of the stochastic nature of the theoretical framework, the $\vec{w}(t)$ can drift even if the value stays close to reward (**Supplementary Materials, Section C**). With each new sample of $\vec{u}(t)$, $m(t)$, and $r^*(t)$, the $\vec{w}(t)$ simply pushes value toward the new hyperplane defined by this sample. Thus, $\vec{w}(t)$ may not return to past positions, possibly drifting according to a random-walk-like trajectory.

Understanding the Fast and Slow Phases of Learning

What are the underlying reasons for the fast and slow phases of learning observed in **Figure 3**? Considering that values follow a gradient descent (section “Learning Dynamics of Aesthetic Weights”), we look toward the error between value and reward for an answer. As **Figure 4B** shows, the error function has a hammock-like shape when plotted against balance and complexity weights. Consequently, the error function varies rapidly along one direction and slowly along its perpendicular. This shape leads to differences in gradients across regions of the function. Thus, if the aesthetic weights start at a point with an especially large error, they will experience a large gradient, descending fast toward the minimum of the function (red line in **Figure 4B**). If instead they start at a point with an especially small error, they will descend slowly toward the minimum (green line). Ultimately, as the weights approach the minimum of



the error function, gradients get smaller and the convergence becomes more stochastic. Thus, weights become more sensitive to variations of sensory inputs, rewards, and motivations. When gradients are steep, weights tend to move to reduce the error rapidly even in presence of input variations. Such initial fast approach is consistent with the fast learning-rate explained in section “Learning Dynamics of Aesthetic Weights.” Hence, the initial conditions may dictate an initial fast approach to the shallow bed, and a more slowly stochastic dependence once there. The direction of gradient descent in the shallow bed is typically different from in the initial fast phase. These different directions lead to curved trajectories (yellow line). Such curved trajectories explain the complex shape of the phase plot (Figure 3B).

Understanding Apparent Competition Between Aesthetic Weights

What is the reason for the apparent competition between balance and complexity weights during the slow phase of learning in Figure 3? A simple hypothesis is that the apparent competition arises from the negative correlation between balance and complexity, i.e., between the components of \vec{u} (Figure 2A). However, inspection of Eq. 7, suggests alternate hypotheses beyond the negative correlation between balance and complexity. For example, they have different reward structures (Figures 2B,C), possibly leading the weight of one becoming more relevant than the other is. Finally, because motivation affects balance and complexity in different manners, it too, could create an apparent competition (Figure 2D). To test these hypotheses, we ran six new simulations varying input correlation, reward structures, and motivation functions. These simulations eliminated the negative correlation between balance and complexity, made the reward structures identical, or set the motivation to a constant independent of balance and complexity. The results of these simulations appear in Figure 5.

When we eliminated the negative correlation between balance and complexity (standard parameters, except that $\rho = 0$), the apparent competition between their weights did not vanish (Figure 5A). Consequently, this negative correlation is not a necessary condition for the apparent competition. However, the negative correlation affects the apparent competition, because it becomes weaker when we eliminate this correlation, and we see slightly larger balance than complexity weights. Similarly, having different reward structures is not a necessary condition for the apparent competition. If we make the reward structures for balance and complexity identical (both linear as in Figure 2B), the apparent competition remains (Figure 5B). This change leads to an initial rise in both weights followed by an overwhelming relative increase in the balance weight. Finally, in Figure 5C, we remove the effect of motivation from the simulation, by setting $m \equiv 1$. This change results in an isotropic cloud, showing that the shape of the motivation function is a major contributor to the apparent competition. An appropriate motivation function may even be a necessary condition for the apparent competition.

Can the negative correlation between balance and complexity, different reward structures, or the shape of the motivation function be a sufficient condition for the apparent competition between the aesthetic weights? To answer this question, we eliminated two of these conditions at a time. We thus left only one condition in place in each simulation. As seen in Figure 5D, when $m \equiv 1$, and the reward structures are similar for balance and complexity, there is no apparent competition between the weights. Hence, the apparent competition vanishes although the negative correlation is still present. Similarly, when $m \equiv 1$ and we eliminate the negative correlation between balance and complexity, the apparent competition vanishes. It disappears although we still have differences in reward structures (Figure 5E). Thus, neither the negative correlation nor the difference in reward structures is a sufficient condition for the apparent competition. In contrast, the apparent competition

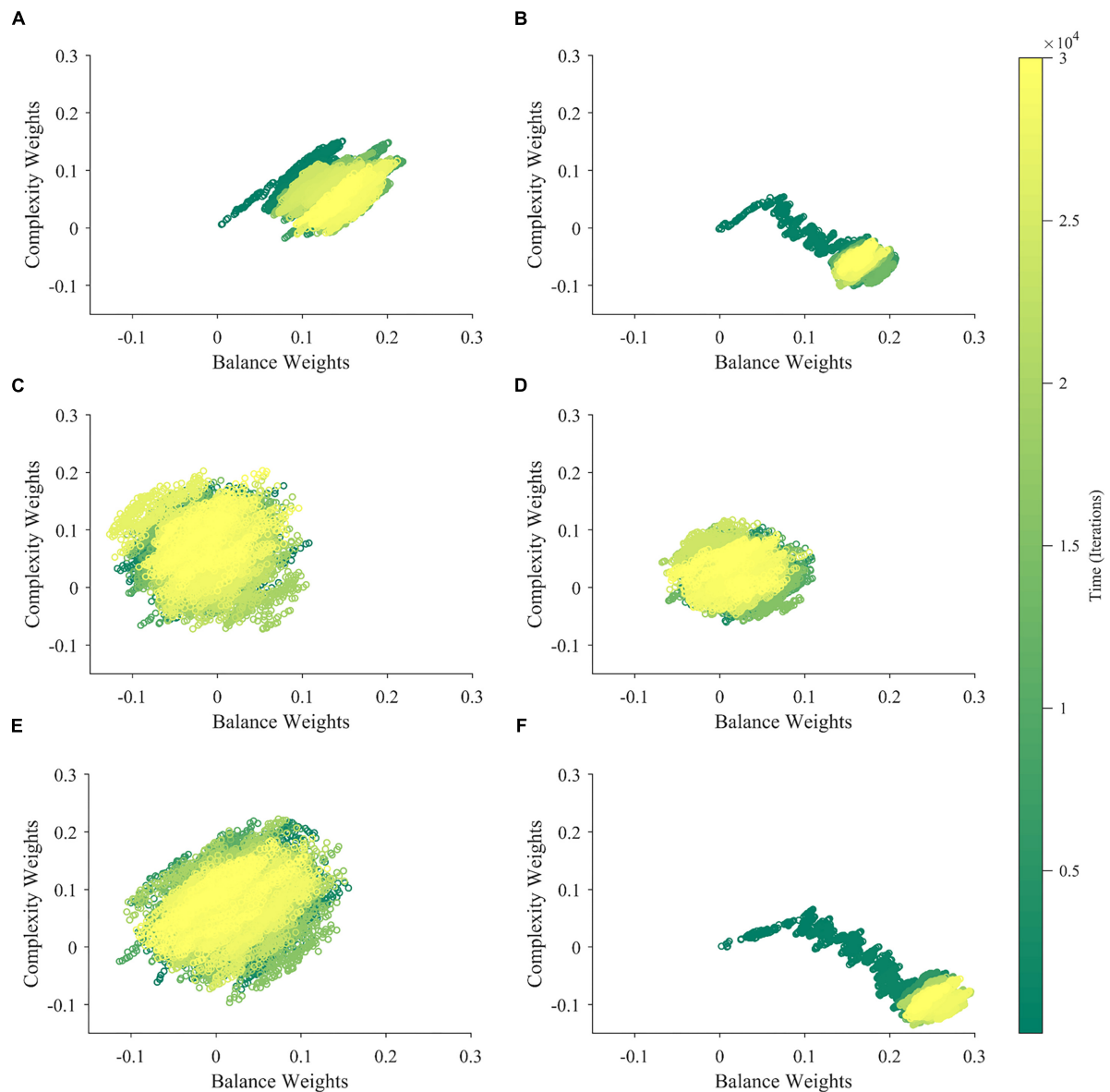


FIGURE 5 | Contributions of the Properties of \bar{u} , r^* , and m to the Apparent Competition between Aesthetic Weights. **(A)** Eliminating the negative correlation between the components of \bar{u} weakens but does not kill the apparent competition. **(B)** Making the reward function r^* similar for balance and complexity may even strengthen apparent competition. **(C)** Making m a constant leads to the virtual elimination of apparent competition. **(D)** When both m is a constant, and the reward function r^* is similar for balance and complexity, there is no apparent competition. **(E)** The same happens when both m is a constant and we eliminate the negative correlation between the components of \bar{u} . **(F)** In contrast, apparent competition remains and can even become stronger when both we eliminate the negative correlation between the components of \bar{u} and the reward function r^* is similar for balance and complexity. Thus, the main factor determining apparent competition in our illustrative model may be the shape of the motivation function.

continues when we eliminate the correlation and the difference in reward structure, leaving the shape of the motivation function intact. This result thus gives further evidence that the appropriateness of this shape may be a sufficient condition for the apparent competition.

Overall, the key factor for the apparent competition between aesthetic values in our illustrative model may be the motivation function. It generates the apparent competition by modulating both sensory sampling and reward. Negative correlations

between the components of the sensory inputs do play a role in the apparent competition but a lesser one.

The Role of Motivation and Reward on Aesthetic Individuality

An important consequence of our theoretical framework is that different individuals develop distinct aesthetic weights. If two individuals were from different societies or cultures, they would

tend to have differences in their learning parameters. These differences are illustrated by the blue boxes Society 1 and 2 in **Figure 1**. Mathematically, these individuals would have different B parameters in Eqs. 5 and 6. However, even if individuals came from the same society, their learning parameters would tend to be distinct (blue boxes labeled Individual 1 and 2 in **Figure 1**). Again, in Eqs. 5 and 6, these individuals would have different \bar{I}_u and \bar{I}_m parameters. In **Figure 6**, we illustrate through computer simulations how this individuality emerges.

To illustrate the effect of individualized learning of aesthetic value in a society, we modeled a scenario for the case of motivation for complexity. While we could have investigated motivation for balance as well, existing research shows different personality traits can account for changes in preference for complexity (Furnham and Bunyan, 1988). Furthermore, it has been shown that the motivation for complexity can be experimentally manipulated (Tinio and Leder, 2009). Thus, we sought to see what happens when motivation for complexity changes across individuals. To test this hypothesis, we varied the peak complexity of the motivation function (**Figures 2D, 6A**). The larger the peak complexity is, the more motivation the individual must act on high complexity. As seen in **Figure 6B**, changing this form of motivation has a direct effect on learned aesthetic weights. Specifically, when the motivation is shifted toward high complexities, late aesthetic weights weight for balance becomes weaker. In contrast, those for complexity become stronger. Hence, the three different individuals in **Figure 6B** (in terms of peak complexity) express different learned aesthetic weights. Moreover, because our theoretical framework is stochastic, the aesthetic weights form clouds in limited regions of the so-called neuroaesthetic space (Aleem et al., 2017). The separation and partial overlap between these clouds is like what we observe for different artists (inset of **Figure 6B**).

In turn, to illustrate how social variations may influence aesthetic learning of individuals, we modulated the reward structure for balance (**Figures 2B, 6C**). The larger the slope of this structure, the more social reward an individual got with highly balanced sensory inputs. As a result, we can see in the **Figure 6D** that the changing of the reward structure has an appropriate effect on aesthetic weights. Increasing the balance slope expectedly increases the weights toward balance, reducing those for complexity. Again, the three different individuals in **Figure 6D** (in terms of social reward structure) express different learned aesthetic weights. Finally, once again, our theoretical framework is stochastic. Consequently, the aesthetic weights form clouds in limited regions of the neuroaesthetic space as also seen in the analysis of aesthetic weights in portraits by master painters of the Early Renaissance (Aleem et al., 2017).

In conclusion, individuality in aesthetic learning emerges in the theoretical framework through variations in either cultural norms or individual motivational states.

Beauty and the Emergence of the Peak-Shift Effect

Any theory for aesthetic learning must account after convergence for as many relevant properties in the literature as possible.

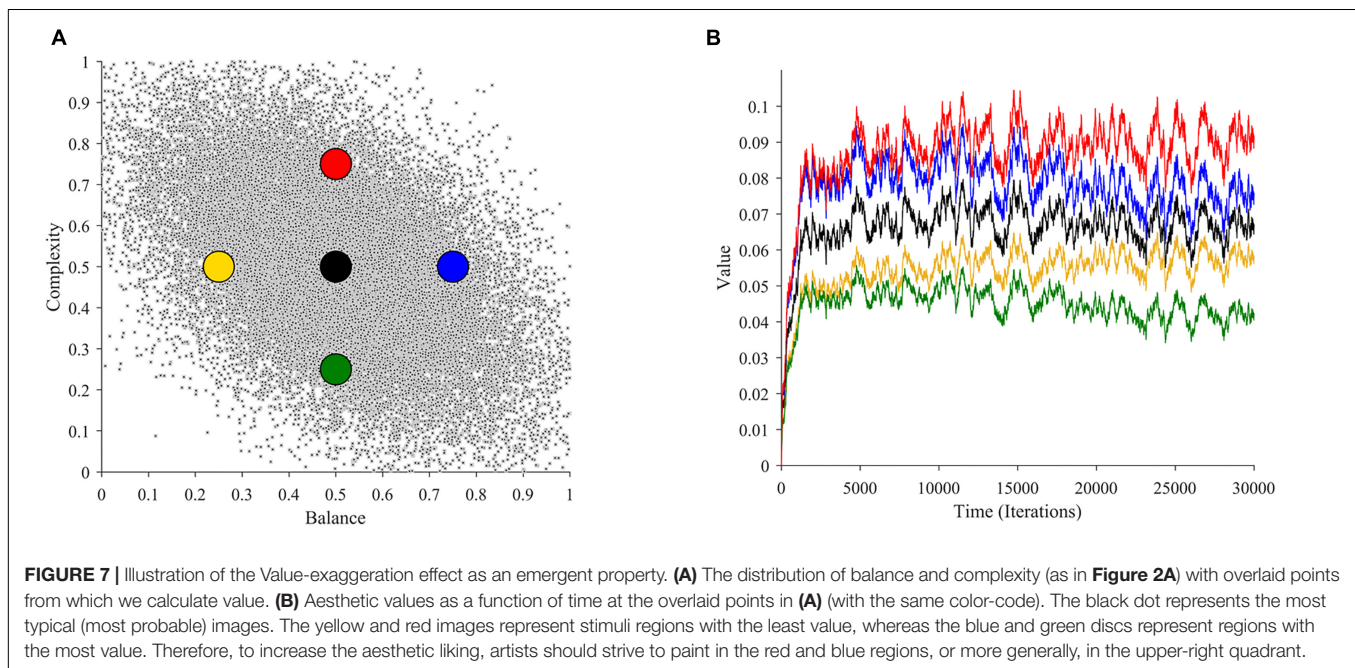
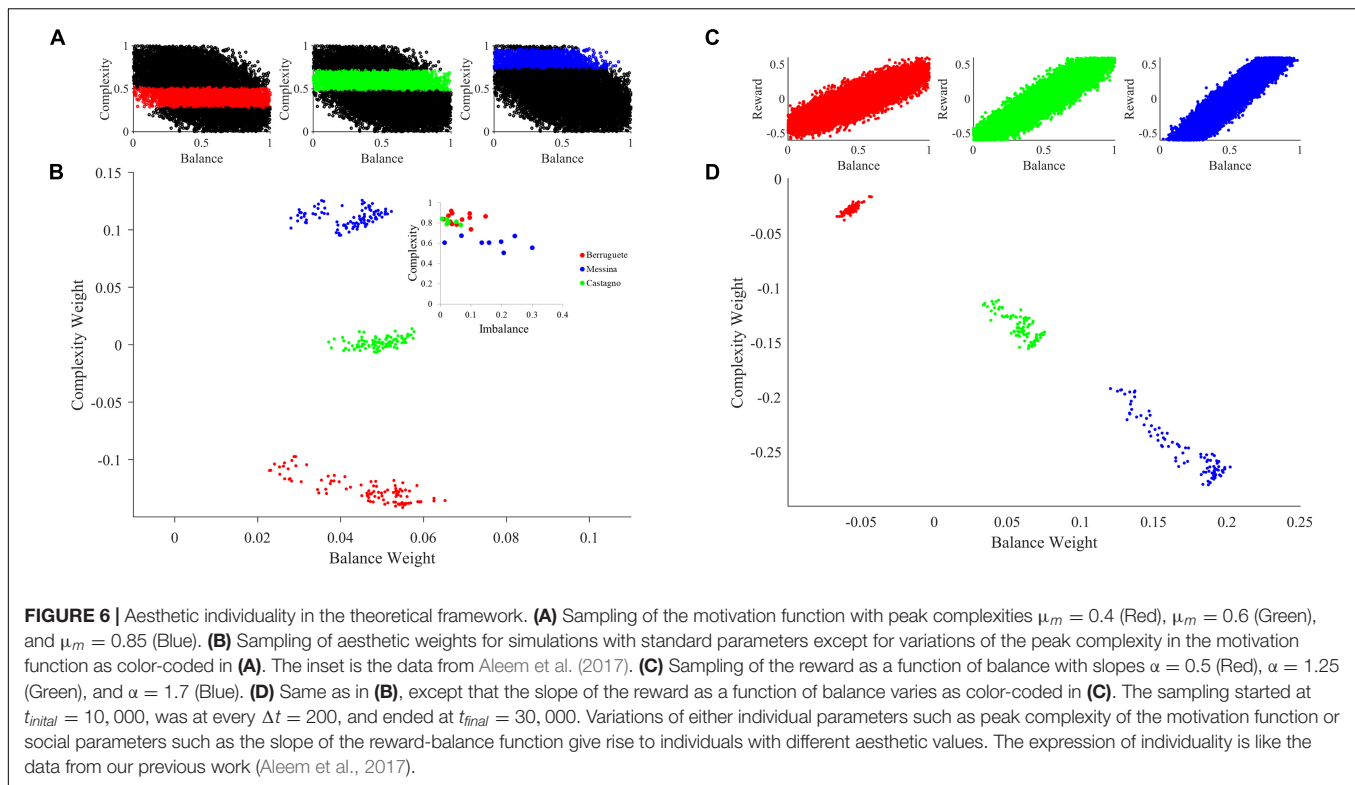
So far, we have discussed the dynamics of learning aesthetic weights. However, we have not yet explored the amount of value possible in different regions of the neuroaesthetic space. This exploration naturally leads us to the broader question of creation of art and beauty. Where does beauty, or in our case, regions of highest value, exist? We looked to the literature for existing hypotheses on this question. One of the most prominent hypotheses in this regard has to do with the “peak-shift” effect (Ramachandran and Hirstein, 1999). It supposes that the beauty of an object is partly owed to the exaggeration of some of its characteristics. According to the hypothesis, if an attribute signals value normally, exaggerating that attribute would lead to greater value. This effect is theorized to explain the tendency of artists to exaggerate variables that contribute to aesthetic emotions. Accordingly, visual artists should tend to exaggerate certain statistical properties like symmetry, complexity or even certain facial features as compared to what one observes typically (Costa and Corazza, 2006; Graham and Redies, 2010; Aleem et al., 2017). Thus, because of this exaggeration effect, beauty is not merely copying reality.

In this section, we study whether and why our theoretical framework is consistent with such a peak-shift effect. To perform this study, we calculated aesthetic values of images with different complexities and balances in our simulations. We used these calculations to compare the aesthetic values of images with the most typical statistics with those with less probability of occurring. The comparison appears in **Figure 7**.

As seen in **Figure 7A**, there are many possible regions of the neuroaesthetic space that a certain scene or painting could occupy. However, not all these regions are identical in terms of value. For example, the region indicated by the black dot represents what is most typical. If we turn to **Figure 7B**, we see that this region leads to the learning of moderate overall value. It is apparent that there are regions with greater or lesser value. In our example in **Figure 7A**, the yellow and green dots represent regions with lower value in relation to the black dot. In turn, the regions around the red and blue dots yield greater value than the regions around the black dot. Consequently, if an artist wants to maximize value, they would be keen to paint with attributes in the regions of the upper-right quadrant of this neuroaesthetic space. In our case, artists would thus tend to exaggerate the complexity and balance concurrently to increase the aesthetic appeal of their work.

Mathematical analyses show that this value-exaggeration result is a general property for our theoretical framework (**Supplementary Materials, Section D**). Thus, this result is applicable beyond the parameters of the simulations in **Figure 7**. The analysis also extends the results for broad classes of learning models that are nonlinear, that is, not following Eq. 1. If the properties of these learning models and of the probability distribution of sensory signals obey general conditions, then the value-exaggeration result will hold. The linear model in Eq. 4 will almost always obey these conditions.

In conclusion, our analyses support the peak-shift effect. Our framework does so by predicting that the most typical inputs are not necessarily the sources of highest predicted value.



Other possible inputs yield more aesthetic value than reality, thus appearing more beautiful. As an important extension, our theoretical framework predicts that the value landscape of possible inputs is different for each individual. Our explanation for this effect is thus that what matters for aesthetics is not the statistics of sensory inputs but their values to perceivers.

DISCUSSION

The field of neuroaesthetics has progressed rapidly lately, especially with regards to the understanding of the “what” and “where” of aesthetic preferences. However, one of the biggest remaining questions has to do with “how” we develop aesthetic

values in the first place (Göksun et al., 2014). The origin of certain preferences such as that for contour, symmetry, or contrast can be explained by evolutionary accounts (Reber et al., 2004; Bar and Neta, 2006). However, even these seemingly innate preferences are subject to experience dependent changes (Germine et al., 2015; Huang et al., 2018). Therefore, it is important to consider the dynamics of aesthetic preferences over time. Here, we presented a theoretical contribution to the understanding of these dynamics. Our theoretical framework proposes that learning forms a large component of aesthetic values. We operationalized this proposal through a computational model of reinforcement learning. We discuss our interpretation of some of the important findings of this modeling effort in greater detail here. For a quick summary of the results, please refer to the beginning of the Results section.

Interpretations of the Simulation Results

Our results suggest that the time course of aesthetic learning has different phases. Such a multi-phase result is expected from a reward-based learning model. If complete prior inexperience is assumed, then a rapid initial phase is bound to happen, followed by a slower one. These fast-then-slow learning dynamics are characteristic of reinforcement-learning frameworks, where the change of weights is proportional to the error (Dayan and Abbott, 2001; Sutton and Barto, 2018). The fast phase corresponds to the large initial errors, but when they decrease, the learning rate slows down. Our decision for setting the initial weights at zero allows us to understand general principles of learning and to see that it has distinct phases. Such fast aesthetic learning should occur mostly early on in life. However, whether we are born without any aesthetic preconceptions is an open question. What is the likelihood that we come across a completely novel stimuli in our adult lives? More likely, we see previously learned attributes in a novel way, as for example, a drummer seeing a table for the first time. Or take abstract art, which combines familiar visual primitives in a novel way. In these cases, some type of initial prediction of value may exist, albeit with low confidence and high noise. Here too, our theoretical framework would predict that as soon as learning occurs it would be multi-phasic, rapid initially, and then coming to a consolidation.

A surprising result of our model was the apparent competition between aesthetic variables in the second phase of learning. We investigated the source of this apparent competition in our illustrative model and found that motivation is the main driving force. Therefore, motivation appears to be crucial in guiding different trajectories of learning. While motivation in our model seems like a simple gating mechanism, the results show that motivation has complex consequences on learning. This is due to the probabilistic dependence of motivation on sensory inputs and social rewards. Hence, motivation allows for state-dependent learning, accounting for aesthetic diversity and individuality. This result makes intuitive sense, as while we may all start from similar points, our motivations to engage with certain objects and environments will be different. Therefore, different motivations will lead us to divergent paths. These motivations in turn will be influenced by many internal and external factors including the environment and social standards. The interaction and

co-dependence of these factors leads to many unique outcomes from a shared starting point.

A unique aspect of our theoretical framework and thus, of our illustrative model is their statistical nature. Coupled with the nonlinearities in the model, this statistical nature leads to many important and surprising results. One such result is the apparent competitive interactions mentioned above. Another important result is that the learned aesthetic weights are stochastic, that is, we should not expect them to be constant and stable, but to fluctuate over time. In our illustrative examples with only two variables, the aesthetic values eventually converge stochastically around a fixed point. The situation should be more complex in the real world, because the number of variables would be higher, possibly leading to multiple fixed points instead of just one. Such a multiplicity would give the appearance of multiple possible aesthetic stable states. Another complexity stems from the key difference between aesthetic weights and values. Aesthetic weights are different from the values, as different weights can lead to same values. Therefore, any apparent fixed-point in aesthetic weights could drift over time adding more variability to aesthetic preferences. The associated aesthetic values, as seen in **Figure 7** are stochastic as well. Therefore, our theoretical framework makes a surprising prediction that aesthetic preferences are not the same from one moment to another. This goes against the common assumption that our preferences are relatively stable and thus, that we only need to account for them once. In support of this, converging evidence is beginning to call the long-held assumption of preference stability into question (Höfel and Jacobsen, 2003; Chen and Risen, 2010; Kościński, 2010; McManus et al., 2010; Pugach et al., 2017). The observed instability is commonly attributed to noise within the internal sampling of subjective values. Our theoretical framework makes an additional prediction that the values themselves may be stochastic.

Limitations and Outlook

At first glance, our theoretical framework for aesthetic learning may seem too reductionist. Aesthetic experiences are complex and many factors are at play. We propose that low-level features and reward-based learning forms just one component of acquiring and using aesthetic values. Our theoretical framework does not address other important aspects for aesthetic emotions, such as semantics, attention, and memory (Leder et al., 2004). We acknowledge these factors play a role in the formation of aesthetic values and their omission is not to undermine them. Instead, we chose to limit the complexity of our theoretical framework at this first iteration to serve as a basic building block on which to incorporate the aforementioned factors. However, even a model based on low-level features can still be highly informative on aesthetic preferences of individuals, as recently demonstrated by Iigaya et al. (2020). Additionally, a reinforcement learning circuit is easily amenable to additional factors, for example Leong et al. incorporate attention directly into the reinforcement-learning circuitry computing subjective value, as we did with motivation (Leong et al., 2017).

A factor that features prominently in studies of aesthetic preference formation but not considered in our framework here is familiarity (Zajonc, 1968). Numerous studies show

that increased familiarity tends to improve appraisal, with no apparent reward (Leder et al., 2004; Park et al., 2010; Lindell and Mueller, 2011). However, we argue that familiarity may either be intrinsically rewarding or a promoter to value. For instance, every new exposure leading to familiarity improves our processing of the stimuli. This improvement likely facilitates object recognition and ability to extract semantic and emotional content, which is rewarding (Reber et al., 2004). At a conceptual level, familiarity is intimately tied with novelty, which may also be intrinsically rewarding (Biederman and Vessel, 2006). However, the relationship between these two factors is not entirely clear. For example, roles of familiarity and novelty may be dependent on context (Tinio and Leder, 2009; Park et al., 2010). To help understand this, future iterations of our framework would benefit by incorporating these related factors. For example, one could incorporate an aspect of reward that is contingent both on the novelty and the number of exposures to certain stimuli, as suggested by Biederman and Vessel (2006). As of now, we do not differentiate between these aspects of reward, but this distinction is important and a necessary addition for the future.

As far as implementation, our model assumes a linear relationship between aesthetic weights and values (Eq. 1). Biologically, this linearity is not reflective of typical reward-related synapses (Schultz, 2015). Recent assessments of the field of neuroaesthetics have signaled the need for a new conception of aesthetics that incorporates distributed processing and non-linear recurrent networks (Leder and Nadal, 2014; Nadal and Chatterjee, 2019). While we agree, we suggest that linearity is a suitable starting point. Recent work comparing a linear rule versus a deep neural network to predict subjective aesthetic value found that both fared comparably (Iigaya et al., 2020). We argue that this also applies for our theoretical framework for aesthetic learning. We have shown that most nonlinearities can explain the value-exaggeration effect (section “Beauty and the Emergence of the Peak-shift Effect” and Section D of **Supplementary Materials**). Most of our other results would likely be similar if we used a monotonic nonlinear relationship between aesthetic weights and values. For example, Eq. 4 is all that we need to explain the fast-then-slow dynamics of aesthetic learning. From that equation, when the error is large or small, so is the rate of learning, regardless of whether the relationship between aesthetic weights and values is linear. Likewise, Eq. 4 is all that we need to explain the near optimality of the learned value, because this equation predicates the tendency of the minimization of error. Consequently, near optimality should occur regardless of whether the relationship between aesthetic weights and values is linear. We can make similar arguments for the non-necessity of linearity for almost all the results in this paper. Thus, assuming a linear relationship between aesthetic weights and values is not a major problem for the validity of our results.

Other limitations are not with the theoretical framework but with the illustrative model. For example, we limit the sensory inputs in our model to two visual statistics. While this simplifies our simulations of the model, it is not reflective of the external world, where a deluge of variables is at play. These variables may all exist in some complicated multi-dimensional space, which we have previously termed the “neuroaesthetic space” (Aleem et al.,

2017). Future implementations would certainly have to increase the dimensionality and type of sensory inputs into the model.

Overall, our model, while limited, provide a platform for further research, such as by increasing the richness of the model in the many ways mentioned above. Equally as important are efforts to test empirically the predictions of the model. More developmental and longitudinal empirical studies of aesthetic preferences are needed. For example, one could conduct extensive reinforcement learning studies to determine how learning modulates subjective values over long periods as shown by Wimmer et al. (2018). Similarly, one could empirically test the prediction of temporal variability in aesthetic values.

Compatibility With Existing Frameworks

Where does our contribution fall into place within the existing theories of aesthetics? Most of the extant theoretical frameworks for aesthetics aim to explain the phenomena at hand. That is not to say that some of these frameworks do not consider the importance of temporal aspects, albeit implicitly. In particular, learning is a key part of many of the existing influential theoretical frameworks. For example, in Chatterjee and Vartanian’s “Aesthetic Triad” model, aspects of learning and reward make two out of the three nodes (Chatterjee and Vartanian, 2014). Others have made reward-based learning central to their theories. For example, in formulating the “Aesthetic Preference Formation” model, Skov defines nodes associated with sensory stimuli, reward prediction, learning, and context amongst others (Skov, 2010). Like us, he stresses the involvement of a reinforcement learning mechanism that is not unique to aesthetics.

In regards to the time frames of learning, Nadal and Chatterjee describe three time-scales influencing aesthetic preferences (Nadal and Chatterjee, 2019). Our model is most like their middle-range time scale, which concerns “individual experience and cultural context.” In a similar vein, Vessel and colleagues build on the reward circuitry with an explicit emphasis on time (Biederman and Vessel, 2006; Vessel and Rubin, 2010). Like us, they implicate associative processes as a central mechanism of time-dependent preference formation. In their view, aesthetic preferences are shaped by the temporal summation of their associative components. In contrast to us, their theory mainly focuses on mechanisms of shared associations. For example, most people will have favorable memories of beaches, therefore leading to a large consensus of preference. However, our model accounts for this preference effect as well by incorporating social statistics of rewards. Another important aspect of Vessel and colleagues work is that “deeper” and more meaningful rewards will lead to stronger preferences. While we do not consider this aspect in our framework, it would be a compatible addition.

Unlike the theories mentioned above, our theoretical framework is specified in a fully computational manner. However, other theories are computational, too! Perhaps the earliest computational theory of aesthetics comes from Martindale. This theory largely focuses on pleasure, formulating that the enjoyment derived from an aesthetic stimulus is proportional to the number of cognitive units activated by it, as envisioned in a neural network (Martindale, 1984). Another

computational theory comes from Schmidhuber, who contends that aesthetic preferences are largely driven by intrinsic reward (Schmidhuber, 2010). According to him, when we learn new things and improve our predictions of the world, we maximize this reward. This idea is similar to the theory put forth by Van de Cruys and Wagemans who propose that aesthetic value results from the resolution of prediction errors caused by ambiguities in art (Van de Cruys and Wagemans, 2011). However, all of these theories are primarily focused on the nature of the aesthetic experience and its ensuing reward. While these theories use reinforcement learning circuitry as their basis, unlike us, they do not explicitly consider the learning of aesthetic values over time. But both Schmidhuber and Van de Cruys and Wagemans do argue that experience with different environments over time will lead to differences in predictions, accounting for individual and cultural differences. Nevertheless, unlike these theories, we do not concern ourselves with the nature of reward. In our view, rewards form aesthetic values, whether the rewards are external or internal.

In our framework, the predicted reward, often termed “value” is akin to aesthetic value. This is arguably the most important assumption of our framework. Thus, we propose that our prior experiences with an object influence the values assigned to the many attributes of that object. When we sense these attributes in a new sensory input, their associated values will influence our aesthetic preference for that input (Barrett and Bar, 2009). An illustration of this aesthetic value transfer effect is shown by Strauss et al. (2013). They found that preferences for two-dimensional color patches were systematically altered by looking at positive or negative objects of the same color. More direct studies using instrumental conditioning show that preferences for cues proportionally coincide with their ability to predict reward, even when subliminal (Pessiglione et al., 2008). Therefore, we contend that as certain stimulus attributes signal reward, they themselves become secondary reinforcers, and hence obtain aesthetic value. For example, humans and other animals may have initially preferred symmetry because of its health cues, such as in faces (Treder, 2010). However, after eons of evolution, symmetry may now be a secondary reinforcer itself, signaling value independently (Pombo and Velasco, 2019). Neuroimaging studies support this viewpoint, showing that secondary reinforcers activate similar regions associated with pleasure as do primary reinforcers (Sescousse et al., 2013).

Whether the pleasure from aesthetic value is phenomenologically distinct from other pleasures is an open question (Christensen, 2017; Nadal and Skov, 2018; Menninghaus et al., 2019; Skov and Nadal, 2020). On the one hand, neuroimaging studies show that a range of rewarding, pleasure-giving experiences are processed in the same brain regions. This common processing of reward allows us to make value-based decisions across various goods (Levy and Glimcher, 2012). On the other hand, network-based studies of deeply rewarding phenomena show the concurrent role of other brain processes (van Elk et al., 2019). For example, the default-mode network has been shown to be modulated by intense aesthetic experiences (Vessel et al., 2012). Thus, a subjectively deep experience is likely to activate different brain networks, yet simultaneously be under the constrain of the neurobiological

roots of “basic” pleasures. The broader implications of these differences remain to be discovered.

What Is Beautiful According to Our Theoretical Framework?

“Beauty is nature’s brag.” Thus, the poet John Milton wrote in praise of the beauty that one often experiences in nature (Milton, 1858). We hear of such experiences commonly, but not all natural scenes are pleasant or breathtaking. For example, some scenes may be repulsive to some people by including a rotting corpse, an approaching snake, or a spider web. What is it that makes some natural scenes beautiful? Following from the discussion in section “Compatibility with Existing Frameworks,” our theoretical framework proposes that when a certain natural scene appears beautiful, it does so, because its statistics elicit high positive value. Our results showed that when looking at the overall value landscape, certain regions that are far away from the norm will correspond to higher value (Figure 7). These are the regions that our theoretical framework may consider “beautiful.” Accordingly, only the minority of scenes might be truly beautiful by eliciting high values. These are the scenes that exaggerate high-value attributes. This sentiment is similar to Ramachandran and Hirstein’s application of the “peak-shift” effect to beauty, proposing that it often comes from exaggeration (Ramachandran and Hirstein, 1999). We thus may ask, to what “exaggerated” natural scenes was Milton referring to? Perhaps, Milton’s scenes would have some sort of exaggerated statistics related to attributes that were innate, that is, formed due to evolutionary pressures (Aleem et al., 2019). Alternatively, according to the theoretical framework here, Milton’s scenes would have exaggerated sensory statistics related with positive experiences in his youth.

Why would the brain evolve a mechanism that prefers exaggeration rather than the most common reality? We argue that any learning system such as the brain would likely prefer exaggeration if its goal is to maximize reward. Consequently, perhaps evolution has allowed our ancestors to choose actions that maximize value. However, there are limits to exaggeration. For example, our results show that exaggeration in the wrong direction will lead to lower than normal aesthetic values, or what one may consider “ugly.”

In sum, our theoretical framework extends the peak-shift hypothesis through individualized value exaggeration. According to the framework, the aesthetic weights that maximize reward are not universal across all individuals. Each person has an individual set of near-optimal aesthetic weights according to personal motivations, and social and environmental contexts. Neuroaesthetic-space regions of high value, or beauty, to one may be regions of low value or ugly to another. We conclude that the different senses of beauty across individuals are not arbitrary, but stem from a personalized near optimization of values.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

HA and NG were involved in developing the framework and writing and revising the manuscript. NG developed the equations and carried out the mathematical analyses. HA carried out the simulations. IC-H helped to edit and revise the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the research funds allocated to NG from the Georgetown University.

REFERENCES

- Aleem, H., Correa-Herran, I., and Grzywacz, N. M. (2017). Inferring master painters' esthetic biases from the statistics of portraits. *Front. Hum. Neurosci.* 11:94. doi: 10.3389/fnhum.2017.00094
- Aleem, H., Pombo, M., Correa-Herran, I., and Grzywacz, N. M. (2019). "Is beauty in the eye of the beholder or an objective truth? A neuroscientific answer," in *Mobile Brain–Body Imaging and the Neuroscience of Art, Innovation and Creativity*, eds J. L. Contreras-Vidal, D. Robleto, J. G. Cruz-Garza, J. M. Azorin, and C. S. Nam (Cham: Springer Nature Switzerland AG).
- Averbeck, B. B., and Costa, V. D. (2017). Motivational neural circuits underlying reinforcement learning. *Nat. Neurosci.* 20, 505–512. doi: 10.1038/nn.4506
- Bar, M., and Neta, M. (2006). Humans prefer curved visual objects. *Psychol. Sci.* 17, 645–648. doi: 10.1111/j.1467-9280.2006.01759.x
- Barrett, L. F., and Bar, M. (2009). See it with feeling: affective predictions during object perception. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1325–1334. doi: 10.1098/rstb.2008.0312
- Berlyne, D. E. (1971). *Aesthetics and Psychobiology*. New York, NY: Appleton-Century-Crofts.
- Biederman, I., and Vessel, E. A. (2006). Perceptual pleasure and the brain: a novel theory explains why the brain craves information and seeks it through the senses. *Am. Sci.* 94, 247–253.
- Brown, S., and Dissanayake, E. (2009). The arts are more than aesthetics: neuroaesthetics as narrow aesthetics. *Neuroaesthetics* 43:57.
- Brown, S., Gao, X., Tisdelle, L., Eickhoff, S. B., and Liotti, M. (2011). Naturalizing aesthetics: brain areas for aesthetic appraisal across sensory modalities. *Neuroimage* 58, 250–258. doi: 10.1016/j.neuroimage.2011.06.012
- Chatterjee, A., and Vartanian, O. (2014). Neuroaesthetics. *Trends Cogn. Sci.* 18, 370–375.
- Chen, M. K., and Risen, J. L. (2010). How choice affects and reflects preferences: revisiting the free-choice paradigm. *J. Pers. Soc. Psychol.* 99:573. doi: 10.1037/a0020217
- Christensen, J. F. (2017). Pleasure junkies all around! Why it matters and why 'the arts' might be the answer: a biopsychological perspective. *Proc. R. Soc. B Biol. Sci.* 284:20162837. doi: 10.1098/rspb.2016.2837
- Costa, M., and Corazza, L. (2006). Aesthetic phenomena as supernormal stimuli: the case of eye, lip, and lower-face size and roundness in artistic portraits. *Perception* 35, 229–246. doi: 10.1068/p3449
- Craig, A. D. (2009). How do you feel–now? The anterior insula and human awareness. *Nature Rev. Neurosci.* 10, 59–70. doi: 10.1038/nrn2555
- Dabney, W., Kurth-Nelson, Z., Uchida, N., Starkweather, C. K., Hassabis, D., Munos, R., et al. (2020). A distributional code for value in dopamine-based reinforcement learning. *Nature* 577, 671–675. doi: 10.1038/s41586-019-1924-6
- Dayan, P., and Abbott, L. F. (2001). *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. Cambridge, MA: Massachusetts Institute of Technology Press.
- Dechesne, F., Di Tosto, G., Dignum, V., and Dignum, F. (2013). No smoking here: values, norms and culture in multi-agent systems. *Artif. Intell. Law* 21, 79–107. doi: 10.1007/s10506-012-9128-5

ACKNOWLEDGMENTS

We would like to thank the members of the Interdisciplinary Program in Cognitive Science at Georgetown University for their comments on the early versions of this work. We also thank Rick Pike and Helen Ryan for their administrative support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnhum.2020.00345/full#supplementary-material>

- Donderi, D. C. (2006). Visual complexity: a review. *Psychol. Bull.* 132, 73–97.
- Field, D. J. (1994). What is the goal of sensory coding? *Neural Comp.* 6, 559–601. doi: 10.1162/neco.1994.6.4.559
- Furnham, A., and Bunyan, M. (1988). Personality and art preferences. *Eur. J. Pers.* 2, 67–74. doi: 10.1002/per.2410020106
- Germine, L., Russell, R., Bronstad, P. M., Blokland, G. A., Smoller, J. W., Kwok, H., et al. (2015). Individual aesthetic preferences for faces are shaped mostly by environments, not genes. *Curr. Biol.* 25, 2684–2689. doi: 10.1016/j.cub.2015.08.048
- Göksun, T., Kranjec, A., and Chatterjee, A. (2014). "The development of visual art preferences," in *Proceedings of the 23rd International Association of Empirical Aesthetics*, New York, NY.
- Graham, D. J., and Redies, C. (2010). Statistical regularities in art: relations with visual coding and perception. *Vision Res.* 50, 1503–1509. doi: 10.1016/j.visres.2010.05.002
- Güçlütürk, Y., Jacobs, R. H., and Lie, R. V. (2016). Liking versus complexity: decomposing the inverted U-curve. *Front. Hum. Neurosci.* 10:112. doi: 10.3389/fnhum.2016.00112
- Höfel, L., and Jacobsen, T. (2003). Temporal stability and consistency of aesthetic judgments of beauty of formal graphic patterns. *Percept. Motor Skills* 96, 30–32. doi: 10.2466/pms.2003.96.1.30
- Howard, J. D., and Kahnt, T. (2017). Identity-specific reward representations in orbitofrontal cortex are modulated by selective devaluation. *J. Neurosci.* 37, 2627–2638. doi: 10.1523/jneurosci.3473-16.2017
- Huang, Y., Xue, X., Spelke, E., Huang, L., Zheng, W., and Peng, K. (2018). The aesthetic preference for symmetry dissociates from early-emerging attention to symmetry. *Sci. Rep.* 8:6263.
- Iigaya, K., Yi, S. I., Wahle, A., Tanwisuth, K., and O'Doherty, J. P. (2020). Aesthetic preference for art emerges from a weighted integration over hierarchically structured visual features in the brain. *bioRxiv* [Preprint]. doi: 10.1101/2020.02.09.940353
- Kościński, K. (2010). Do they know what they like? Intra-individual variation of female facial preferences. *J. Evol. Psychol.* 8, 23–55. doi: 10.1556/jep.8.2010.1.4
- Lacey, S., Hagtvædt, H., Patrick, V. M., Anderson, A., Stilla, R., Deshpande, G., et al. (2011). Art for reward's sake: visual art recruits the ventral striatum. *Neuroimage* 55, 420–433. doi: 10.1016/j.neuroimage.2010.11.027
- Leder, H., Belke, B., Oeberst, A., and Augustin, D. (2004). A model of aesthetic preference and aesthetic judgments. *Br. J. Psychol.* 95, 489–508. doi: 10.1348/0007126042369811
- Leder, H., and Nadal, M. (2014). Ten years of a model of aesthetic appreciation and aesthetic judgments: the aesthetic episode—Developments and challenges in empirical aesthetics. *Br. J. Psychol.* 105, 443–464. doi: 10.1111/bjop.12084
- Leong, Y. C., Radulescu, A., Daniel, R., DeWoskin, V., and Niv, Y. (2017). Dynamic interaction between reinforcement learning and attention in multidimensional environments. *Neuron* 93, 451–463. doi: 10.1016/j.neuron.2016.12.040
- Levy, D. J., and Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Curr. Opin. Neurobiol.* 22, 1027–1038. doi: 10.1016/j.conb.2012.06.001

- Lindell, A. K., and Mueller, J. (2011). Can science account for taste? Psychological insights into art appreciation. *J. Cogn. Psychol.* 23, 453–475. doi: 10.1080/20445911.2011.539556
- Little, A. C., Apicella, C. L., and Marlowe, F. W. (2007). Preferences for symmetry in human faces in two cultures: data from the UK and the Hadza, an isolated group of hunter-gatherers. *Proc. Biol. Sci.* 274, 3113–3117. doi: 10.1098/rspb.2007.0895
- Little, A. C., DeBruine, L. M., and Jones, B. C. (2011). Exposure to visual cues of pathogen contagion changes preferences for masculinity and symmetry in opposite-sex faces. *Proc. R. Soc. B Biol. Sci.* 278, 2032–2039. doi: 10.1098/rspb.2010.1925
- Martindale, C. (1984). The pleasures of thought: a theory of cognitive hedonics. *J. Mind Behavior* 5, 49–80.
- Masuda, T., Gonzalez, R., Kwan, L., and Nisbett, R. E. (2008). Culture and aesthetic preference: comparing the attention to context of East Asians and Americans. *Pers. Soc. Psychol. Bull.* 34, 1260–1275. doi: 10.1177/0146167208320555
- McManus, I., Cook, R., and Hunt, A. (2010). Beyond the golden section and normative aesthetics: why do individuals differ so much in their aesthetic preferences for rectangles? *Psychol. Aesthet. Creat. Arts* 4, 113–126. doi: 10.1037/a0017316
- Menninghaus, W., Wagner, V., Wassiliwizky, E., Schindler, I., Hanich, J., Jacobsen, T., et al. (2019). What are aesthetic emotions? *Psychol. Rev.* 126, 171–195.
- Milton, J. (1858). *Comus a Mask by John Milton*, G. Abingdon: Routledge.
- Nadal, M., and Chatterjee, A. (2019). Neuroaesthetics and art's diversity and universality. *Wiley Interdiscip. Rev. Cogn. Sci.* 10:e1487. doi: 10.1002/wcs.1487
- Nadal, M., and Skov, M. (2018). The pleasure of art as a matter of fact. *Proc. R. Soc. B Biol. Sci.* 285:20172252. doi: 10.1098/rspb.2017.2252
- Nelson, L. D., and Morrison, E. L. (2005). The symptoms of resource scarcity: judgments of food and finances influence preferences for potential partners. *Psychol. Sci.* 16, 167–173. doi: 10.1111/j.0956-7976.2005.00798.x
- Park, D. C., and Huang, C.-M. (2010). Culture wires the brain: a cognitive neuroscience perspective. *Perspect. Psychol. Sci.* 5, 391–400. doi: 10.1177/1745691610374591
- Park, J., Shimojo, E., and Shimojo, S. (2010). Roles of familiarity and novelty in visual preference judgments are segregated across object categories. *Proc. Natl. Acad. Sci. U.S.A.* 107, 14552–14555. doi: 10.1073/pnas.1004374107
- Pessiglione, M., Petrovic, P., Daunizeau, J., Palminteri, S., Dolan, R. J., and Frith, C. D. (2008). Subliminal instrumental conditioning demonstrated in the human brain. *Neuron* 59, 561–567. doi: 10.1016/j.neuron.2008.07.005
- Pombo, M., and Velasco, C. (2019). How aesthetic features convey the concept of brand premiumness. *PsyArXiv*. [Preprint]. doi: 10.31234/osf.io/7kpwz
- Pouget, A., Beck, J. M., Ma, W. J., and Latham, P. E. (2013). Probabilistic brains: knowns and unknowns. *Nat. Neurosci.* 16, 1170–1178. doi: 10.1038/nn.3495
- Pugach, C., Leder, H., and Graham, D. J. (2017). How stable are human aesthetic preferences across the lifespan? *Front. Hum. Neurosci.* 11:289. doi: 10.3389/fnhum.2017.00289
- Ramachandran, V. S., and Hirstein, W. (1999). The science of art: a neurological theory of aesthetic experience. *J. Conscious. Stud.* 6, 15–51.
- Reber, R., Schwarz, N., and Winkielman, P. (2004). Processing fluency and aesthetic pleasure: is beauty in the perceiver's processing experience? *Pers. Soc. Psychol. Rev.* 8, 364–382. doi: 10.1207/s15327957pspr0804_3
- Rosenbaum, S. (1961). Moments of a truncated bivariate normal distribution. *J. R. Stat. Soc. Series B (Methodological)* 23, 405–408. doi: 10.1111/j.2517-6161.1961.tb00422.x
- Sartwell, C. (2012). "Beauty," in *The Stanford Encyclopedia of Philosophy* (Winter, 2017 Edn), ed E. N. Zalta. Stanford, CA: Metaphysics Research Lab, Stanford University.
- Schmidhuber, J. (2010). Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *IEEE Trans. Auton. Ment. Dev.* 2, 230–247. doi: 10.1109/tamd.2010.2056368
- Schultz, W. (2015). Neuronal reward and decision signals: from theories to data. *Physiol. Rev.* 95, 853–951. doi: 10.1152/physrev.00023.2014
- Senzaki, S., Masuda, T., and Nand, K. (2014). Holistic versus analytic expressions in artworks: cross-cultural differences and similarities in drawings and collages by Canadian and Japanese school-age children. *J. Cross Cult. Psychol.* 45, 1297–1316. doi: 10.1177/0022022114537704
- Sescousse, G., Caldú, X., Segura, B., and Dreher, J.-C. (2013). Processing of primary and secondary rewards: a quantitative meta-analysis and review of human functional neuroimaging studies. *Neurosci. Biobehav. Rev.* 37, 681–696. doi: 10.1016/j.neubiorev.2013.02.002
- Silvia, P. J., Henson, R. A., and Templin, J. L. (2009). Are the sources of interest the same for everyone? Using multilevel mixture models to explore individual differences in appraisal structures. *Cogn. Emot.* 23, 1389–1406. doi: 10.1080/02699930902850528
- Skov, M. (2010). "The pleasure of art," in *Pleasures of the Brain*, eds M. Kringelbach and K. Berridge (Oxford: Oxford University Press), 270–283.
- Skov, M., and Nadal, M. (2020). There are no aesthetic emotions: comment on Menninghaus et al. (2019). *Psychol. Rev.* 127, 640–649. doi: 10.1037/rev0000187
- Sorokowski, P., Sorokowska, A., and Witzel, C. (2014). Sex differences in color preferences transcend extreme differences in culture and ecology. *Psychon. Bull. Rev.* 21, 1195–1201. doi: 10.3758/s13423-014-0591-8
- Steciuch, C. C., Kopatich, R. D., Feller, D. P., Durik, A. M., and Millis, K. (2019). Don't go with your gut: exploring the role of motivation in aesthetic experiences. *Psychol. Aesthet. Creat. Arts*. doi: 10.1037/aca0000259
- Strauss, E. D., Schloss, K. B., and Palmer, S. E. (2013). Color preferences change after experience with liked/disliked colored objects. *Psychon. Bull. Rev.* 20, 935–943. doi: 10.3758/s13423-013-0423-2
- Sutton, R. S., and Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT press.
- Tinio, P. P., and Leder, H. (2009). Just how stable are stable aesthetic features? Symmetry, complexity, and the jaws of massive familiarization. *Acta Psychol. (Amst)* 130, 241–250. doi: 10.1016/j.actpsy.2009.01.001
- Treder, M. S. (2010). Behind the looking-glass: a review on human symmetry perception. *Symmetry* 2, 1510–1543. doi: 10.3390/sym2031510
- Van de Cruys, S., and Wagemans, J. (2011). Putting reward in art: a tentative prediction error account of visual art. *i-Perception* 2, 1035–1062. doi: 10.1068/i0466aap
- van Elk, M., Arciniegas Gomez, M. A., van der Zwaag, W., van Schie, H. T., and Sauter, D. (2019). The neural correlates of the awe experience: Reduced default mode network activity during feelings of awe. *Human Brain Mapp.* 40, 3561–3574.
- Vartanian, O., and Skov, M. (2014). Neural correlates of viewing paintings: evidence from a quantitative meta-analysis of functional magnetic resonance imaging data. *Brain Cogn.* 87, 52–56. doi: 10.1016/j.bandc.2014.03.004
- Vessel, E. A., and Rubin, N. (2010). Beauty and the beholder: highly individual taste for abstract, but not real-world images. *J. Vis.* 18, 11–14.
- Vessel, E. A., Starr, G. G., and Rubin, N. (2012). The brain on art: intense aesthetic experience activates the default mode network. *Front. Hum. Neurosci.* 6:66. doi: 10.3389/fnhum.2012.00066
- Wainwright, M. J., and Simoncelli, E. P. (2000). Scale mixtures of Gaussians and the statistics of natural images. *Adv. Neural Inform. Process. Syst.* 12, 855–861.
- Wang, T., Mo, L., Mo, C., Tan, L. H., Cant, J. S., Zhong, L., et al. (2015). Is moral beauty different from facial beauty? Evidence from an fMRI study. *Soc. Cogn. Affect. Neurosci.* 10, 814–823. doi: 10.1093/scan/nsu123
- Wilson, A., and Chatterjee, A. (2005). The assessment of preference for balance: introducing a new test. *Empirical Stud. Arts* 23, 165–180. doi: 10.2190/b1lr-mvf3-f36x-xr64
- Wimmer, G. E., Li, J. K., Gorgolewski, K. J., and Poldrack, R. A. (2018). Reward learning over weeks versus minutes increases the neural representation of value in the human brain. *J. Neurosci.* 38, 7649–7666. doi: 10.1523/jneurosci.0075-18.2018
- Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *J. Pers. Soc. Psychol.* 9, 1–27. doi: 10.1037/h0025848

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

Copyright © 2020 Aleem, Correa-Herran and Grzywacz. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Relationships Between Trait Creativity and Resting-State EEG Microstates Were Modulated by Self-Esteem

Xin Wu^{1,2*†}, Jiajia Guo^{1,2†}, Yufeng Wang^{1,2}, Feng Zou^{1,2}, Peifang Guo¹, Jieyu Lv³ and Meng Zhang^{1,2*}

¹ School of Psychology, Xinxiang Medical University, Xinxiang, China, ² Cognitive, Emotional, and Behavioral Lab, Xinxiang Medical University, Xinxiang, China, ³ Department of Psychology, Central University of Finance and Economics, Beijing, China

OPEN ACCESS

Edited by:

Surjo R. Soekadar,
Charité – Universitätsmedizin Berlin,
Germany

Reviewed by:

Pierpaolo Croce,
University of Studies G. d'Annunzio
Chieti and Pescara, Italy
María del Mar Molero,
University of Almería, Spain

*Correspondence:

Xin Wu
uking05@126.com
Meng Zhang
mengzhang.1985@163.com

[†]These authors have contributed
equally to this work

Specialty section:

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

Received: 25 June 2020

Accepted: 13 October 2020

Published: 11 November 2020

Citation:

Wu X, Guo J, Wang Y, Zou F,
Guo P, Lv J and Zhang M (2020) The
Relationships Between Trait Creativity
and Resting-State EEG Microstates
Were Modulated by Self-Esteem.
Front. Hum. Neurosci. 14:576114.
doi: 10.3389/fnhum.2020.576114

Numerous studies find that creativity is not only associated with low effort and flexible processes but also associated with high effort and persistent processes especially when defensive behavior is induced by negative emotions. The important role of self-esteem is to buffer negative emotions, and individuals with low self-esteem are prone to instigating various forms of defensive behaviors. Thus, we thought that the relationships between trait creativity and executive control brain networks might be modulated by self-esteem. The resting-state electroencephalogram (RS-EEG) microstates can be divided into four classical types (MS1, MS2, MS3, and MS4), which can reflect the brain networks as well as their dynamic characteristic. Thus, the Williams Creative Tendency Scale (WCTS) and Rosenberg Self-esteem Scale (RSES) were used to investigate the modulating role of self-esteem on the relationships between trait creativity and the RS-EEG microstates. As our results showed, self-esteem consistently modulated the relationships between creativity and the duration and contribution of MS2 related to visual or imagery processing, the occurrence of MS3 related to cingulo-opercular networks, and transitions between MS2 and MS4, which were related to frontoparietal control networks. Based on these results, we thought that trait creativity was related to the executive control of bottom-up processing for individuals with low self-esteem, while the bottom-up information from vision or visual imagery might be related to trait creativity for individuals with high self-esteem.

Keywords: creativity, self-esteem, resting-state EEG, microstates, trait creativity

INTRODUCTION

Creativity refers to the tendency to imagine and produce something novel (i.e., original) and unexpected, yet still appropriate (i.e., effective and useful) (Sternberg, 1999; Kaufman and Sternberg, 2010). In fact, creativity can be divided into those aspects related to personality and cognition (Rhodes, 1961; Gough, 1976; Amabile, 1996; Runco, 2007; Piffer, 2012). Williams (1969) suggested a cognitive-affective model of creativity and developed a corresponding creativity assessment packet (CAP) (Williams, 1969, 1980). The CAP included a divergent thinking (creative cognition) test and a divergent feeling (trait creativity) test (including four aptitude elements: imagination, risk taking, curiosity, and challenge) (Williams, 1993; Hwang et al., 2007;

Liu et al., 2011). Based on previous findings, trait creativity is a set of aptitudes or personality variables that influence an individual's creativity, while creative cognition refers to cognitive processes and metacognitive strategies during creative production, such as divergent thinking (Satzinger et al., 1999; Zeng et al., 2009). Although it is emphasized that creativity is a function of flexibility, creativity can be also achieved through persistence, which means that creative productions can be acquired by hard work, perseverance, and exploration of a few cognitive categories or perspectives (Schooler et al., 1993; Finke, 1996; Simonton, 1997; Dietrich, 2004; Nijstad et al., 2010).

Some cognitions of creativity had been confirmed by using divergent thinking tasks. It had been argued that generation of creative ideas require associative processes which include processes of freely and spontaneously forming associations between elements, as well as controlled processes which include inhibiting unsuitable ideas and evaluating and selecting creative ideas (Bendetowicz et al., 2017; Benedek et al., 2017). Beaty et al. (2017) also suggested that creative cognitions require the dynamic interactions between default and cognitive control networks, which reflected that both bottom-up and top-down processes are necessary to generate creative ideas. These opinions might fit with the blind variation and selective retention of creativity (Campbell, 1960; Simonton, 2011; Beaty et al., 2017), which implies the associative processes for blind variation and the controlled processes for selective retention. Other studies also found that the suppression of bottom-up irrelevant information is necessary when semantic information is retrieved and integrated to generate creative ideas (Fink et al., 2009, 2010; Wu et al., 2015). Sternberg (1999) suggested that trait creativity can have an impact on creative problem-solving ability. Individuals with certain creative traits (e.g., curiosity and imagination) can be more creative than those without these characteristics (Oldham and Cummings, 1996; Feist, 1998; Piffer, 2012). Moreover, creative individuals had higher gray matter volume in the right posterior middle temporal gyrus (related to representations of semantic concepts) (Li et al., 2015). On these perspectives, trait creativity might be related to the neural networks found in the creative cognitions to some extent.

Self-esteem is an attitude based on positive and negative self-evaluations (Rosenberg, 1965) and reflects the positive aspect of self-concept (Campbell et al., 1996). Individuals with high self-esteem tend to believe themselves to be capable and worthy, so they are more likely to express ideas that differ from others and are more willing to share creative ideas (Thatcher and Brown, 2010). The generative and flexible thinking associated with creativity can aid in successfully crafting self-serving justifications that allow individuals to maintain positive self-views (Carson et al., 2003; Gino and Ariely, 2012; Antinori et al., 2017). However, terror management theory suggests that the important role of self-esteem is to buffer anxieties induced by social threat, such as death threats (Greenberg et al., 1997; Pyszczynski, 2004) and negative feedback (Brown, 2010). Individuals with low self-esteem are prone to instigating various forms of defensive behavior to bolster their self-worth (Pyszczynski, 2004). Moreover, creativity can be achieved by persistence when defensive behavior is induced by negative mood

states (such as fear and anxiety) (Baas et al., 2011; Roskes et al., 2012). Thus, creativity might be achieved by the function of flexibility for high-self-esteem individuals, while creativity might be achieved by the function of persistence for low-self-esteem individuals. It had been suggested that flexibility is associated with low effort, low resource demands, high speed, and efficient processing (Evans, 2003; Winkielman et al., 2003; Dietrich, 2004; De Dreu et al., 2008; Oppenheimer, 2008), while persistence is associated with high effort, perseverance, and a slower speed of operation (Evans, 2003; Winkielman et al., 2003; De Dreu et al., 2008). Thus, trait creativity might be related to controlled processes for individuals with low self-esteem, while it is related to associative processes for individuals with high self-esteem.

Previous studies had confirmed that functional networks can be depicted by spontaneous brain activities (Raichle and Mintun, 2006; Raichle, 2010), which might imply that the influence of self-esteem on trait creativity might be investigated by analyzing brain activity under the resting state. It had been found that some functional networks were confirmed by using the resting-state functional magnetic resonance imaging (RS-fMRI), such as default modal network, attentional network, salient network, and visual network (Andrews-Hanna et al., 2006; Raichle and Mintun, 2006; Fox et al., 2007). In addition, spontaneous brain activities are also investigated by the resting-state electroencephalogram (RS-EEG), where the RS-EEG microstates are used to depict the brain networks by using the signal from all electrodes (Dierks et al., 1997; Stevens and Kircher, 1998; Lehmann et al., 2005, 2010; Kikuchi et al., 2011; Schlegel et al., 2012). RS-EEG microstates are also seen as the "atoms of thought" and can be divided into four typical microstates (Lehmann et al., 1998; Khanna et al., 2014). When the evidences from RS-fMRI and RS-EEG are combined, the relevant brain networks are confirmed, which indicates that MS1 was related to the bilateral superior temporal gyrus and middle temporal gyrus, which were linked to semantic processes or phonological processing; MS2 was associated with the extrastriate cortex, which might be related to visual processing and visual images; MS3 was associated with positive BOLD activation in cingulo-opercular brain networks, which were related to salient or attention control; and MS4 was associated with right-lateralized dorsal and ventral attentional networks (Lehmann et al., 1987; Britz et al., 2010; Musso et al., 2010; Yuan et al., 2012).

Considering the high time resolution of EEG, RS-EEG microstates can also provide more dynamic characteristics of the brain networks relative to RS-fMRI. Specifically, duration is the time coverage of each microstate; occurrence is the average number of occurrences per microstate in a second; contribution is the total duration of each microstate, accounting for the total resting EEG duration; the possibility of transition between any two microstates is related to the information flow between them (Britz et al., 2010; Khanna et al., 2014; Gao et al., 2017). Moreover, the characteristics are related to the altered mental states under experimental conditions. Seitzman et al. (2017) found that the occurrence and contribution of MS2 and the duration of MS1 were modulated by the eye-open or eye-close condition; the occurrence and contribution of MS4 were increased under attentional tasks; the transition between MS3

and MS1 was also decreased under attentional tasks. Zappasodi et al. (2019) also found that the microstates related to visual (MS2) and default-mode network (MS3) were modulated by visuospatial tasks, which reflect that the contribution of MS2 was significantly increased under visuospatial tasks, while the contribution of MS2 was significantly decreased. Moreover, Santarnecchi et al. (2017) found that the RS-EEG microstates of MS2 and MS3 were related to fluid intelligence, where they found that the occurrences of MS2 and MS3 were significantly negatively related to fluid intelligence, and the contribution of MS2 was negatively associated with the increase of fluid intelligence after training it.

According to previous studies (De Dreu et al., 2008; Bendetowicz et al., 2017; Benedek et al., 2017), we speculated that trait creativity for individuals with low or high self-esteem might rely on different cognitive processes. Specifically, trait creativity might be related to the controlled processes for individuals with low self-esteem, while it is related to associative processes for individuals with high self-esteem. Thus, relationships between trait creativity and RS-EEG microstates might be modulated by self-esteem. In the present study, the modulated roles of self-esteem in the relationships between creativity and RS-EEG microstates were investigated by using the Williams Creative Tendency Scale (WCTS) and Rosenberg Self-esteem Scale (RSES) to measure the creativity and self-esteem, respectively, and combining the RE-EEG microstate analysis. Previous studies considered that the RS-EEG microstates were calculated based on the alpha band activities and with the inhibition of modality-specific processing, increasing the characteristics of MS2 and MS3 (Milz et al., 2016; Santarnecchi et al., 2017). Thus, we hypothesize that the temporal characteristics of sensory input (such as MS1 and MS2) might be positively related to trait creativity for individuals with low self-esteem, which make them inhibit bottom-up irrelevant information. However, these characteristics might be negatively related to trait creativity for individuals with high self-esteem, which make them generate more associations. In addition, the possibility of transitions between top-down control system (MS3 or MS4) and sensory input (such as MS1 and MS2) might be higher for those with low self-esteem relative to high self-esteem, which makes it easier for them to control the bottom-up irrelevant information.

MATERIALS AND METHODS

Subjects

Three hundred thirty-six right-handed subjects recruited in Xinxiang Medical University (72% male, 28% female; mean age = 18.3, SD = 0.84) participated in the study. Subjects had no history of neurological or psychiatric disease and did not take any medication that could affect the experiment. All participants gave written informed consent to participate in the study which was approved by the ethics committee of Xinxiang Medical University. One subject's data were deleted due to data record error.

Materials

Williams Creative Tendency Scale

The WCTS was used to measure trait creativity (revised by Lin Xingtai of Taiwan Normal University). The WCTS is composed of 50 items, and the subjects were asked to respond to a 3-point Likert-type scale ranging from 1 (totally disagree) to 3 (totally agree). According to Williams (1994), WCTS can be divided into four subscales, namely, curiosity (13 items; e.g., "I would like to know what other people think"), imagination (13 items; e.g., "If the final page of a storybook is missing, I will make up the story's ending myself"), challenge (12 items; e.g., "I like unusual things"), and risk taking (12 items; e.g., "Trying a new game or activity is an interesting thing"). Reliability analysis showed that the reliability coefficients of the total score of the scale were between 0.569 and 0.678. In this study, the alpha reliability for the WCTS was 0.866 according to our sample.

Rosenberg Self-Esteem Scale

Participants completed a measure of self-esteem: RSES (Rosenberg, 1965). The scale is a self-assessment measure of self-esteem commonly used at home and abroad, which consists of 10 items. All 10 items are rated on a 4-point scale ranging from 1 (not very true of me) to 4 (very true of me). On a scale of 10 to 40, higher scores indicate higher levels of self-esteem and self-acceptance. Previous studies have reported alpha reliability for the RSES ranging from 0.72 to 0.88 (Gray-Little and Carels, 1997). In this study, the alpha reliability for the RSES was 0.816 according to our sample.

RS-EEG Data Acquisition

During RS-EEG recording (6 min in duration), subjects were asked to open their eyes and focus on the "+" appearing in the center of the screen quietly without moving their body or head. The RS-EEG data were recorded by using the Neuro Scan Product with 64 Ag-AgCl scalp sites according to the international 10–20 system in an elastic cap. During recording, all electrodes were referenced to Cz and re-referenced off-line to linked mastoids. Channels for horizontal and vertical EEG were computed off-line from electrodes recorded from the outer canthi of the eyes and from above and below the right eye, respectively. Electrode impedance was kept below 5 k Ω . EEG was sampled online with a frequency of 500 Hz DC amplifiers with a band-pass filter of 0.1–100 Hz.

RS-EEG Microstate Preprocessing

The EEG data were preprocessed using EEGLAB¹ in MATLAB 2018b². Data were filtered off-line by a band-pass filter of 2–20 Hz and were run through an independent component analysis (ICA). Artifacts produced by blinks, eye movements, eye drift, head movements, power-line interference, or electrocardiograph were rejected. The artifact-free data were recomputed against the average, according to previous studies (Lehmann et al., 1987, 2005; Koenig et al., 1999; Gao et al., 2017). Then the data were segmented into 180 epochs with an epoch length of

¹<https://sccn.ucsd.edu/eeqlab/index.php>

²<http://cn.mathworks.com/>

2,000 ms. EEG epochs with amplitude values exceeding $\pm 80 \mu\text{V}$, which might be contaminated by strong muscle artifacts, were manually rejected.

The RS-EEG microstates were calculated according to previous studies (Lehmann et al., 1987, 2005; Koenig et al., 1999; Gao et al., 2017). First, the global field power (GFP), which was defined as the EEG potential variance across scalp electrodes, was calculated, and only the topographies at peaks of the GFP were further analyzed. Then, based on previous studies (Tibshirani and Walther, 2005), atomize–agglomerate hierarchical clustering (AAHC) was performed to analyze the microstates with the polarity of each topographical map being disregarded, which was a modified k -means to create unique clusters for EEG microstate analysis. After that, a cross-validation criterion was used to identify the optimal number of template maps, which was the best solution to find the minimal number of template maps explaining the maximal variance in cluster analysis (Pascual-Marqui et al., 1995; Britz et al., 2010). According to our data, four clusters (MS1, MS2, MS3, and MS4) were found, and the explained variance was 0.786 ± 0.033 (see **Figure 1**), which was the same as that found in most studies of RS-EEG microstate (e.g., Lehmann et al., 1987, 2005; Britz et al., 2010; Gao et al., 2017). The global map dissimilarity (GMD) was used as a criterion to fit all original maps of each subject into the four prototype maps, where each time point was fitted and labeled with the one cluster it correlated best (Van de Ville et al., 2010). Finally, the labeled data were used to compute the temporal characteristics, namely, duration, occurrence, and contribution of each microstate, as well as the probability of transition between them.

Statistical Analysis

The total scores of the WCTS and RSES were imported into SPSS for correlation analysis, and the correlation between creativity and self-esteem as well as the correlation between the four dimensions of creativity (risk taking, challenge, curiosity, and imagination) and self-esteem was obtained. Then, the total score of the scales (WCTS and RSES) and the data of the duration, occurrence, and contribution of four RS-EEG microstates were imported into SPSS to analyze the relationship between creativity and microstates and between self-esteem and microstates. Finally, the modulating role of RSES in the relationship between the total score of WCTS and the duration, occurrence, and contribution of microstates and the transitions between them were analyzed using Model 1 of PROCESS 3.0 (Hayes, 2018) with a statistical threshold of $p < 0.05$ (FDR corrected).

RESULTS

The Relationship Between Self-Esteem and Creativity

Since some data were more than three standard deviations from the mean, we chose to exclude them and ended up with data from 334 subjects available. The total score of RSES is significantly positively correlated with the total score of WCTS ($r = 0.262$, $p < 0.001$) (see **Figure 2**), as well as the four subscales of WCTS [risk taking ($r = 0.294$, $p < 0.001$), challenge ($r = 0.316$, $p = 0.001$),

TABLE 1 | The relationships between self-esteem and creativity ($n = 334$).

	Risk taking	Curiosity	Imagination	Challenge	Creativity
SES	0.2949 (<0.001)	0.186 (0.001)	0.112 (0.040)	0.316 (<0.001)	0.262 (<0.001)

The bolded and italic values mean that the relationship was significant at $p < 0.05$ with Bonferroni correction.

and curiosity ($r = 0.186$, $p < 0.001$), imagination ($r = 0.112$, $p < 0.05$) (see **Table 1**). In addition, the original score of the AU task was not significantly correlated to the total score of RSES.

The Relationship Between Self-Esteem, Creativity, and Microstates

Through the correlation analysis of creativity and its different dimensions with various types of microstates, we found no significant correlation between creativity and microstates. On a regular basis, we also analyzed the relationship between self-esteem and various types of microstates; however, there was no significant correlation. Further analysis had shown that age and gender had no effect on the relationships.

The Modulating Role of RSES in Creativity of WCTS

After sex and age were controlled for, the interaction of the total score of RSES \times the mean duration of MS2 was significant [$F(1,329) = 17.691$, $p < 0.001$, $\Delta R^2 = 0.046$] (see **Table 2**). Simple slope analysis results showed that when the total RSES was lower (mean -1 sd), the total score of WCTS was significantly positively correlated with the duration of MS2 ($\beta = 0.218$, $t = 3.192$, $p < 0.005$) and that when the total RSES score was higher (mean $+1$ sd), the total score of WCTS was significantly negatively correlated with the duration of MS2 ($\beta = -0.204$, $t = -2.650$, $p < 0.01$). Johnson–Neyman results showed that the total score of WCTS was significantly positively correlated with the duration of MS2 when the total score of RSES was below 27 (mean -0.486 sd, 31.04% of our sample) and that the total score of WCTS was significantly negatively correlated with the duration of MS2 when the total score of RSES was above 30 (mean $+0.580$ sd, 23.88% of our sample) (see **Figure 3**).

After sex and age were controlled for, the interaction of the total score of RSES \times the mean contribution of MS2 was significant [$F(1,329) = 9.598$, $p < 0.005$, $\Delta R^2 = 0.026$] (see **Table 2**). Simple slope analysis results showed that when the total RSES was lower (mean -1 sd), the total score of WCTS was significantly positively correlated with the contribution of MS2 ($\beta = 0.211$, $t = 2.835$, $p = 0.005$) and that when the total RSES score was higher (mean $+1$ sd), the total score of WCTS was not significantly correlated with the contribution of MS2 ($\beta = -0.095$, $t = -1.376$, $p = 0.170$). Johnson–Neyman results showed that the total score of WCTS was significantly positively correlated with the contribution of MS2 when the total score of RSES was below 27 (mean -0.486 sd, 31.04% of our sample) and that the total score of WCTS was significantly negatively correlated with the contribution of MS2 when the total score of RSES was above 32 (mean $+1.414$ sd, 14.63% of our sample) (see **Figure 3**).

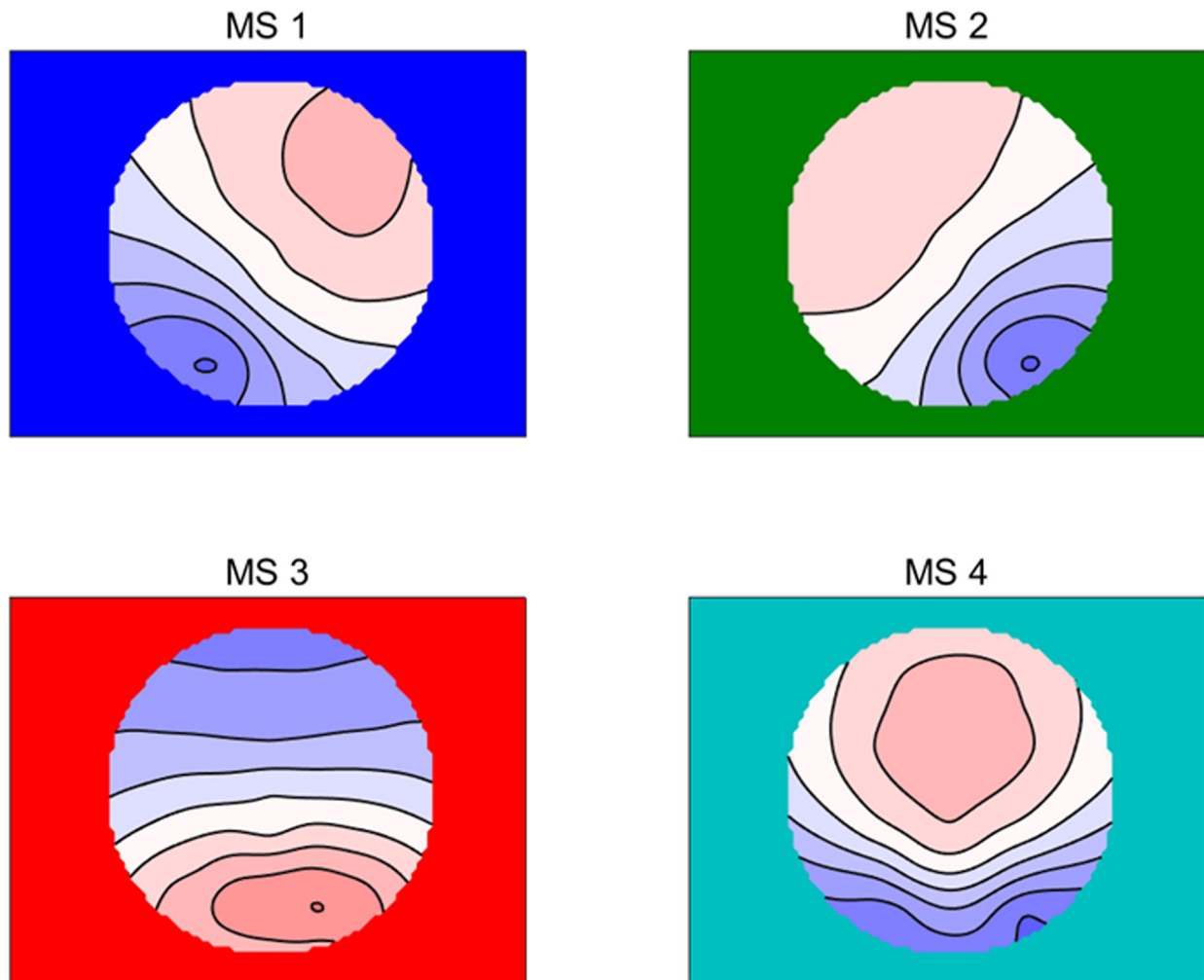


FIGURE 1 | The four microstate topographic maps are RS-EEG microstate Type A (MS1), Type B (MS2), Type C (MS3), and Type D (MS4).

After sex and age were controlled for, the interaction of the total score of RSES \times the mean occurrence of MS3 was significant [$F(1,329) = 8.061, p = 0.005, \Delta R^2 = 0.022$] (see **Table 2**). Simple slope analysis results showed that when the total RSES was lower (mean -1 sd), the total score of WCTS was significantly negatively correlated with the occurrence of MS3 ($\beta = -0.179, t = -2.417, p < 0.05$); when total RSES was higher (mean $+1$ sd), the total score of WCTS was not significantly correlated with the occurrence of MS3 ($\beta = 0.148, t = -2.417, p = 0.070$). Johnson–Neyman results showed that the total score of WCTS was significantly negatively correlated with the occurrence of MS3 when the total score of RSES was below 25 (mean -0.643 sd, 17.61% of our sample) and that the total score of WCTS was significantly positively correlated with the occurrence of MS3 when the total score of RSES was above 32 (mean $+1.414$ sd, 14.63% of our sample) (see **Figure 3**).

After sex and age were controlled for, the interaction of the total score of RSES \times the possibility of transition from MS2 to MS4 [possibility ($MS2toMS4$)] was significant [$F(1,330) = 10.122,$

$p < 0.005, \Delta R^2 = 0.028$] (see **Table 2**). Simple slope analysis results showed that when the total RSES was lower (mean -1 sd), the total score of WCTS was significantly positively correlated with possibility ($MS2toMS4$) ($\beta = 0.239, t = 2.952, p < 0.005$) and that when total RSES was higher (mean $+1$ sd), the total score of WCTS was not significantly correlated with possibility ($MS2toMS4$) ($\beta = -0.087, t = -1.217, p = 0.224$). Johnson–Neyman results showed that the total score of WCTS was significantly positively correlated with the duration of MS2 when the total score of RSES was below 27 (mean -0.486 sd, 31.04% of our sample) and that the total score of WCTS was significantly negatively correlated with the duration of MS2 when the total score of RSES was above 32 (mean $+1.414$ sd, 14.63% of our sample) (see **Figure 4**).

After sex and age were controlled for, the interaction of the total score of RSES \times the possibility of transition from MS4 to MS2 [possibility ($MS4toMS2$)] was significant [$F(1,329) = 6.766, p < 0.01, \Delta R^2 = 0.018$] (see **Table 2**). Simple slope analysis results showed that when the total RSES was lower (mean -1 sd), the total score of WCTS was significantly positively correlated with

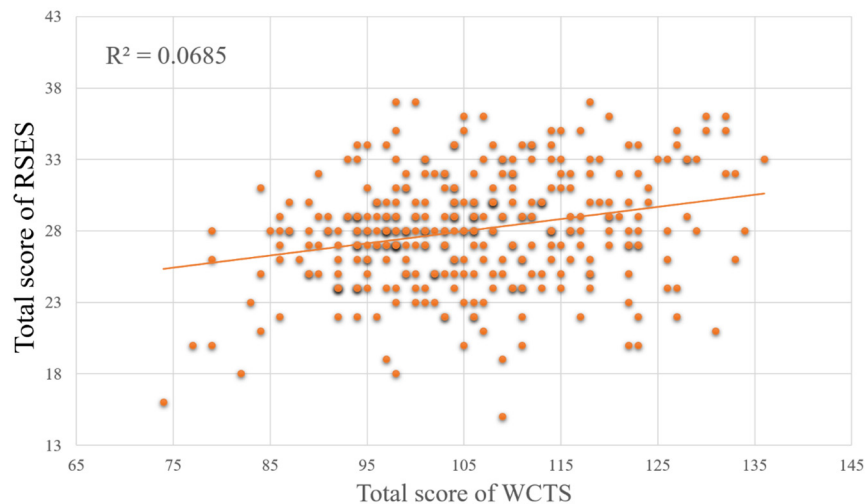


FIGURE 2 | The relationships between the total score of WCTS and total score of RSES.

TABLE 2 | The interaction parameters of RS-EEG microstate and total score of RSES when predicting total score of WCTS after controlling sex and age.

	ΔR^2	df1	df2	F-value	p (Uncorrected)	p (FDR-corrected)
Duration						
MS1	0.004	1	329	1.548	0.214	0.278
MS2	0.046	1	329	17.691	< 0.001	0.005
MS3	0.007	1	329	2.440	0.119	0.226
MS4	0.007	1	329	2.411	0.122	0.226
Occurrence						
MS1	0.014	1	329	5.090	0.024	0.082
MS2	0.002	1	329	0.060	0.807	0.807
MS3	0.022	1	329	8.061	0.005	0.030
MS4	0.012	1	329	4.372	0.037	0.111
Contribution						
MS1	0.007	1	329	2.397	0.123	0.226
MS2	0.026	1	329	9.598	0.002	0.024
MS3	0.003	1	329	1.122	0.290	0.316
MS4	0.001	1	329	0.299	0.585	0.610
Transition						
MS1 to MS2	0.005	1	329	1.623	0.204	0.278
MS1 to MS3	0.004	1	329	1.274	0.260	0.297
MS1 to MS4	0.016	1	329	5.887	0.016	0.064
MS2 to MS1	0.004	1	329	1.514	0.220	0.278
MS2 to MS3	0.004	1	329	1.295	0.256	0.297
MS2 to MS4	0.023	1	329	8.652	0.004	0.030
MS3 to MS1	0.009	1	329	3.178	0.076	0.182
MS3 to MS2	0.006	1	329	2.202	0.139	0.226
MS3 to MS4	0.005	1	329	1.633	0.202	0.278
MS4 to MS1	0.010	1	329	3.739	0.054	0.144
MS4 to MS2	0.018	1	329	6.766	0.009	0.047
MS4 to MS3	0.006	1	329	2.179	0.141	0.226

The bold values are significant in $p < 0.05$ with FDR-corrected.

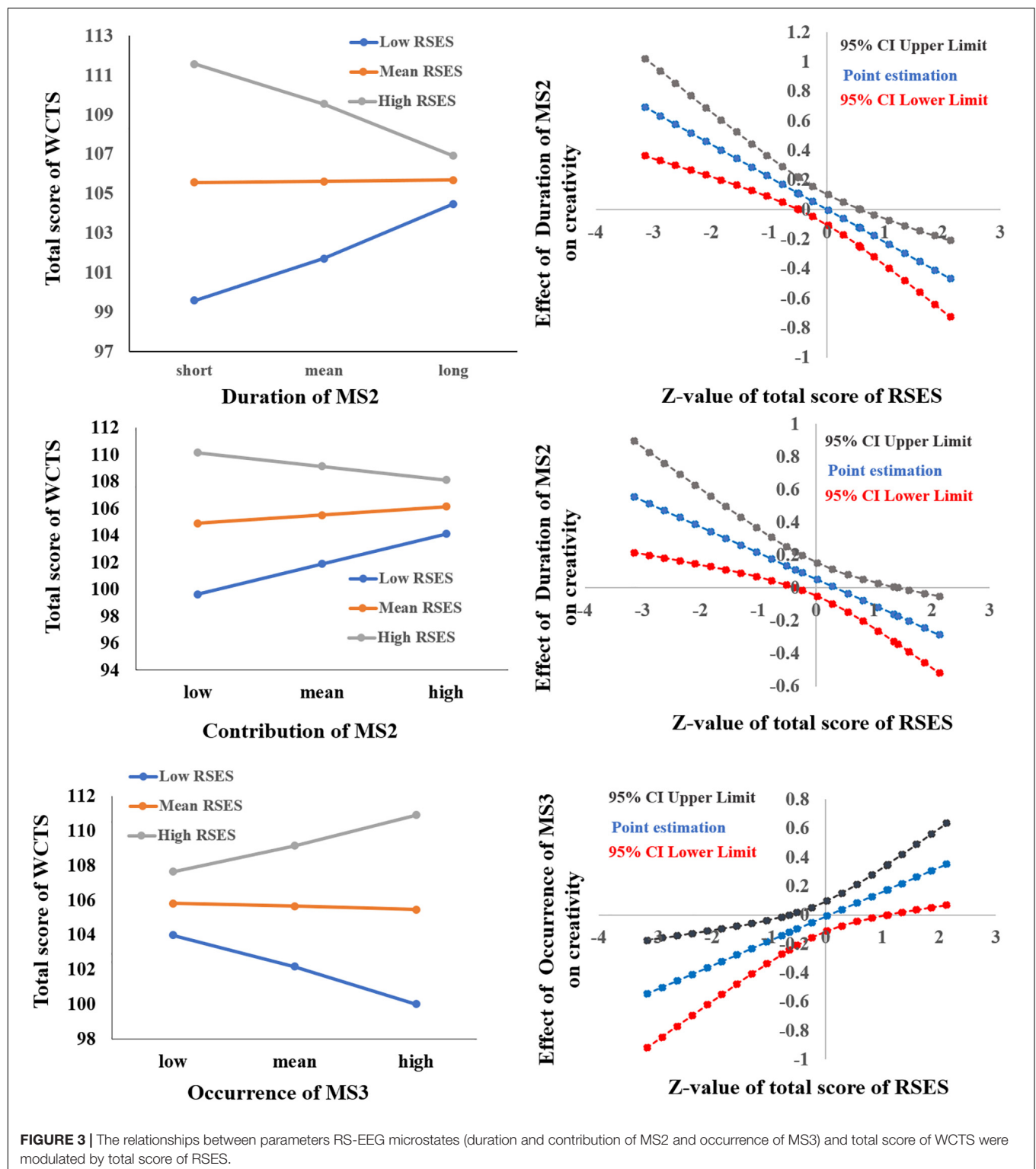
possibility (MS4toMS2) ($\beta = -0.203$, $t = 2.527$, $p < 0.05$) and that when total RSES was higher (mean + 1 sd), the total score of WCTS was not significantly correlated with possibility (MS4toMS2)

($\beta = -0.082$, $t = -1.151$, $p = 0.251$). Johnson–Neyman results showed that the total score of WCTS was significantly negatively correlated with the duration of MS2 when the total score of RSES was below 27 (mean – 0.405 sd, 31.04% of our sample) and that the total score of WCTS was significantly negatively correlated with the duration of MS2 when the total score of RSES was above 36 (mean + 1.956 sd, 2.39% of our sample) (see Figure 4).

DISCUSSION

In the present study, the modulating role of self-esteem in the relationship between creativity and RS-EEG microstates was investigated using WCTS and RSES combined with RS-EEG microstate analysis. Consistent with the previous studies, this experiment also proved the positive correlation between trait creativity and self-esteem (Jaquish and Ripple, 1981; Goldsmith and Matherly, 1988; Yau, 2011). Importantly, the RS-EEG microstate results showed that RSES could modulate the relationship between WCTS creativity and the duration and contribution of MS2, the occurrence of MS3, and the possibility (MS4toMS2).

In the previous series of studies, it had been suggested that a wide breadth of attention could facilitate creative performance (Mendelsohn and Griswold, 1964, 1966; Mendelsohn and Lindholm, 1972; Mendelsohn, 1976), which means that the greater the number and range of stimuli attended to at any one time, the more chances there are to generate creative ideas (Kasof, 1997; Memmert and Roth, 2007). Previous studies from visual attention found that the activation of the visual cortex (striated and extrastriated cortex) could be influenced by visual attention, which means that the activation of the visual cortex can be modulated by operating both through the facilitation of visual processing at the attended location and through inhibition of unattended stimulus representations (Slotnick et al., 2003).



Moreover, it had been found that the extrastriate cortex was also activated when creative tasks (such as alternative uses) were performed (Fink et al., 2009, 2010) and that gray matter density in the visual cortex was positively correlated with creativity (Fink et al., 2014; Wu et al., 2015). Previous studies had found that MS2

was negatively associated with the activation of the extrastriate cortex, which might imply that individuals with a short duration of MS2 possess a stronger function of visual processing or visual images (Britz et al., 2010; Khanna et al., 2015; Gao et al., 2017). Therefore, for individuals with higher self-esteem, the duration

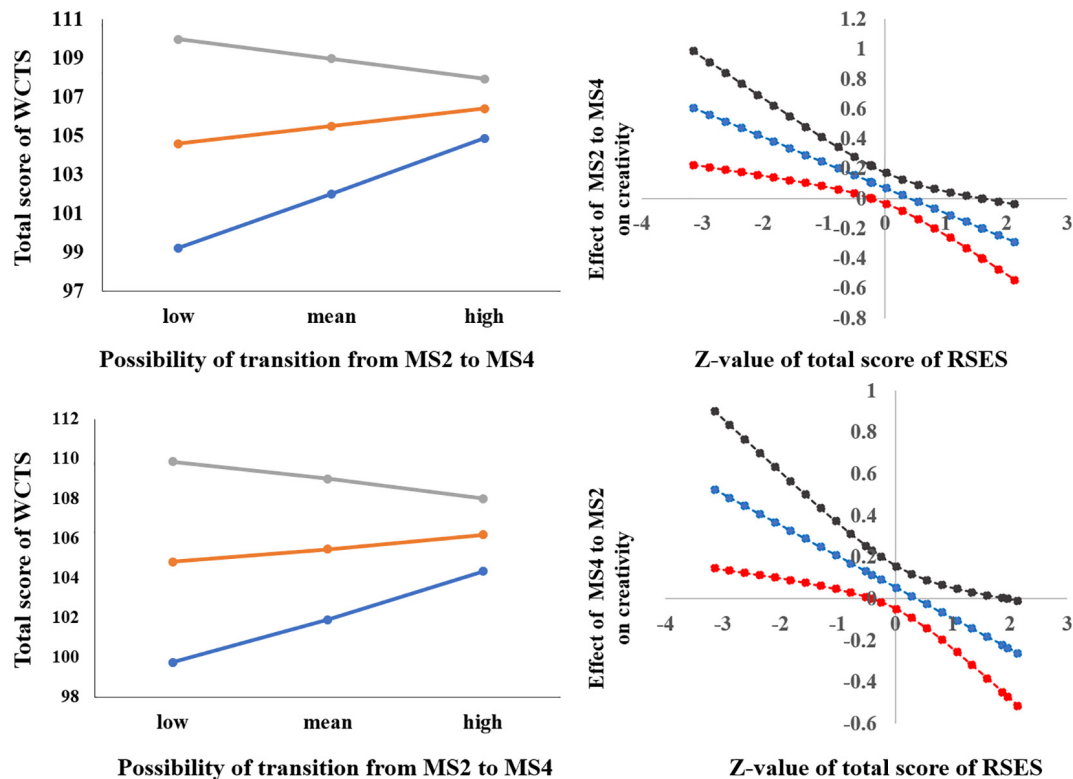


FIGURE 4 | The relationships between parameters RS-EEG microstates (transition from MS2 to MS4 and transition from MS4 to MS2) and total score of WCTS were modulated by total score of RSES.

of MS2 was negatively correlated with the total score of WCTS, which might reflect that a strong function of the visual cortex could make individuals attend more elements at one time and make individuals more creative.

According to the model of a dual pathway to creativity, creative ideas can be generated by the functions of flexibility and persistence (Nijstad et al., 2010). It had been suggested that inhibition of irrelevant bottom-up cognitive processes was required for creativity (e.g., Fink et al., 2009, 2010; Wu et al., 2015), especially when the function of persistence was induced under threat conditions (Baas et al., 2011; Roskes et al., 2012). According to terror management theory (Greenberg et al., 1997) and sociometer theory (Leary et al., 1995), the important role of self-esteem is to buffer negative emotions induced by death threats (death anxiety) or social threats (social rejection). Thus, bottom-up cognitive processes might be a disadvantage to creativity for individuals with low self-esteem. Consistent with this opinion, our results showed that the duration of MS2 was positively correlated with the total score of WCTS when individuals have low self-esteem. Therefore, we thought that irrelevant bottom-up cognitive processes might be more prone to being inhibited as the duration of MS2 increases, which is good for individuals with low self-esteem as this enables them to generate creative ideas.

Now that inhibition of irrelevant bottom-up cognitive processes is required for creativity (e.g., Fink et al., 2009, 2010;

Wu et al., 2015), the transitions between MS4 (executive control) and MS2 (visual processes) found in this study might also reflect that persistence is needed for creativity in individuals with low self-esteem. Previous studies had found that MS4 was related to right-lateralized frontoparietal networks, which might be related to dorsal and ventral attention networks (Britz et al., 2010). It had been confirmed that the right dorsal frontal-parietal networks were involved in top-down control, while the ventral frontal-parietal networks were related to information-capture attention in the bottom-up manner (Cabeza et al., 2008, Cabeza et al., 2014). It was further found that the ventral frontal-parietal networks were related to the phasic and adaptive aspects of cognitive control (moment-to-moment executive control), while dorsal frontal-parietal networks were related to top-down selective attention to specific stimulus features (Sadaghiani et al., 2010, 2012; Sadaghiani and Kleinschmidt, 2016). According to our results, the possibility of transitions between MS2 and MS4 was positively correlated to the total score of WCTS for individuals with low self-esteem; at the same time, the possibility of transition from MS2 to MS4 was negatively correlated to the total score of WCTS for individuals with high self-esteem. Thus, the trait creativity for individuals with low self-esteem might be dependent on the moment-to-moment information to attention in a bottom-up manner, but the trait creativity for individuals with high self-esteem might be dependent on the top-down selective attention to specific stimulus features.

The pursuit of self-esteem is a fundamental human need (Taylor and Brown, 1988; Solomon et al., 1991), but the consequences of pursuing self-esteem may produce the risk of failure in verifying individuals' abilities, qualities, and self-worth and make them experience uncertainty (Crocker and Park, 2004). Moreover, individuals who chronically experience real or imagined rejection are prone to developing lower self-esteem relative to individuals feeling accepted and included in their social environment (Dandeneau and Baldwin, 2004). It had been found that the dorsal anterior cingulate cortex (dACC), the right ventral lateral prefrontal cortex (rVLPFC), and the anterior insular (AI) were more active during rejection than during inclusion (Eisenberger et al., 2003, 2007; Slavich and Epel, 2010; Masten et al., 2011; Rotge et al., 2014). Moreover, RS-EEG MS3 was related to the cingulo-opercular networks, which include the dACC and insula (Katayama et al., 2007; Seeley et al., 2007; Britz et al., 2010). Therefore, the negative effect of occurrence of RS-EEG MS3 on the trait creativity for individuals with low self-esteem might reflect that with the functions of cingulo-opercular networks increasing, individuals might be prone to being influenced by social threat and make them develop lower trait creativity.

This experiment investigated the modulating effect of self-esteem on creativity and RS-EEG microstates. The findings suggest that self-esteem modulates the relationship between creativity and the duration and contribution of MS2, the occurrence of MS3, and the possibility (MS4toMS2). Based on these results, we thought trait creativity was related to automatic or bottom-up cognitive processes for individuals with high self-esteem, while inhibition of irrelevant information could facilitate creativity for individuals with low self-esteem. Moreover, social threat experiences might have a detrimental effect on creativity for individuals with low self-esteem. Though there were some important and robust evidences for us to understand the relationships between creativity and RS-EEG microstates, several limitations should be considered. Firstly, only sex and age were controlled, and some other potential factors for creativity (such as intelligence and personality) were not controlled. Secondly, complex cognitive processes could be related to

creativity; however, only some of them have been reflected by our results. Thirdly, due to undergraduates being selected in this study, it might be cautious to explore other groups with different ages (such as children and old adults). Therefore, more detailed experiments and advanced paradigms should be used in future studies to determine the cognitive meanings of each microstate to further investigate the relationships between microstates and creativity.

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 Xinxiang Medical University ethics committee. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

XW, JG, and MZ designed the experiment. JG wrote this manuscript. YW, FZ, and PG collected and analyzed the data. XW, JL, and MZ revised a manuscript. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (31600927 and 81830040), the Youth Foundation of Social Science and Humanity, China Ministry of Education (19YJCZH179), and the Key Scientific Research Projects of Colleges and Universities in Henan Province (20A190001).

REFERENCES

- Amabile, T. M. (1996). *Creativity in Context*. New York, NY: Westview.
- Andrews-Hanna, J. R., Smallwood, J., and Spreng, R. N. (2006). The default network and self-generated thought: component processes, dynamic control, and clinical relevance. *Ann. N.Y. Acad. Sci.* 1316, 29–52. doi: 10.1111/nyas.12360
- Antinori, A., Carter, O. L., and Smillie, L. D. (2017). Seeing it both ways: Openness to experience and binocular rivalry suppression. *J. Res. Pers.* 68, 15–22. doi: 10.1016/j.jrp.2017.03.005
- Baas, M., De Dreu, C. K., and Nijstad, B. A. (2011). When prevention promotes creativity: the role of mood, regulatory focus, and regulatory closure. *J. Pers. Soc. Psychol.* 100:794. doi: 10.1037/a0022981
- Beatty, R. E., Christensen, A. P., Benedek, M., Silvia, P. J., and Schacter, D. L. (2017). Creative constraints: brain activity and network dynamics underlying semantic interference during idea production. *Neuroimage* 148, 189–196. doi: 10.1016/j.neuroimage.2017.01.012
- Bendetowicz, D., Urbanski, M., Garcin, B., Foulon, C., Levy, R., Bréchemier, M. L., et al. (2017). Two critical brain networks for generation and combination of remote associations. *Brain* 1, 217–233. doi: 10.1093/brain/awx294
- Benedek, M., Kenett, Y. N., Umdasch, K., Anaki, D., Faust, M., Neubauer, A. C., et al. (2017). How semantic memory structure and intelligence contribute to creative thought: a network science approach. *Think. Reason.* 23, 1–26.
- Britz, J., Ville, D. V. D., and Michel, C. M. (2010). Bold correlates of eeg topography reveal rapid resting-state network dynamics. *Neuroimage* 52, 1162–1170. doi: 10.1016/j.neuroimage.2010.02.052
- Brown, J. D. (2010). High self-esteem buffers negative feedback: once more with feeling. *Cogn. Emot.* 24, 1389–1404. doi: 10.1080/02699930903504405
- Cabeza, R., Ciaramelli, E., Olson, I. R., and Moscovitch, M. (2008). The parietal cortex and episodic memory: an attentional account. *Nat. Rev. Neurosci.* 9:613. doi: 10.1038/nrn2459
- Cabeza, R., Mazuz, Y. S., Stokes, J., Kragel, J. E., Woldorff, M. G., Ciaramelli, E., et al. (2014). Overlapping parietal activity in memory and perception: evidence for the attention to memory model. *J. Cogn. Neurosci.* 23, 3209–3217. doi: 10.1162/jocn_a_00065

- Campbell. (1960). Blind variation and selective retention in creative thought as in other knowledge processes. *Psychol. Rev.* 67:380. doi: 10.1037/h0040373
- Campbell, J. D., Trapnell, P. D., Heine, S. J., Katz, I. M., Lavallee, L. F., and Lehman, D. R. (1996). Self-concept clarity: measurement, personality correlates, and cultural boundaries. *J. Pers. Soc. Psychol.* 70:141. doi: 10.1037/0022-3514.70.1.141
- Carson, S. H., Peterson, J. B., and Higgins, D. M. (2003). Decreased latent inhibition is associated with increased creative achievement in high-functioning individuals. *J. Pers. Soc. Psychol.* 85, 499–506. doi: 10.1037/0022-3514.85.3.499
- Crocker, J., and Park, L. E. (2004). The costly pursuit of self-esteem. *Psychol. Bull.* 130:392. doi: 10.1037/0033-2909.130.3.392
- Dandeneau, S. D., and Baldwin, M. W. (2004). The inhibition of socially rejecting information among people with high versus low self-esteem: the role of attentional bias and the effects of bias reduction training. *J. Pers. Soc. Psychol.* 23, 584–603. doi: 10.1521/jscp.23.4.584.40306
- De Dreu, C. K. W., Baas, M., and Nijstad, B. A. (2008). Hedonic tone and activation level in the mood-creativity link: toward a dual pathway to creativity model. *J. Pers. Soc. Psychol.* 94, 739–756. doi: 10.1037/0022-3514.94.5.739
- Dierks, T., Jelic, V., Julin, P., Maurer, K., Wahlund, L. O., Almkvist, O., et al. (1997). EEG-microstates in mild memory impairment and Alzheimer's disease: possible association with disturbed information processing. *J. Neural. Transm.* 104, 483–495. doi: 10.1007/bf01277666
- Dietrich, A. (2004). The cognitive neuroscience of creativity. *Psychon. Bull. Rev.* 11, 1011–1026.
- Eisenberger, N. I., Lieberman, M. D., and Williams, K. D. (2003). Does rejection hurt? an fMRI study of social exclusion. *Science* 302, 290–292. doi: 10.1126/science.1089134
- Eisenberger, N. I., Way, B. M., Taylor, S. E., Welch, W. T., and Lieberman, M. D. (2007). Understanding genetic risk for aggression: clues from the brain's response to social exclusion. *Biol. Psychiatry* 61, 1100–1108. doi: 10.1016/j.biopsych.2006.08.007
- Evans, J. S. B. T. (2003). In two minds: dual-process accounts of reasoning. *Trends Cogn. Sci.* 7, 454–459. doi: 10.1016/j.tics.2003.08.012
- Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Pers. Soc. Psychol. Rev.* 2, 290–309. doi: 10.1207/s15327957pspr0204_5
- Fink, A., Grabner, R. H., Gebauer, D., Reishofer, G., Koschutnig, K., and Ebner, F. (2010). Enhancing creativity by means of cognitive stimulation: evidence from an fMRI study. *Neuroimage* 52, 1687–1695. doi: 10.1016/j.neuroimage.2010.05.072
- Fink, A., Graif, B., and Neubauer, A. C. (2009). Brain correlates underlying creative thinking: EEG alpha activity in professional vs. novice dancers. *Neuroimage* 46, 854–862. doi: 10.1016/j.neuroimage.2009.02.036
- Fink, A., Koschutnig, K., Hutterer, L., Steiner, E., Benedek, M., Weber, B., et al. (2014). Gray matter density in relation to different facets of verbal creativity. *Brain Struct. Funct.* 219, 1263–1269. doi: 10.1007/s00429-013-0564-0
- Finke, R. A. (1996). Imagery, creativity, and emergent structure. *Conscious. Cogn.* 5, 381–393. doi: 10.1006/ccog.1996.0024
- Fox, M. D., Snyder, A. Z., Vincent, J. L., and Raichle, M. E. (2007). Intrinsic fluctuations within cortical systems account for intertrial variability in human behavior. *Neuron* 56, 171–184. doi: 10.1016/j.neuron.2007.08.023
- Gao, Z. K., Cai, Q., Yang, Y. X., Dong, N., and Zhang, S. S. (2017). Visibility graph from adaptive optimal kernel time-frequency representation for classification of epileptiform EEG. *Int. J. Neural Syst.* 27:1750005. doi: 10.1142/s0129065717500058
- Gino, F., and Ariely, D. (2012). The dark side of creativity: original thinkers can be more dishonest. *J. Pers. Soc. Psychol.* 102, 445–459. doi: 10.1037/a0026406
- Goldsmith, R. E., and Matherly, T. A. (1988). Creativity and self-esteem: a multiple operationalization validity study. *J. Psychol.* 122, 47–56. doi: 10.1080/00223980.1988.10542942
- Gough, H. G. (1976). Studying creativity by means of word association tests. *J. Appl. Psychol.* 61, 348–353. doi: 10.1037/0021-9010.61.3.348
- Gray-Little, B., and Carels, R. A. (1997). The effect of racial dissonance on academic self-esteem and achievement in elementary, junior high school students. *J. Res. Adoles.* 7, 109–131. doi: 10.1207/s15327795jra0702_1
- Greenberg, J., Solomon, S., and Pyszczynski, T. (1997). *Terror Management Theory Of Self-Esteem And Cultural Worldviews: Empirical Assessments and Conceptual Refinements*. In *Advances in Experimental Social Psychology*, Vol. 29. Cambridge, MA: Academic Press, 61–139.
- Hayes, J. (2018). *Process Models of Change*. The Theory and Practice of Change Management.
- Hwang, W. Y., Chen, N., Dung, J., and Yang, Y. (2007). Multiple representation skills and creativity effects on mathematical problem solving using a multimedia whiteboard system. *Educ. Technol. Soc.* 10:191.
- Jaquish, G. A., and Ripple, R. E. (1981). Cognitive creative abilities and self-esteem across the adult life-span. *Hum. Dev.* 24, 110–119. doi: 10.1159/000272654
- Kasof, J. (1997). Creativity and breadth of attention. *Creat. Res. J.* 10, 303–315. doi: 10.1207/s15326934crj1004_2
- Katayama, M., Kawaguchi, T., Berger, M. S., and Pieper, R. O. (2007). Dna damaging agent-induced autophagy produces a cytoprotective adenosine triphosphate surge in malignant glioma cells. *Cell Death. Differ.* 14, 548–558. doi: 10.1038/sj.cdd.4402030
- Kaufman, J. C., and Sternberg, R. J. (eds) (2010). *The Cambridge Handbook of Creativity*. Cambridge: Cambridge University Press.
- Khanna, A., Pascual-Leone, A., and Farzan, F. (2014). Reliability of resting-state microstate features in electroencephalography. *PLoS One* 9:e114163. doi: 10.1371/journal.pone.0114163
- Khanna, A., Pascual-Leone, A., Michel, C. M., and Farzan, F. (2015). microstates in resting-state eeg: current status and future directions. *Neurosci. Biobehav. Rev.* 49, 105–113. doi: 10.1016/j.neubiorev.2014.12.010
- Kikuchi, M., Koenig, T., Munesue, T., Hanaoka, A., and Strik, W. D. (2011). EEG microstate analysis in drug-naïve patients with panic disorder. *PLoS One* 6:e22912. doi: 10.1371/journal.pone.0022912
- Koenig, T., Lehmann, D., Merlo, M. C. G., Kochi, K., Hell, D., and Koukkou, M. (1999). A deviant eeg brain microstate in acute, neuroleptic-naïve schizophrenics at rest. *Eur. Arch. Psychiatry Clin. Neurosci.* 249, 205–211. doi: 10.1007/s004060050088
- Leary, M. R., Tambor, E. S., Terdal, S. K., and Downs, D. L. (1995). Self-esteem as an interpersonal monitor: the sociometer hypothesis. *J. Pers. Soc. Psychol.* 68:518. doi: 10.1037/0022-3514.68.3.518
- Lehmann, D., Faber, P. L., Galderisi, S., Herrmann, W. M., Kinoshita, T., Koukkou, M., et al. (2005). EEG microstate duration and syntax in acute, medication-naïve, first-episode schizophrenia: a multi-center study. *Neuroimaging* 138, 141–156. doi: 10.1016/j.pscychresns.2004.05.007
- Lehmann, D., Ozaki, H., and Pal, I. (1987). Eeg alpha map series: brain microstates by space-oriented adaptive segmentation. *Electroencephalogr. Clin. Neurophysiol.* 67:288.
- Lehmann, D., Pascual-Marqui, R. D., Strik, W. K., and Koenig, T. (2010). Core networks for visual-concrete and abstract thought content: a brain electric microstate analysis. *Neuroimage* 49, 1073–1079. doi: 10.1016/j.neuroimage.2009.07.054
- Lehmann, D., Strik, W. K., Henggeler, B., König, T., and Koukkou, M. (1998). Brain electric microstates and momentary conscious mind states as building blocks of spontaneous thinking: I. Visual imagery and abstract thoughts. *Int. J. Psychophysiol.* 29, 1–11. doi: 10.1016/s0167-8760(97)00098-6
- Li, M., Zhang, P., Li, S., and Wenxing, L. (2015). Learning through empowering leadership: can we achieve cross-level creativity simultaneously? *Leadersh. Org. Dev. J.* 2015, 16833–16833. doi: 10.5465/ambpp.2015.16833abstract
- Liu, M.-J., Shih, W.-L., and Ma, L.-Y. (2011). Are children with Asperger syndrome creative in divergent thinking and feeling? A brief report. *Res. Autism Spectr. Disord.* 5, 294–298. doi: 10.1016/j.rasd.2010.04.011
- Masten, C. L., Eisenberger, N. I., Borofsky, L. A., McNealy, K., Pfeifer, J. H., and Dapretto, M. (2011). Subgenual anterior cingulate responses to peer rejection: a marker of adolescents' risk for depression. *Dev. Psychopathol.* 23, 283–292. doi: 10.1017/s0954579410000799
- Memmert, D., and Roth, K. (2007). The effects of non-specific and specific concepts on tactical creativity in team ball sports. *J. Sports Sci.* 25, 1423–1432. doi: 10.1080/02640410601129755
- Mendelsohn, G., and Griswold, B. (1964). Differential use of incidental stimuli in problem solving as a function of creativity. *J. Abnorm. Soc. Psychol.* 68, 431–436. doi: 10.1037/h0040166
- Mendelsohn, G. A. (1976). Associative and attentional processes in creative performance. *J. Pers.* 44, 341–369. doi: 10.1111/j.1467-6494.1976.tb00127.x

- Mendelsohn, G. A., and Griswold, B. B. (1966). Assessed creative potential, vocabulary level, and sex as predictors of the use of incidental cues in verbal problem solving. *J. Pers. Soc. Psychol.* 4:423. doi: 10.1037/h0023783
- Mendelsohn, G. A., and Lindholm, E. P. (1972). Individual differences and the role of attention in the use of cues in verbal problem solving. *J. Personal.* 40, 226–241.
- Milz, P., Faber, P. L., Lehmann, D., König, T., Kochi, K., and Pascualmarqui, R. D. (2016). The functional significance of EEG microstates—Associations with modalities of thinking. *Neuroimage* 125, 643–656. doi: 10.1016/j.neuroimage.2015.08.023
- Musso, F., Brinkmeyer, J., Mobascher, A., Warbrick, T., and Winterer, G. (2010). Spontaneous brain activity and eeg microstates. *a novel eeg/fmri analysis approach to explore resting-state networks*. *Neuroimage* 52, 1149–1161. doi: 10.1016/j.neuroimage.2010.01.093
- Nijstad, B. A., De Dreu, C. K. W., Rietzschel, E. F., and Baas, M. (2010). The dual pathway to creativity model: creative ideation as a function of flexibility and persistence. *Eur. Rev. Soc. Psychol.* 21, 34–77. doi: 10.1080/10463281003765323
- Oldham, G. R., and Cummings, A. (1996). Employee creativity: personal and contextual factors at work. *Acad. Manag. J.* 39, 607–634. doi: 10.5465/256657
- Oppenheimer, D. M. (2008). The secret life of fluency. *Trends Cogn. Sci.* 12, 237–241. doi: 10.1016/j.tics.2008.02.014
- Pascual-Marqui, R. D., Michel, C. M., and Lehmann, D. (1995). Segmentation of brain electrical activity into microstates: model estimation and validation. *IEEE Trans. Biomed. Eng.* 42, 658–665. doi: 10.1109/10.391164
- Piffer, D. (2012). Can creativity be measured? An attempt to clarify the notion of creativity and general directions for future research. *Think. Skills Creat.* 7, 258–264. doi: 10.1016/j.tsc.2012.04.009
- Pyszczynski, T. (2004). What are we so afraid of?: a terror management theory perspective on the politics of fear. *Soc. Res.* 71, 827–848.
- Raichle, (2010). Two views of brain function. *Trends Cogn. Sci.* 14, 180–190. doi: 10.1016/j.tics.2010.01.008
- Raichle, M. E., and Mintun, M. A. (2006). Brain work and brain imaging. *Annu. Rev. Neurosci.* 29, 449–476. doi: 10.1146/annurev.neuro.29.051605.112819
- Rhodes, M. (1961). An analysis of creativity. *Phi Delta Kappan* 42, 305–310.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: University Press.
- Roskes, M., De Dreu, C. K., and Nijstad, B. A. (2012). Necessity is the mother of invention: avoidance motivation stimulates creativity through cognitive effort. *J. Pers. Soc. Psychol.* 103:242. doi: 10.1037/a0028442
- Rotge, J.-Y., Lemogne, C., Hinfrey, S., Huguet, P., and Fossati, P. (2014). A meta-analysis of the anterior cingulate contribution to social pain. *Soc. Cogn. Affect. Neurosci.* 10, 19–27.
- Runco, M. A. (2007). *Creativity: Theories and Themes: Research, Development, and Practice*. Amsterdam: Elsevier Academic Press.
- Sadaghiani, S., and Kleinschmidt, A. (2016). Brain networks and α -oscillations: structural and functional foundations of cognitive control. *Trends Cogn. Sci.* 20, 805–817. doi: 10.1016/j.tics.2016.09.004
- Sadaghiani, S., Scheeringa, R., Lehongre, K., Morillon, B., Giraud, A.-L., D'Esposito, M., et al. (2012). Alpha-band phase synchrony is related to activity in the fronto-parietal adaptive control network. *J. Neurosci.* 32, 14305–14310. doi: 10.1523/jneurosci.1358-12.2012
- Sadaghiani, S., Scheeringa, R., Lehongre, K., Morillon, B., Giraud, A.-L., and Kleinschmidt, A. (2010). Intrinsic connectivity networks, α oscillations, and tonic alertness: a simultaneous electroencephalography/functional magnetic resonance imaging study. *J. Neurosci.* 30, 10243–10250. doi: 10.1523/jneurosci.1004-10.2010
- Santarnecchi, E., Khanna, A. R., Musaeus, C. S., Csy, B., and Authors, H. S. T. (2017). Eeg microstate correlates of fluid intelligence and response to cognitive training. *Brain Topogr.* 30, 502–520. doi: 10.1007/s10548-017-0565-z
- Satzinger, J. W., Garfield, M. J., and Nagasundaram, M. (1999). The creative process: the effects of group memory on individual idea generation. *J. Manag. Inf. Syst.* 15, 143–160. doi: 10.1080/07421222.1999.11518225
- Schlegel, F., Lehmann, D., Faber, P. L., Milz, P., and Gianotti, L. R. (2012). EEG microstates during resting represent personality differences. *Brain Topogr.* 25, 20–26. doi: 10.1007/s10548-011-0189-7
- Schooler, J. W., Ohlsson, S., and Brooks, K. (1993). Thoughts beyond words: when language overshadows insight. *J. Exp. Psychol. Gen.* 122, 166–183. doi: 10.1037/0096-3445.122.2.166
- Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., et al. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *J. Neurosci.* 27, 2349–2356. doi: 10.1523/JNEUROSCI.5587-06.2007
- Seitzman, B. A., Abell, M., Bartley, S. C., Erickson, M. A., Bolbecker, A. R., Hetrick, W. P., et al. (2017). Cognitive manipulation of brain electric microstates. *Neuroimage* 146, 533–543. doi: 10.1016/j.neuroimage.2016.10.002
- Simonton, D. K. (1997). Creative productivity: a predictive and explanatory model of career trajectories and landmarks. *Psychol. Rev.* 104, 66–89. doi: 10.1037/0033-295x.104.1.66
- Simonton, K. D. (2011). Creativity and discovery as blind variation: campbell's (1960) bvsr model after the half-century mark. *Rev. Gen. Psychol.* 15, 158–174. doi: 10.1037/a0022912
- Slavich, G. M., and Epel, E. S. (2010). *The Stress and Adversity Inventory (STRAIN): An Automated System for Assessing Cumulative Stress Exposure*. Los Angeles, CA: University of California.
- Slotnick, S. D., Schwarzbach, J., and Yantis, S. (2003). Attentional inhibition of visual processing in human striate and extrastriate cortex. *Neuroimage* 19, 1602–1611. doi: 10.1016/s1053-8119(03)00187-3
- Solomon, S., Greenberg, J., and Pyszczynski, T. (1991). *A Terror Management Theory of Social Behavior: The Psychological Functions Of Self-Esteem And Cultural Worldviews*. In *Advances in Experimental Social Psychology*, Vol. 24. Cambridge, MA: Academic Press, 93–159.
- Sternberg, R. J. (ed.) (1999). *Handbook of Creativity*. Cambridge, MA: Cambridge University Press.
- Stevens, A., and Kircher, T. (1998). Cognitive decline unlike normal aging is associated with alterations of EEG temporo-spatial characteristics. *Eur. Arch. Psychiatry Clin. Neurosci.* 248, 259–266. doi: 10.1007/s004060050047
- Taylor, S. E., and Brown, J. D. (1988). Illusion and well-being: a social psychological perspective on mental health. *Psychol. Bull.* 103:193. doi: 10.1037/0033-2909.103.2.193
- Thatcher, S. M. B., and Brown, S. A. (2010). Individual creativity in teams: the importance of communication media mix. *Decis. Support Syst.* 49, 290–300. doi: 10.1016/j.dss.2010.03.004
- Tibshirani, R., and Walther, G. (2005). Cluster validation by prediction strength. *J. Comp. Graph. Statist.* 14, 511–528. doi: 10.1198/106186005x59243
- Van de Ville, D., Britz, J., and Michel, C. M. (2010). EEG microstate sequences in healthy humans at rest reveal scale-free dynamics. *Proc. Natl. Acad. Sci. U.S.A.* 107, 18179–18184. doi: 10.1073/pnas.1007841107
- Williams, B. R. (1994). Neurobiology, cellular and molecular biology, and psychosomatic medicine. *Psychos. Med.* 56, 308–315. doi: 10.1097/00006842-199407000-00006
- Williams, F. E. (1969). Models for encouraging creativity in the classroom by integrating cognitive-affective behaviors. *Educ. Technol.* 9, 7–13.
- Williams, F. E. (1980). *Creativity Assessment Packet (CAP): manual*. Buffalo, NY: DOK Publishers.
- Williams, F. E. (1993). *Creativity Assessment Packet Examiners Manual*. Austin, TX: PROED.
- Winkielman, P., Schwarz, N., Fazendeiro, T., and Reber, R. (2003). “The hedonic marking of processing fluency: implications for evaluative judgement,” in *The Psychology of Evaluation: Affective Processes in Cognition and Emotion*, eds J. Musch and K. C. Klauer (Mahwah, NJ: Erlbaum, Inc), 189–217.
- Wu, X., Yang, W., Tong, D., Sun, J., Chen, Q., Wei, D., et al. (2015). A meta-analysis of neuroimaging studies on divergent thinking using activation likelihood estimation. *Hum. Brain Mapp.* 36, 2703–2718. doi: 10.1002/hbm.22801
- Yau, C. (2011). An essential interrelationship: healthy self-esteem and productive creativity. *J. Creat. Behav.* 25, 154–161. doi: 10.1002/j.2162-6057.1991.tb01365.x
- Yuan, H., Zotev, V., Phillips, R., Drevets, W. C., and Bodurka, J. (2012). Spatiotemporal dynamics of the brain at rest — exploring eeg microstates as electrophysiological signatures of bold resting state networks. *Neuroimage* 60, 2062–2072. doi: 10.1016/j.neuroimage.2012.02.031
- Zappasodi, F., Perrucci, M. G., Saggino, A., Croce, P., Mercuri, P., Romanelli, R., et al. (2019). EEG microstates distinguish between cognitive components of fluid reasoning. *Neuroimage* 189, 560–573. doi: 10.1016/j.neuroimage.2019.01.067

Zeng, L., Salvendy, G., and Zhang, M. (2009). Factor structure of web site creativity. *Comp. Hum. Behav.* 25, 568–577. doi: 10.1016/j.chb.2008.12.023

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

Copyright © 2020 Wu, Guo, Wang, Zou, Guo, Lv and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Attentive Cursor Dataset

Luis A. Leiva^{1*} and Ioannis Arapakis²

¹ Department of Communications and Networking, Aalto University, Espoo, Finland, ² Telefónica Research, Madrid, Spain

Keywords: aimed movements, attention, demographics, web search, mouse cursor

1. INTRODUCTION

We introduce a large-scale dataset of mouse cursor movements that can be used to predict user attention, infer demographics information, and analyze fine-grained movements. Attention is a finite resource, so people spend their time on things they find valuable, especially when browsing online. Objective measurements of attentional processes are increasingly sought after by researchers, advertisers, and other key stakeholders from both academia and industry. With every click, digital footprints are created and logged, providing a detailed record of a person's online activity. However, click data provide an incomplete picture of user interaction, as they inform mainly about a users' end choice. A user click is often preceded by several valuable interactions, such as scrolling, hovers, aimed movements, etc. and thus having access to this kind of data can lead to an overall better understanding of the user's cognitive processes. For example, previous work has evidenced that when the mouse cursor is motionless, the user is processing information (Hauger et al., 2011; Huang et al., 2011; Diriye et al., 2012; Boi et al., 2016), i.e., essentially "users first focus and then execute actions" (Martín-Albo et al., 2016). We have collected mouse cursor tracking logs from near 3K subjects performing a transactional search task that together account for roughly 2 h worth of interaction data. Our dataset has associated attention labels and five demographics attributes that may help researchers to conduct several analysis, like the ones we discuss later in this section.

Research in mouse cursor tracking has a long track record. Chen et al. (2001) were among the first ones to note a relationship between gaze position and cursor position during web browsing. Mueller and Lockerd (2001) investigated the use of mouse tracking to create compelling visualizations and model the users' interests. It has been argued that mouse movements can reveal subtle patterns like reading (Hauger et al., 2011) or hesitation (Martín-Albo et al., 2016), and can help the user regain context after an interruption (Leiva, 2011a). Others have also noted the utility of mouse cursor analysis as a low-cost and scalable proxy of eye tracking (Huang et al., 2012; Navalpakkam et al., 2013). Several works have investigated closely the utility of mouse cursor data in web search (Arapakis et al., 2015; Lagun and Agichtein, 2015; Liu et al., 2015; Arapakis and Leiva, 2016; Chen et al., 2017) and web page usability evaluation (Arroyo et al., 2006; Atterer et al., 2006; Leiva, 2011b), two of the most prominent use cases of this technology. Mouse biometrics is another active research area that has shown promise in controlled settings (Lu et al., 2017; Krátky and Chudá, 2018). Researchers have started to analyze mouse movements on websites for the detection of neurodegenerative disorders (White et al., 2018; Gajos et al., 2020). In practice, commercial web search engines often use mouse cursor tracking to improve search results (Huang et al., 2011, 2012), optimize page design (Leiva, 2012; Diaz et al., 2013), and offer better recommendations to their users (Speicher et al., 2013). In what follows, we provide a brief survey of what others have accomplished by analyzing mouse cursor movements in web search tasks. These analyses highlight potential use cases of our dataset, thereby allowing researchers to investigate similar environments and behaviors.

OPEN ACCESS

Edited by:

Claudio De Stefano,
University of Cassino, Italy

Reviewed by:

Hugo Gamboa,
New University of Lisbon, Portugal
Nicole Cilia,
University of Cassino, Italy

*Correspondence:

Luis A. Leiva
firstname.lastname@aalto.fi

Specialty section:

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

Received: 25 May 2020

Accepted: 19 October 2020

Published: 16 November 2020

Citation:

Leiva LA and Arapakis I (2020) The
Attentive Cursor Dataset.
Front. Hum. Neurosci. 14:565664.
doi: 10.3389/fnhum.2020.565664

1.1. Inferring Interest

For a long time, commercial search engines have been interested in how users interact with Search Engine Result Pages (SERPs), to anticipate better placement and allocation of ads in sponsored search or to optimize the content layout. Early work considered simple, coarse-grained features derived from mouse cursor data to be surrogate measurements of user interest (Goecks and Shavlik, 2000; Claypool et al., 2001; Shapira et al., 2006). Follow-up research transitioned to more fine-grained mouse cursor features (Guo and Agichtein, 2008, 2010) that were shown to be more effective. These approaches have been directed at predicting open-ended tasks like search success (Guo et al., 2012) or search satisfaction (Liu et al., 2015). Mouse cursor position is mostly aligned to eye gaze, especially on SERPs (Guo and Agichtein, 2012; Lagun et al., 2014a), and that can be used as a good proxy for predicting good and bad abandonment (Diriye et al., 2012; Brückner et al., 2020).

1.2. Inferring Visual Attention

Mouse cursor tracking has been used to survey the visual focus of the user, thus revealing valuable information regarding the distribution of user attention over the various SERP components. Despite the technical challenges that may arise from this analysis, previous work has shown the utility of mouse movement patterns to measure within-content engagement (Arapakis et al., 2014a; Carlton et al., 2019) and predict reading experiences (Hauger et al., 2011; Arapakis et al., 2014b). Lagun et al. (2014a) introduced the concept of motifs, or frequent cursor subsequences, in the estimation of search result relevance. Similarly, Liu et al. (2015) applied the motifs concept to SERPs and predicted search result utility, searcher effort, and satisfaction at the search task level. Boi et al. (2016) proposed a method for predicting whether the user is actually looking at the content pointed by the cursor, exploiting the mouse cursor data and a segmentation of the web page contents. Lastly, Arapakis and Leiva (2016) investigated user engagement with direct displays on SERPs and provided further evidence that supports the utility of mouse cursor data for measuring user attention at a display-level granularity (Arapakis and Leiva, 2020; Arapakis et al., 2020).

1.3. Inferring Emotion

The connection between mouse cursor movements and the underlying psychological states has been a topic of research since the early 90s (Card et al., 1987; Accot and Zhai, 1997). Some studies have investigated the utility of mouse cursor data for predicting the user's emotional state. For example, Zimmermann et al. (2003) investigated the effect of induced affective states on the motor-behavior of online shoppers and found that the total duration of mouse cursor movements and the number of velocity changes were associated to the experienced arousal. Kaklauskas et al. (2009) created a system that extracts physiological and motor-control parameters from mouse cursor interactions and then triangulated those with psychological data taken from self-reports, to correlate the users' emotional state and productivity. In a similar line, Azcarraga and Suarez (2012) combined electroencephalography signals and mouse cursor interactions to predict self-reported emotions like frustration,

interest, confidence and excitement. Yamauchi (2013) studied the relationship between mouse cursor trajectories and generalized anxiety in human subjects. Lastly, Kapoor et al. (2007) predicted whether a user experiences frustration, using an array of affective-aware sensors.

1.4. Inferring Demographics

Prior work has linked age with motor control and pointing performance in tasks that involve the use of a computer mouse (Walker et al., 1997; Bohan and Chaparro, 1998; Hsu et al., 1999; Smith et al., 1999; Jastrzembski et al., 2003; Lindberg et al., 2006). Overall, aging is marked by a decline in motor control abilities, therefore it is expected to affect the users' pointing performance and, by extension, how they move the computer mouse. For example, Smith et al. (1999) observed that older people incurred in longer mouse movement times, more sub-movements, and more pointing errors than the young. These findings underline potential age effects on the way a mouse device is used in an online search task. Prior research has also noted sensory-motor differences due to gender (Landauer, 1981; Chen and Chen, 2008; Yamauchi et al., 2015), such as significant variation in the cursor movement distance, pointing time, and cursor patterns. The cause of these variations has been attributed to gender-based differences in how users move a mouse cursor or to different cognitive mechanisms (perceptual and spatial processes) involved in motor control.

Others have also examined the extent to which mouse cursor movements can help identify gender and age (Yamauchi and Bowman, 2014; Kratky and Chuda, 2016; Pentel, 2017), however the experimental settings have limited generalizability, either because the tasks are not well-connected to typical activities that users perform online, such as web search, because the data include multiple samples per participant, thereby increasing the risks of information leakage, or because researchers could not verify their ground-truth data. In our dataset, we limit the training samples to exactly one mouse cursor trajectory per participant, who are verified, high-quality crowdworkers.

2. METHOD

We ran an online crowdsourcing study that reproduced the conditions of a transactional search task. Participants were presented with a simulated information need that explained that they were interested in purchasing some product for them or a friend. Overall, the study consisted of three parts, to be described later: (1) pre-task guidelines, (2) the web search task, and (3) a post-task questionnaire.

2.1. Participants

We recruited participants from the FIGURE EIGHT crowdsourcing platform¹. They were of mixed nationality (e.g., American, Belgian, British, German) and had diverse educational backgrounds (see Table 1). All participants were proficient in English and were experienced (Level 3) contributors, i.e., they had a proven track record of successfully completed

¹<https://www.figure-eight.com>

TABLE 1 | Demographics information from our dataset.

Age	Count	Gender	Count	Nationality	Count	Education	Count	Income	Count
18–23	380	Male	1,605	USA	1,755	High school	593	<25K	881
24–29	716	Female	1,118	VEN	251	College	472	25–34K	446
30–35	590	NA	14	GBR	209	Bachelor's	704	35–49K	367
36–41	417			CAN	66	Graduate	499	50–74K	394
42–47	223			EGY	37	Master's	399	75–99K	249
48–53	174			UKR	31	Doctorate	30	100–149K	145
54–59	132			IND	29	NA	40	150–249K	42
60–65	63			SRB	27			>250K	23
+66	24			RUS	25			NA	190
NA	18			...					

tasks and of a different variety, thus being considered very reliable contributors.

2.2. Materials

Starting from Google Trends², we selected a subset of the Top Categories and Shopping Categories that were suitable representatives of transactional tasks. Then, we extracted the top search queries issued in the US during the last 12 months. Next, we narrowed down our search query collection to 150 representative popular queries. The final collection of transactional queries was repeated as many times needed to produce the desired number of search sessions for the final dataset.

Using this final selection of search queries, we produced the static version of the corresponding Google SERPs and injected custom JavaScript code that allowed us to capture all client-side user interactions. For this, we used EVTRACK³, an open source JavaScript event tracking library derived from the smt2e mouse tracking system (Leiva and Vivó, 2013). EVTRACK can capture browser events either via event listeners (the event is captured as soon as it is fired) or via event polling (the event is captured at fixed-time intervals). We captured `mousemove` events via event polling, every 150 ms to avoid unnecessary data overhead (Leiva and Huang, 2015), and all the other browser events (e.g., `load`, `click`, `scroll`) via event listeners. Whenever an event was recorded, we logged the following information: mouse cursor position (*x* and *y* coordinates), timestamp, event name, XPath of the DOM element that relates to the event, and the DOM element attributes (if any).

All queries triggered some form of advertisements on the SERPs, according to three different formats: “native” (organic ads) or “bundled” (direct display ads). All SERPs included one or more native ads together with one bundled ad. The native advertisements could appear either at the top or bottom position of the SERP, whereas the bundled ads could appear either at the top-left or top-right position. We ensured that only one ad was visible per condition and participant at a time. This was possible by instrumenting each downloaded SERP with custom JavaScript

code that removed all ads excepting one that would be selected for a given participant. In any case, native bottom-most ads were not shown to the participants.

2.3. Pre-task Guidelines

Participants were instructed to read carefully the terms and conditions of the study which, among other things, informed them that they should perform the task from a desktop or laptop computer using a computer mouse (and refrain from using a touchpad, tablet, or mobile device) and that their browsing activity would be logged. Moreover, participants consented to share their browsing data and their (anonymized) responses for later analysis.

Participants were asked to act naturally and choose anything that would best answer a given search query, since all “clickable” elements (e.g., result links, images, etc.) on the SERP were considered valid answers. The instructions were followed by a brief search task description using this template: “*You want to buy <noun> (for you or someone else as a gift) and you have submitted the search query <noun> to Google Search. Please browse the search results page and click on the element that you would normally select under this scenario.*” The template was populated with the corresponding <noun> entities, based on the assigned query.

Participants were allowed as much time as they needed to examine the SERP and proceed with the search task, which would conclude whenever they clicked on any SERP element. The payment for the participation was \$0.20. Participants could also opt out at any moment, in which case they were not compensated. Each participant could take the study only once.

2.4. Task Procedure

Each participant was presented with a search task description, then provided with a predefined search query (selected at random from our pool of queries) and the corresponding SERP, and they were asked to click on any element of the page that best solved the task. This way, we ensured that participants interacted with the same pool of web search queries and avoided any unaccounted systematic bias due to query quality variation. All possible combinations of query and ad style (i.e., format and position)

²<https://trends.google.com/trends/>

³<https://github.com/luileito/evtrack>

were pre-computed so that whenever a new user accessed the study, they were assigned one of these combinations at random.

Participants accessed the instrumented SERPs through a dedicated web server that did not alter the look and feel of the original SERPs. This allowed us to capture fine-grained user interactions while ensuring that the content of the SERPs remained consistent with the original version. Each participant was allowed to perform the search task only once to avoid introducing possible carry over effects and, thus, altering their browsing behavior in subsequent search tasks. In sum, each participant was exposed only to a single condition; i.e., a unique combination of query and ad style. Finally, at the end of the study participants had to copy a unique code and paste it on FIGURE EIGHT in order to have their job validated.

2.5. Post-task Questionnaire

Upon concluding the search task, participants were asked to answer a series of questions. The questions were forced-choice type and allowed multi-point response options.

The first question asked the degree to which the user noticed the advertisements shown on the SERP: *While performing the search task, to what extent did you pay attention to the advertisement?* We used a 5-point Likert-type scale to collect the labels: 1 (“Not at all”), 2 (“Not much”), 3 (“I can’t decide”), 4 (“Somewhat”), and 5 (“Very much”). In practice, these scores should be collapsed to binary labels (true/false), but we felt it was necessary to use a 5-point Likert-type scale for several reasons. First, using 2-point scales often results in highly skewed data (Johnson et al., 1982). Second, it is important to leave room for neutral responses, because some users may not want to say one way or another, otherwise this can produce response biases. But 3-point scales can lead more users to stay neutral, because the remaining options can be seen as “too extreme.” Therefore, we opted for a 5-point scale, which leaves more room for “soft responses” and in addition is easy to understand. With this scoring scheme, therefore, we are confident that eventual binary labels would actually reflect positive and negative user votes.

The questionnaire also comprised the following demographics-related questions:

1. *What is your gender?* [Male, Female, Prefer not to say]
2. *What is your age group?* [18–23, 24–29,..., 60–65, +66, Prefer not to say]
3. *What is your native language?* [Pull-down list, Prefer not to say]
4. *What is your education level?* [High school, College,..., Doctorate, Prefer not to say]
5. *What is your current income?* [25K, 35K,..., 250K, Prefer not to say]

3. VALIDATION AND FILTERING

Crowdsourcing studies offer several advantages over *in-situ* methods of experimentation (Mason and Suri, 2012), such as access at a larger and more diverse pool of participants with stable availability, collection of real usage data at a relatively large scale, and a low-cost alternative to the more expensive laboratory-based

experiments. On the downside, experimenters have to account for potential threats to ecological validity, distractions in the physical environment of the participant, and privacy issues, to name a few. Still, crowdsourcing allows for exploring a wider range of parameters in a more controlled manner as compared to in-the-wild large-scale studies.

We collected self-reported ground-truth labels in a similar vein to previous work (Feild et al., 2010; Lagun et al., 2014b; Liu et al., 2015; Arapakis and Leiva, 2016) which also administered post-task questionnaires. To mitigate and discount low-quality responses, several preventive measures were put into practice, such as introducing test (gold-standard) questions to our tasks, selecting experienced contributors with high accuracy rates, and monitoring their task completion time, thus ensuring the internal validity of our experiment.

Starting from a set of 3,223 participants who initially accessed the study, we filtered automatically those who did not finish it (138 cases) as well as participants who did not move their mouse at all (176 cases). We concluded to a dataset with 2,909 observations comprising at least one mouse movement, together with their associated browser’s and user’s metadata. See **Table 1** for a summary of the available demographics information.

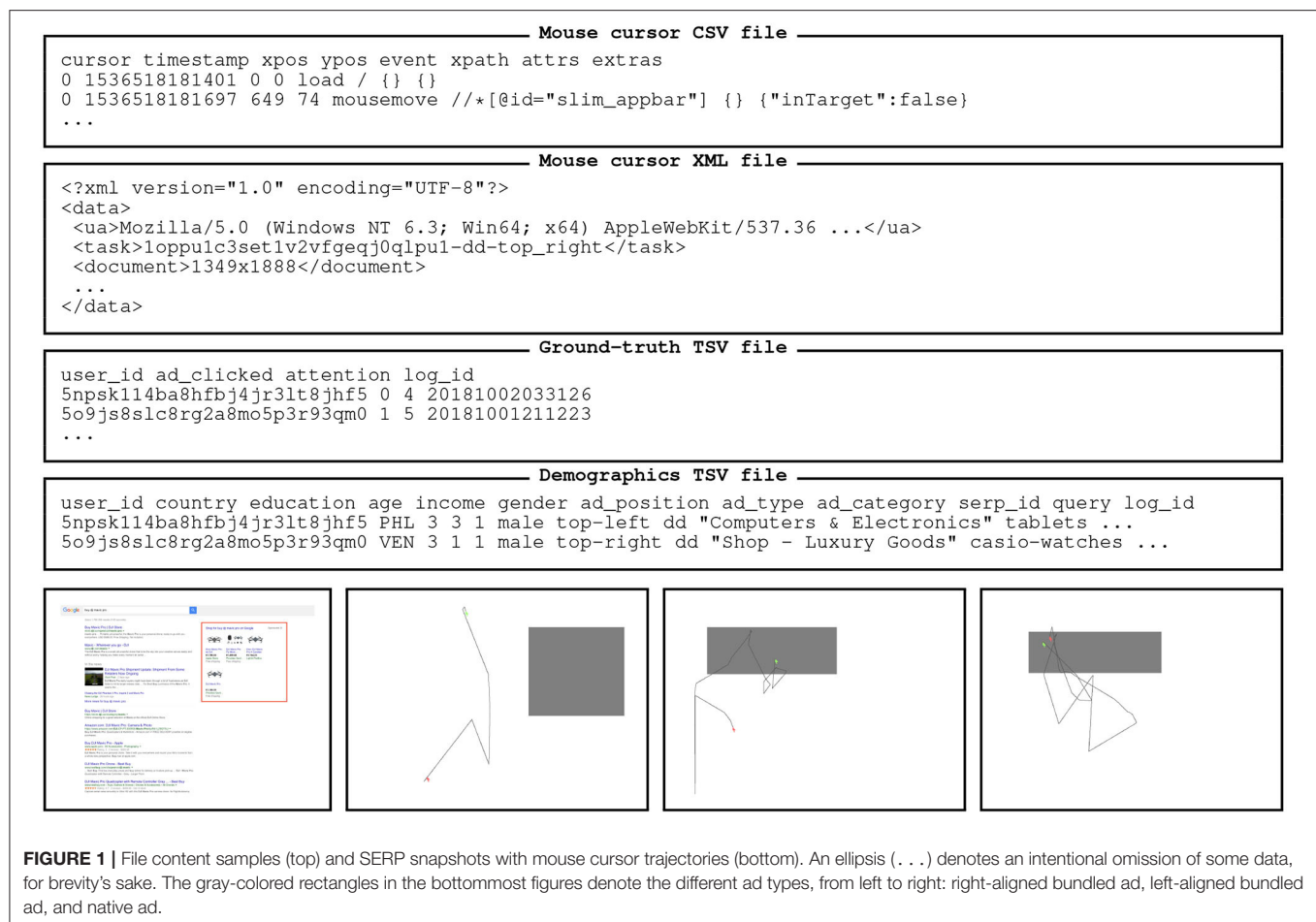
There are 92 unique combinations of query and ad style, each of which assessed by 32 users on average ($SD = 17$ users). There are 1,942 observations from the attended condition (self-reported Likert-type score ≥ 4), 776 observations from the non-attended condition (score ≤ 2), and 191 observations from the neutral condition (score of 3). The average mouse cursor trajectory has 15.78 coordinates ($SD = 16.5$, $min = 1$, $max = 222$), which is around the same order of magnitude as reported in similar studies (Huang et al., 2011; Leiva and Huang, 2015; Arapakis and Leiva, 2016).

Excepting the automatic filtering procedure explained above, our data is in raw form and therefore some columns require further processing. For example, most columns pertaining demographics information are stored as integers, therefore researchers should consult **Table 1** to retrieve the corresponding categorical labels. We also recommend researchers to apply other filtering methods, depending on the nature of their experiments, such as collapsing the ground-truth attention labels from the original 1–5 scale to a binary scale (Arapakis and Leiva, 2020; Arapakis et al., 2020) or ignoring cursor trajectories having <5 coordinates, which in most cases would correspond to 1 s of interaction data.

3.1. Data Format

The Attentive Cursor dataset includes the following resources:

1. A folder with mouse tracking log files, as recorded by the EVTRACK software:
 - a. Browser events: space-delimited files (CSV) with information about each event type (8 columns).
 - b. Browser metadata: XML files with information about the user’s browser (e.g., viewport size).
2. A TSV file with ground-truth labels (4 columns).



3. A tab-delimited file (TSV) with user's demographics and stimulus condition (12 columns).
4. A folder with all SERPs in HTML format.
5. A README file with a detailed explanation of each resource.

Figure 1 provides some examples of the kind of data that researchers can find in our dataset. We provide the URL to the repository in the "Data Availability Statement" section below.

4. CONCLUSION

We have presented a large-scale, in-the-wild dataset of mouse cursor movements in web search, with associated ground-truth labels about user's attention and demographics attributes. The dataset represents real-world behavior of individuals completing a transactional web search task. What makes this dataset both unique and challenging is the fact that there is only one observation per user. It is not possible to leak information from any data splits; e.g., training, validation, and testing splits typically used in machine learning studies. It is our hope that the dataset will foster research in several scientific domains, including, e.g., information retrieval, movement science, and psychology.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

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

ACKNOWLEDGMENTS

Two manuscripts using a post-processed version of this dataset have been recently published by the authors (Arapakis and Leiva, 2020; Arapakis et al., 2020). LAL acknowledged the support from the Finnish Center for Artificial Intelligence (FCAI).

REFERENCES

- Accot, J., and Zhai, S. (1997). "Beyond fitts' law: Models for trajectory-based HCI tasks," in *Proceedings of CHI* (Atlanta, GA), 295–302. doi: 10.1145/258549.258760
- Arapakis, I., Lalmas, M., Cambazoglu, B. B., Marcos, M. C., and Jose, J. M. (2014a). User engagement in online news: Under the scope of sentiment, interest, affect, and gaze. *J. Assoc. Inf. Sci. Technol.* 65, 1988–2005. doi: 10.1002/asi.23096
- Arapakis, I., Lalmas, M., and Valkanas, G. (2014b). "Understanding within-content engagement through pattern analysis of mouse gestures," in *Proceedings of CIKM*, 1439–1448. doi: 10.1145/2661829.2661909
- Arapakis, I., and Leiva, L. A. (2016). "Predicting user engagement with direct displays using mouse cursor information," in *Proceedings of SIGIR* (Pisa), 599–608. doi: 10.1145/2911451.2911505
- Arapakis, I., and Leiva, L. A. (2020). "Learning efficient representations of mouse movements to predict user attention," in *Proceedings of SIGIR*. doi: 10.1145/3397271.3401031
- Arapakis, I., Leiva, L. A., and Cambazoglu, B. B. (2015). "Know your onions: understanding the user experience with the knowledge module in web search," in *Proceedings of CIKM* (Melbourne, VIC), 1695–1698. doi: 10.1145/2806416.2806591
- Arapakis, I., Penta, A., Joho, H., and Leiva, L. A. (2020). A price-per-attention auction scheme using mouse cursor information. *ACM Trans. Inf. Syst.* 38:13. doi: 10.1145/3374210
- Arroyo, E., Selker, T., and Wei, W. (2006). "Usability tool for analysis of web designs using mouse tracks," in *Proceedings of CHIEA* (Montréal, QC), 357–362. doi: 10.1145/1125451.1125557
- Atterer, R., Wnuk, M., and Schmidt, A. (2006). "Knowing the user's every move: user activity tracking for website usability evaluation and implicit interaction," in *Proceedings of WWW* (Edinburgh), 203–212. doi: 10.1145/1135777.1135811
- Azcarraga, J., and Suarez, M. T. (2012). "Predicting academic emotions based on brainwaves, mouse behaviour and personality profile," in *Proceedings of PRICAI* (Berlin; Heidelberg: Springer), 728–733. doi: 10.1007/978-3-642-32695-0_64
- Bohan, M., and Chaparro, A. (1998). Age-related differences in performance using a mouse and trackball. *Hum. Factors* 42, 152–155. doi: 10.1177/154193129804200202
- Boi, P., Fenu, G., Spano, L. D., and Vargiu, V. (2016). Reconstructing user's attention on the web through mouse movements and perception-based content identification. *ACM Trans. Appl. Percept.* 13, 15:1–15:21. doi: 10.1145/2912124
- Brückner, L., Arapakis, I., and Leiva, L. A. (2020). "Query abandonment prediction with deep learning models of mouse cursor movements," in *Proceedings of CIKM*. doi: 10.1145/3340531.3412126
- Card, S. K., English, W. K., and Burr, B. J. (1987). "Evaluation of mouse, rate controlled isometric joystick, step keys, and text keys, for text selection on a CRT," in *Human-Computer Interaction*, eds R. M. Baecker and W. A. S. Buxton (Taylor & Francis), 386–392.
- Carlton, J., Brown, A., Jay, C., and Keane, J. (2019). "Inferring user engagement from interaction data," in *Proceedings of CHI EA*, 1212, 1–6. doi: 10.1145/3290607.3313009
- Chen, M. C., Anderson, J. R., and Sohn, M. H. (2001). "What can a mouse cursor tell us more? Correlation of eye/mouse movements on web browsing," in *Proceedings of CHI EA*, 281–282. doi: 10.1145/634067.634234
- Chen, R. C. C., and Chen, T.-K. (2008). The effect of gender-related difference on human-centred performance using a mass assessment method. *IJCAT* 32, 322–333. doi: 10.1504/IJCAT.2008.021387
- Chen, Y., Liu, Y., Zhang, M., and Ma, S. (2017). User satisfaction prediction with mouse movement information in heterogeneous search environment. *IEEE Trans. Knowl. Data. Eng.* 29, 2470–2483. doi: 10.1109/TKDE.2017.2739151
- Claypool, M., Le, P., Wased, M., and Brown, D. (2001). "Implicit interest indicators," in *Proceedings of IUI*, 33–40. doi: 10.1145/359784.359836
- Diaz, F., White, R., Buscher, G., and Liebling, D. (2013). "Robust models of mouse movement on dynamic web search results pages," in *Proceedings of CIKM* (San Francisco, CA), 1451–1460. doi: 10.1145/2505515.2505717
- Diriye, A., White, R., Buscher, G., and Dumais, S. (2012). "Leaving so soon? Understanding and predicting web search abandonment rationales," in *Proceedings of CIKM*, 1025–1034. doi: 10.1145/2396761.2398399
- Feild, H. A., Allan, J., and Jones, R. (2010). "Predicting searcher frustration," in *Proceedings of SIGIR*, 34–41. doi: 10.1145/1835449.1835458
- Gajos, K. Z., Reinecke, K., Donovan, M., Stephen, C. D., Hung, A. Y., Schmahmann, J. D., et al. (2020). Computer mouse use captures ataxia and parkinsonism, enabling accurate measurement and detection. *Mov. Disord.* 35, 354–358. doi: 10.1002/mds.27915
- Goecks, J., and Shavlik, J. (2000). "Learning users' interests by unobtrusively observing their normal behavior," in *Proceedings of IUI*, 129–132. doi: 10.1145/325737.325806
- Guo, Q., and Agichtein, E. (2008). "Exploring mouse movements for inferring query intent," in *Proceedings of SIGIR* (Singapore), 707–708. doi: 10.1145/1390334.1390462
- Guo, Q., and Agichtein, E. (2010). "Ready to buy or just browsing? Detecting web searcher goals from interaction data," in *Proceedings of SIGIR*, 130–137. doi: 10.1145/1835449.1835473
- Guo, Q., and Agichtein, E. (2012). "Beyond dwell time: estimating document relevance from cursor movements and other post-click searcher behavior," in *Proceedings of WWW*, 569–578. doi: 10.1145/2187836.2187914
- Guo, Q., Lagun, D., and Agichtein, E. (2012). "Predicting web search success with fine-grained interaction data," in *Proceedings of CIKM*, 2050–2054. doi: 10.1145/2396761.2398570
- Hauger, D., Paramythis, A., and Weibelzahl, S. (2011). "Using browser interaction data to determine page reading behavior," in *Proceedings of UMAP* (Berlin; Heidelberg: Springer), 147–158. doi: 10.1007/978-3-642-22362-4_13
- Hsu, S. H., Huang, C. C., Tsuang, Y. H., and Sun, J. S. (1999). Effects of age and gender on remote pointing performance and their design implications. *Int. J. Ind. Ergon.* 23, 461–471. doi: 10.1016/S0169-8141(98)00013-4
- Huang, J., White, R., and Buscher, G. (2012). "User see, user point: gaze and cursor alignment in web search," in *Proceedings of CHI*, 1341–1350. doi: 10.1145/2207676.2208591
- Huang, J., White, R. W., and Dumais, S. (2011). "No clicks, no problem: using cursor movements to understand and improve search," in *Proceedings of CHI*, 1225–1234. doi: 10.1145/1978942.1979125
- Jastrzemski, T., Charness, N., Holley, P., and Feddon, J. (2003). Input devices for web browsing: age and hand effects. *Universal Access Inf.* 4, 39–45. doi: 10.1007/s10209-003-0083-5
- Johnson, S., Smith, P., and Tucker, S. (1982). Response format of the job descriptive index: assessment of reliability and validity by the multitrait-multimethod matrix. *J. Appl. Psychol.* 67, 500–505. doi: 10.1037/0021-9010.67.4.500
- Kaklauskas, A., Krutinis, M., and Seniut, M. (2009). "Biometric mouse intelligent system for student's emotional and examination process analysis," in *Proceedings of ICALT* (Riga), 189–193. doi: 10.1109/ICALT.2009.130
- Kapoor, A., Burleson, W., and Picard, R. W. (2007). Automatic prediction of frustration. *Int. J. Hum. Comput. Stud.* 65, 724–736. doi: 10.1016/j.ijhcs.2007.02.003
- Kratky, P., and Chuda, D. (2016). "Estimating gender and age of web page visitors from the way they use their mouse," in *Proceedings of WWW Companion*, 61–62. doi: 10.1145/2872518.2889384
- Krátký, P., and Chudá, D. (2018). Recognition of web users with the aid of biometric user model. *J. Intell. Inf. Syst.* 51, 621–646. doi: 10.1007/s10844-018-0500-0
- Lagun, D., Ageev, M., Guo, Q., and Agichtein, E. (2014a). "Discovering common motifs in cursor movement data for improving web search," in *Proceedings of WSDM*, 183–192. doi: 10.1145/2556195.2556265
- Lagun, D., and Agichtein, E. (2015). "Inferring searcher attention by jointly modeling user interactions and content salience," in *Proceedings of SIGIR*, 483–492. doi: 10.1145/2766462.2767745
- Lagun, D., Hsieh, C.-H., Webster, D., and Navalpakkam, V. (2014b). "Towards better measurement of attention and satisfaction in mobile search," in *Proceedings of SIGIR*, 113–122. doi: 10.1145/2600428.2609631
- Landauer, A. A. (1981). Sex differences in decision and movement time. *Percept. Mot. Skills* 52, 90–90. doi: 10.2466/pms.1981.52.1.90
- Leiva, L. A. (2011a). "Mousehints: easing task switching in parallel browsing," in *Proceedings of CHI EA*, 1957–1962. doi: 10.1145/1979742.1979861
- Leiva, L. A. (2011b). "Restyling website design via touch-based interactions," in *Proceedings of Mobile HCI (Stockholm)*, 599–604. doi: 10.1145/2037373.2037467
- Leiva, L. A. (2012). "Automatic web design refinements based on collective user behavior," in *Proceedings of CHI EA*, 1607–1612. doi: 10.1145/2212776.2223680

- Leiva, L. A., and Huang, J. (2015). Building a better mousetrap: compressing mouse cursor activity for web analytics. *Inf. Process. Manag.* 51, 114–129. doi: 10.1016/j.ipm.2014.10.005
- Leiva, L. A., and Vivó, R. (2013). Web browsing behavior analysis and interactive hypervideo. *ACM Trans. Web* 7, 20:1–20:28. doi: 10.1145/2529995.2529996
- Lindberg, T., Näsänen, R., and Müller, K. (2006). How age affects the speed of perception of computer icons. *Displays* 27, 170–177. doi: 10.1016/j.displa.2006.06.002
- Liu, Y., Chen, Y., Tang, J., Sun, J., Zhang, M., Ma, S., et al. (2015). “Different users, different opinions: predicting search satisfaction with mouse movement information,” in *Proceedings of SIGIR* (Santiago), 493–502. doi: 10.1145/2766462.2767721
- Lu, H., Rose, J., Liu, Y., Awad, A., and Hou, L. (2017). “Combining mouse and eye movement biometrics for user authentication,” in *Information Security Practices*, eds I. Traoré, A. Awad, and I. Woungang (Cham: Springer), 55–71. doi: 10.1007/978-3-319-48947-6_5
- Martin-Albo, D., Leiva, L. A., Huang, J., and Plamondon, R. (2016). Strokes of insight: user intent detection and kinematic compression of mouse cursor trails. *Inf. Process. Manag.* 52, 989–1003. doi: 10.1016/j.ipm.2016.04.005
- Mason, W., and Suri, S. (2012). Conducting behavioral research on Amazon’s Mechanical Turk. *Behav. Res. Methods* 44, 1–23. doi: 10.3758/s13428-011-0124-6
- Mueller, F., and Lockerd, A. (2001). “Cheese: tracking mouse movement activity on websites, a tool for user modeling,” in *Proceedings of CHI EA*, 279–280. doi: 10.1145/634067.634233
- Navalpakkam, V., Jentzsch, L., Sayres, R., Ravi, S., Ahmed, A., and Smola, A. (2013). “Measurement and modeling of eye-mouse behavior in the presence of nonlinear page layouts,” in *Proceedings of WWW*, 953–964. doi: 10.1145/2488388.2488471
- Pentel, A. (2017). “Predicting age and gender by keystroke dynamics and mouse patterns,” in *Adjunct Proceedings of UPMAP* (Bratislava), 381–385. doi: 10.1145/3099023.3099105
- Shapira, B., Taieb-Maimon, M., and Moskowitz, A. (2006). “Study of the usefulness of known and new implicit indicators and their optimal combination for accurate inference of users interests,” in *Proceedings of SAC* (Dijon), 1118–1119. doi: 10.1145/1141277.1141542
- Smith, M. W., Sharit, J., and Czaja, S. J. (1999). Aging, motor control, and the performance of computer mouse tasks. *Hum. Factors* 41, 389–396. doi: 10.1518/001872099779611102
- Speicher, M., Both, A., and Gaedke, M. (2013). “TellMyRelevance! Predicting the relevance of web search results from cursor interactions,” in *Proceedings of CIKM* (Burlingame, CA), 1281–1290. doi: 10.1145/2505515.2505703
- Walker, N., Philbin, D. A., and Fisk, A. D. (1997). Age-related differences in movement control: adjusting submovement structure to optimize performance. *J. Gerontol. A Biol. Sci. Med. Sci.* 52, 389–396. doi: 10.1093/geronb/52B.1.P40
- White, R., Doraiswamy, P., and Horvitz, E. (2018). Detecting neurodegenerative disorders from web search signals. *NPJ Digit. Med.* 1:8. doi: 10.1038/s41746-018-0016-6
- Yamauchi, T. (2013). “Mouse trajectories and state anxiety: feature selection with random forest,” in *Proceedings of ACII* (Geneva), 399–404. doi: 10.1109/ACII.2013.72
- Yamauchi, T., and Bowman, C. (2014). “Mining cursor motions to find the gender, experience, and feelings of computer users,” in *Proceedings of ICDMW* (Shenzhen), 221–230. doi: 10.1109/ICDMW.2014.131
- Yamauchi, T., Seo, J. H., Jett, N., Parks, G., and Bowman, C. (2015). Gender differences in mouse and cursor movements. *Int. J. Hum. Comput. Interact.* 31, 911–921. doi: 10.1080/10447318.2015.1072787
- Zimmermann, P., Guttormsen, S., Danuser, B., and Gomez, P. (2003). Affective computing—a rationale for measuring mood with mouse and keyboard. *Int. J. Occup. Saf. Ergon.* 9, 539–551. doi: 10.1080/10803548.2003.11076589

Conflict of Interest: IA was employed by the company Telefonica Research, though no payment or services from the institution has been received or requested for any aspect of the submitted work.

The remaining 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.

Copyright © 2020 Leiva and Arapakis. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Think Hard or Think Smart: Network Reconfigurations After Divergent Thinking Associate With Creativity Performance

Hong-Yi Wu¹, Bo-Cheng Kuo², Chih-Mao Huang³, Pei-Jung Tsai⁴, Ai-Ling Hsu^{5,6}, Li-Ming Hsu⁷, Chi-Yun Liu⁸, Jyh-Horng Chen¹ and Changwei W. Wu^{8,9*}

¹ Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei, Taiwan, ² Department of Psychology, National Taiwan University, Taipei, Taiwan, ³ Department of Biological Science and Technology, National Chiao Tung University, Hsinchu, Taiwan, ⁴ Intramural Research Program, Neuroimaging Research Branch, National Institute on Drug Abuse, National Institutes of Health, Baltimore, MLD, United States, ⁵ Department of Radiology, Wan Fang Hospital, Taipei Medical University, Taipei, Taiwan, ⁶ Department of Radiology, School of Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan, ⁷ Department of Radiology and Brain Research Imaging Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States, ⁸ Graduate Institute of Mind, Brain, and Consciousness, Taipei Medical University, Taipei, Taiwan, ⁹ Brain and Consciousness Research Center, Taipei Medical University-Shuang-Ho Hospital, New Taipei, Taiwan

OPEN ACCESS

Edited by:

Jose Luis Contreras-Vidal,
University of Houston, United States

Reviewed by:

Hsiang-Yuan Lin,
University of Toronto, Canada

Der-Yow Chen,
National Cheng Kung
University, Taiwan

*Correspondence:

Changwei W. Wu
sleepbrain@tmu.edu.tw

Specialty section:

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

Received: 09 June 2020

Accepted: 26 October 2020

Published: 20 November 2020

Citation:

Wu H-Y, Kuo B-C, Huang C-M,
Tsai P-J, Hsu A-L, Hsu L-M, Liu C-Y,
Chen J-H and Wu CW (2020) Think
Hard or Think Smart: Network
Reconfigurations After Divergent
Thinking Associate With Creativity
Performance.
Front. Hum. Neurosci. 14:571118.
doi: 10.3389/fnhum.2020.571118

Evidence suggests divergent thinking is the cognitive basis of creative thoughts. Neuroimaging literature using resting-state functional connectivity (RSFC) has revealed network reorganizations during divergent thinking. Recent studies have revealed the changes of network organizations when performing creativity tasks, but such brain reconfigurations may be prolonged after task and be modulated by the trait of creativity. To investigate the dynamic reconfiguration, 40 young participants were recruited to perform consecutive Alternative Uses Tasks (AUTs) for divergent thinking and two resting-state scans (before and after AUT) were used for mapping the brain reorganizations after AUT. We split participants into high- and low-creative groups based on creative achievement questionnaire (CAQ) and targeted on reconfigurations of the two brain networks: (1) default-mode network (DMN) and (2) the network seeded at the left inferior frontal gyrus (IFG) because the between-group difference of AUT-induced brain activation located at the left IFG. The changes of post-AUT RSFCs (DMN and IFGN) indicated the prolonged effect of divergent thinking. More specifically, the alterations of RSFC_{IFG-AG} and RSFC_{IFG-IPL} (AG: angular gyrus, IPL: inferior parietal lobule) in the high-creative group had positive relationship with their AUT performances (originality and fluency), but not found in the low-creative group. Furthermore, the RSFC changes of DMN did not present significant relationships with AUT performances. The findings not only confirmed the possibility of brain dynamic reconfiguration following divergent thinking, but also suggested the distinct IFGN reconfiguration between individuals with different creativity levels.

Keywords: creativity, divergent thinking, functional connectivity, resting-state fMRI, alternative usage task (AUT), creative achievement questionnaire (CAQ)

INTRODUCTION

Creativity is the foundation of originality, the generation of novel ideas when facing a specific problem (Sternberg, 1999; Runco and Jaeger, 2012), and the cornerstone of productivity in human civilization and modern society. However, despite the current surging importance of creativity, its manifestations in the brain involves a complex architecture, the underlying mechanisms of which require an extensive investigation to disentangle (Sternberg and Lubart, 1996). The current psychometric creativity measures are largely based on Guilford's theory, according to which creative people have high ideational fluency and high degrees of novelty (Guilford, 1967). To date, the understanding of the internal process of creativity tended to alternate between the generation of novel ideas (i.e., divergent thinking) and the evaluation of generated ideas (i.e., convergent thinking). In the dual-process conception of creativity (Abraham, 2013; Sowden et al., 2015), divergent thinking represents the acquisition of a certain task and diverts attention away from the task itself, in a highly spontaneous manner to generate ideas, whereas the convergent thinking is associated with deliberate constraints and the verification of illuminated ideas (Christoff et al., 2016). Therefore, the divergent thinking can be regarded as an imperative incubation step before the "eureka!" moment reaches the mind. More importantly, the divergent thinking ability has moderate potential to predict creative achievements in the real world (Plucker, 1999). Based on the operational definitions of creativity, scientific disciplines have endeavored to unveil the sophisticated and fascinating mental processes of divergent thinking in the human brain.

In the neuroscience of creativity, questionnaires and cognitive tasks, such as creative achievement questionnaire (CAQ) for individual creativity achievements and alternative uses task (AUT) or the Torrance Tests of Creative Thinking (TTCT) for divergent thinking, were developed to probe the underlying mental processes of creativity (Plucker, 1999; Sternberg, 1999; Carson et al., 2005). With the support of neuroimaging technologies, the neural substrates of divergent thinking have been progressively disclosed in the literature, although with great diversity across brain regions. For example, Jung et al. discovered that the CAQ was positively correlated with the left lateral orbitofrontal volume and the cortical thickness of the right angular gyrus (Jung et al., 2010). Ellamil et al. demonstrated the involvement of the medial temporal lobe in a creative drawing generation task (Ellamil et al., 2012). Using the AUT, Fink et al. demonstrated that originality was positively associated with the activation of the temporal-parietal junction, medial prefrontal cortex (mPFC) and posterior cingulate cortex (PCC) (Fink et al., 2010). Benedek et al. concluded that the left inferior parietal cortex and left prefrontal regions subserved the flexible integration of previous knowledge for constructing novel and creative ideas (Benedek et al., 2014). Abraham et al. demonstrated that the left inferior frontal gyrus (IFG) and temporal poles played major roles in AUT engagements (Abraham et al., 2012). Moreover, a meta-analysis study revealed that the left IFG was among the most predominantly activated regions associated with idea generation across task-induced brain regions

(Gonen-Yaacovi et al., 2013), suggesting its functional role in the semantic processing and conceptual expansion to expand the acquired concept into novel elements (Ward, 1994; Abraham, 2014; Boccia et al., 2015). In sum, emerging evidence indicates that brain regions in the left IFG, mPFC, parietal lobe, and medial temporal lobe are potentially involved in the mental process of divergent thinking.

Beyond the perspective of creative-task induction, divergent thinking can be taken as a spontaneous-thought process for idea incubation because of its involvement in the dynamic shifts between memory, emotion and attention (Sowden et al., 2015; Christoff et al., 2016). Therefore, a growing body of creativity-related neuroimaging studies, targeting the mental process of divergent thinking, have moved toward uncovering the long-distance brain connections using resting-state functional magnetic resonance imaging (rs-fMRI) (Takeuchi et al., 2012; Beaty et al., 2014; Wei et al., 2014). The default-mode network (DMN) has been determined to play an active role in idea generation, and the executive network appears to support idea evaluation instead (Jung et al., 2013; Beaty et al., 2016; Shi et al., 2018). Interestingly, the involvements of DMN connectivity in divergent thinking accords with the findings concerning task engagements, except for the left IFG. By separating participants into high- and low-creative groups based on their creativity score, Beaty et al. revealed greater connectivity strengths between the left IFG and posterior DMN in the high-creative group in contrast to the low-creative group (Beaty et al., 2014), which implies that the creative abilities lead to the variability of brain functionality. Recently, the DMN was further identified as a member of high-creativity neural circuits using connectome-based predictive modeling and AUT-based fMRI datasets (Beaty et al., 2018). The same study also inferred the plausible dynamicity of the network organization after divergent thinking by revealing the higher predictive power of a high-creativity network on the creativity score during AUT, compared with that under normal resting conditions. Wei et al. further suggested the possibility of dynamic resting-state functional connectivity (RSFC) changes by demonstrating that the low-creative group exhibited stronger RSFC changes between the mPFC and the middle temporal gyrus (MTG) after a creativity-related training task (Wei et al., 2014). Given the dynamic nature of creative thinking, it is speculated that the integrity of creativity-associated brain networks changes dynamically following the divergent thinking. However, are the dynamic network reconfiguration associated with the creativity performance? Do different creativity levels lead to distinct RSFC alterations? These unanswered questions are further addressed to better understand the brain functionality underlying creativity.

Herein, we establish the following hypotheses: (1) AUT-induced RSFC changes are associated with the creativity performances, and (2) individuals with different levels of creativity use distinct strategy of brain reconfigurations to comply with the challenges of AUT. To test these hypotheses, we recruited forty young participants to perform AUTs and compared their RSFC maps between two rs-fMRI sessions (Pre- and Post-AUT). The levels of creativity were assessed through splitting the participants into two groups (high- and

low-creative) based on the CAQ scores and we tested the influence of AUT on RSFC changes for each creativity group.

MATERIALS AND METHODS

Participants

A total of 42 right-handed healthy young adults were recruited in this study. The inclusion criteria were as follows: (a) non-smoking and without drug addiction; (b) not pregnant; (c) no history of neurological and psychiatric disorder history, and (d) no metal body implants. The entire protocol was approved by the Research Ethics Committee of National Taiwan University (NTU-REC No. 201407EM028). Because two participants could not complete the entire experimental procedure, their datasets were excluded in the analysis. Accordingly, the following results are reported for the remaining 40 participants (19 males, mean age = 24.4 ± 2.8 years, range: 21–33 years).

Stimuli of AUT

The stimuli consisted of 56 grayscale images of objects from daily life with background removed. Prior to the actual fMRI experiment, we rated the number of alternative usages for each object stimulus from a separate group of participants ($N = 15$, 12 males, mean age = 23.4 ± 3 years, range: 20–30 years). During the rating, the 56 objects were presented consecutively to the participants with inter-trial intervals (ITIs) of 26 s. The participants were instructed to think and express alternative usages for each object within 20 s. The mean number of answer for each object was 2.4. We categorized the difficulty of AUT into three levels as follows: normal, 1.99–2.99 answers per picture; difficult, <1.99 answers per picture; and easy, >2.99 answers per picture. Subsequently, we used the task materials to design the AUT in MRI environments. All of the objects were presented against a white background for both rating and fMRI experiment.

Experimental Procedure

After consenting, the participants were instructed to complete a CAQ, containing eight aspects of creativity (Carson et al., 2005), and to undergo the fMRI experiments. The fMRI experiment included seven sessions in total: two resting-state, two 0-back and three AUT sessions. The participants underwent the first rs-fMRI scan (Pre-AUT), performed three consecutive AUTs with two 0-back tasks in between, and went through the second rs-fMRI scan (Post-AUT) at the end. The participants only received the instruction to generate novel ideas after viewing pictures in the AUT sessions without practices to ensure the Pre-AUT resting state was irrelevant to divergent thinking. The multiple AUTs were designed to enhance the divergent thinking process with undemanding tasks (0-back) in between (Baird et al., 2012). Herein, we considered all AUTs and 0-backs as instructed divergent-thinking engagements, and adopted the two sessions of resting-state scans (Pre-AUT and Post-AUT) to discern the alterations of RSFC through the divergent thinking process.

fMRI Acquisition and Scanning Parameters

MRI experiments were conducted by a 3T MRI scanner (Prisma, Siemens, Erlangen, Germany) with 20-channel head coil at

National Taiwan University. During the experiment, a high-resolution T1-weighted anatomical image was initially scanned using the 3D-MPRAGE sequence with $256 \times 256 \times 192$ matrix size; $0.93 \times 0.93 \times 0.93$ mm³ resolution; inversion time (TI) = 900 ms; repetition time (TR) = 2,000 ms, echo time (TE) = 2.3 ms; flip angle (FA) = 8°; bandwidth = 200 Hz/pixel; NEX = 1. Total scan time was 6 min 14 sec. The fMRI protocol was using a single-shot gradient-echo echo-planar imaging (GE-EPI) sequence with following imaging parameters: TR = 2 s, TE = 34 ms, FA = 84°, bandwidth = 3,005 Hz/pixel, matrix size = $64 \times 64 \times 33$, and FOV = 210×210 mm². Stimuli were presented via E-prime 2.0 (Psychology Software Tools, Pittsburgh, PA, USA) with a back-projection projector in 800×600 resolution. Participants viewed the stimuli using a mirror mounted on the head-coil and the viewing field was 8.4° (horizontal) by 6.3° (vertical) at a viewing distance of 420 cm. Participants were instructed to respond with a button press using the index and middle fingers of their right hand (Lumina response pad; Cedrus, San Pedro, CA, USA).

During the rs-fMRI scans, the participants were instructed to stay still with their eyes open, to not fall into sleep and to not think of anything in particular. The rs-fMRI scan contained 180 measurements (6 min) per session. During the AUTs, the participants were instructed to view gray-scale pictures of various objects from daily life (through a projector), and to think of as many alternative usages as possible within a fixed duration (20 s). Each AUT session contained 12 trials, including one easy-level picture, one difficult-level picture, and ten normal-level pictures to balance the difficulty level, with fixation and cues of 6 s in between; within each trial one single picture was presented for 20 sec. During the 20 s of the picture presentation, the participants were instructed to press the button immediately when they thought of a special usage; and the button-press numbers and the response time were recorded. Each AUT task contained 156 scans (5 min 12 s) in total. After the cessation of each AUT, the participants were asked to recall the answers retrospectively, and the answers were manually recorded by the experimenter. During the 0-back tasks, one English letter was displayed on screen for 0.5 s, followed by a cross fixation for 1.5 s. The participants were instructed to respond using their right hand to press the right button when the letter appearing on the screen was “X”; otherwise, they were to press the left button. The 0-back fMRI contained 152 scans (5 min 4 s) in each session. The datasets generated for this study are available on request to the corresponding author.

AUT Performance

Beyond the creativity assessment obtained through the CAQ, the two creativity indices of AUT (e.g., originality and fluency) were also evaluated for each participant according to the scoring method of previous studies (Hao et al., 2015). Fluency was estimated on the basis of the total number of ideas given during the AUT trials. Originality was estimated on the basis of statistically infrequent responses using the following means. The generated answers from the AUT were collected into a comprehensive lexicon for comparison. Three trained raters independently assessed the originality of the AUT performance

for every participant. If 10% or less of the participants in the sample gave the same response then it was given a score of “1,” and all other responses received scores of “0.” The inter-rater agreement was with intra-class correlation coefficient of 0.97. Subsequently, the originality scores of the AUT performance assigned by the three raters were averaged for every participant.

fMRI Analyses

The data were processed using Analysis of Functional Neuro Images (AFNI) (Cox, 1996). We processed the task data according to the following steps: (a) motion correction; (b) coregistration; (c) smoothing with full width at half maximum (FWHM) = 6 mm, and (d) normalization to the MNI space. During the AUT sessions, participants' responses were recorded and used as the events for divergent thinking, enabling the event-related fMRI analysis. Using the events convolved with the canonical hemodynamic function, we subsequently applied the generalized linear model (GLM) to generate the activation map for each AUT and calculated the average beta map across three AUTs for each individual. The consequent AUT results were equally divided into two groups ($n = 20$ for each of the high- and low-creative groups) according to their CAQ scores (the cut-off score for the CAQ was 8), and a two-sample t -test was conducted to generate the contrast maps of AUT brain activity.

The rs-fMRI data were processed according to the following steps: (a) field-map correction; (b) motion correction; (c) coregistration; (d) detrending and filtering (0~0.1 Hz); (e) smoothing with FWHM = 5 mm, and (f) normalization; and (g) the physiological noise removal including white matter, cerebrospinal fluid and six motion parameters. Subsequently, the functional connectivity maps before and after AUT were established through the seed-correlation analysis, targeting on two specific brain networks: (1) the bilateral posterior cingulate cortex (PCC) [± 3 , -53 , 26] for assessing DMN (Van Dijk et al., 2010; Yan et al., 2013), and (2) the seeding at the peak of AUT contrasts between the high- and low-creative groups. Subsequently, the group-based RSFC maps underwent the following statistical analyses to estimate the brain reconfigurations.

Statistical Analyses

All group analyses of the fMRI results were conducted on the basis of analysis of variance (ANOVA) and t -test using AFNI. The group difference in brain activity during AUTs were evaluated through a two-sample t -test between the high- and low-creative groups. The two brain networks associated with divergent thinking were compared for RSFC according to the 2 (high- and low-creative) \times 2 (pre- and post-AUT) two-way ANOVA (3dMVM), inclusive of the mean framewise displacement to minimize the motion effect on RSFC findings. The multiple comparisons in the group analysis were corrected through the 3dClustSim approach with the auto-correlation function and the significance level was $p < 0.05$ (uncorrected $p < 0.001$, cluster size = 80 mm^3). To further determine the relationships between creativity scores and the divergent thinking effect of RSFC, the Pearson correlation analysis was used to measure the association between the behavior indices of each individual (originality

and fluency) and the RSFC changes (Post-Pre) from selected regions of interested (ROIs). The centers of the spherical ROIs (radius = 5 mm) were placed at the peak value of the significant changes in the contrast maps. The effect of divergent thinking was assessed by calculating the RSFC differences between Pre- and Post-AUT after Fisher's z transformation of the correlation coefficients. The z -transformed RSFC changes from the selected ROIs were correlated with the creativity indices (originality and fluency) with false-discovery rate (FDR) correction across ROI-wise comparisons. The statistical calculations were performed using SPSS 20 (IBM Statistical Package for the Social Sciences).

RESULTS

Creativity Performances: CAQ and AUT

The average CAQ score for the 40 participants was 10.8 ± 10.4 , ranging from 1 to 62. Regarding the overall AUT performances, the evaluated originality and fluency were 0.33 ± 0.11 and 3.1 ± 1.0 , respectively. The CAQ scores were adjusted by the log-transform (Form and Kaernbach, 2018), and the log(CAQ) scores exhibited a positive trend toward the AUT fluency ($r = 0.30$, $p = 0.06$), without prominent relationships with originality ($r = 0.22$, $p = 0.16$) for all participants. To test the difference in brain activity between high and low creativity individuals, the following results were divided into two groups with equal sample size ($n = 20$ for each group); where their CAQ scores were 16.6 ± 12.2 for the high-creative group and 5.0 ± 1.9 for the low-creative group. In addition, the two groups showed no significant differences in age ($p = 0.06$), gender ($p = 0.53$) and motion during the two rs-fMRI scans (framewise displacement: $p = 0.39$).

Brain Activity Associated With AUT

The AUT-associated brain activations were located at the left superior temporal gyrus, anterior cingulate gyrus, declive of the cerebellum, caudate and occipital visual cortex (**Figure 1**, corrected $p < 0.01$). **Figures 1A,B** illustrate the average activation maps of AUT for the high-creative and low-creative groups, respectively, where the high-creative group exhibited relatively broad spatial extents in AUT engagements relatively. Negative activations were found lateralized to the right hemisphere, specifically located at the right PCC and right inferior parietal lobule. **Figure 1C** depicts the contrast between the two groups by independent t -test (corrected $p < 0.05$). A difference between the groups was observed only in the left IFG (BA47, [-32 , 24 , -14]), which was taken as the seed localization for the subsequent RSFC analyses. **Supplementary Figure 1** shows the overall AUT activation map across all participants and **Supplementary Table 1** lists the detailed description of AUT-related brain activities.

Functional Connectivity: IFGN and DMN

To further elucidate the alterations of brain integrity over the divergent thinking process, we applied a seed-correlation approach to investigate RSFC. Two brain networks were targeted with the seed locations in bilateral PCC for DMN and left IFG for IFG-associated network (IFGN) to assess the group

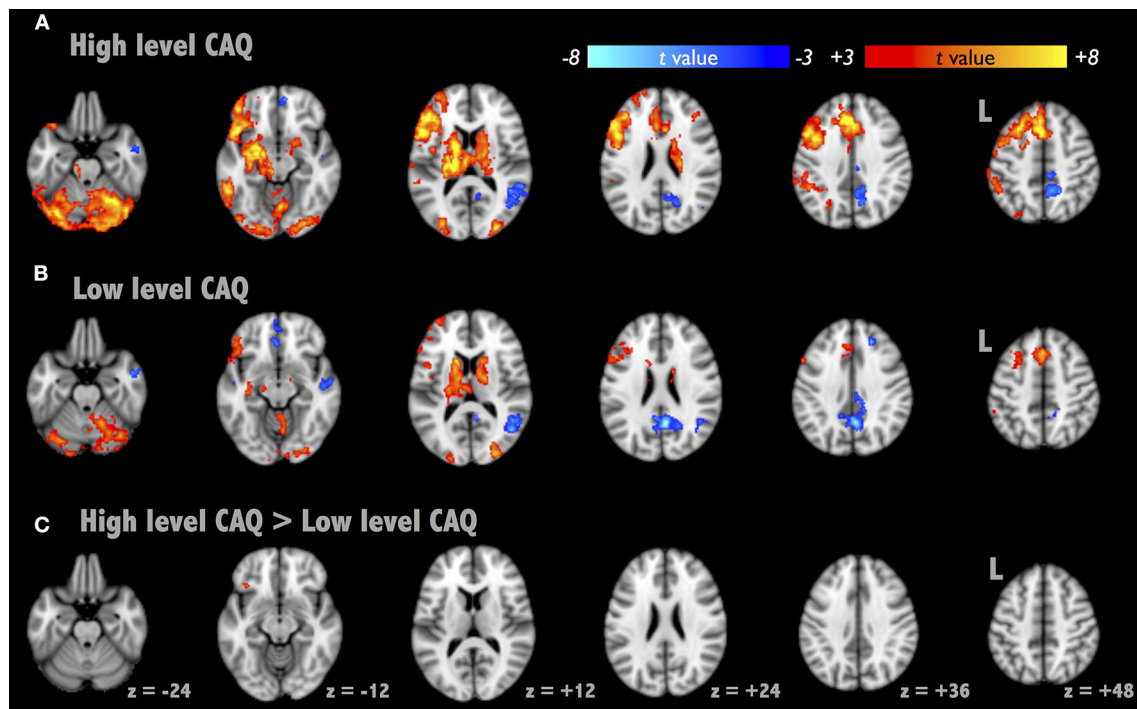


FIGURE 1 | Brain activation maps of Alternative Uses Task (AUT). We presented the activation maps divided into two groups as follows: high- and low-creative [(A,B), respectively, segregated by CAQ scores]. (C) The AUT contrast map between the two groups, where the contrast region subsided in the left inferior frontal gyrus (BA 47).

(high- vs. low-creative) and divergent thinking (Pre- vs. Post-AUT) effects. The upper panel of **Figure 2** illustrates the DMN connectivity maps exhibiting strong connections to the bilateral medial and middle frontal, PCC, and bilateral angular gyrus (corrected $p < 0.01$). The lower panel of **Figure 2** illustrates the IFGN connectivity maps (corrected $p < 0.01$). The IFGN was associated with the bilateral superior frontal, medial frontal, anterior cingulate, insula, superior, and middle temporal gyri. The detailed information of the group \times divergent thinking comparisons (high-creative, low-creative, Pre-AUT and Post-AUT) for IFGN and DMN is listed in **Table 1**. Prior to the AUT engagements (Pre-AUT), the only significant difference in IFGN between the groups was in the right angular gyrus (AG) and right inferior parietal lobule (IPL), but no group difference was observed in DMN at the baseline level. However, the RSFC of both networks changed after performing the AUT. **Table 1** demonstrates that the group difference of RSFC_{DMN} became prominent after AUT engagements, and the RSFC_{IFGN} showed different network reconfigurations between the two groups.

Association Between Functional Connectivity and Creativity Scores

An ROI analysis was conducted to determine the relationship between RSFC changes and creativity scores. The spherical ROIs were placed according to the results in **Table 1**. **Figure 3** presents the RSFC changes between Pre-AUT and Post-AUT and their relationships with AUT originality and fluency in

both groups. The RSFC changes in IFGN at the right AG and right IPL were positively correlated with the AUT performances: originality- $\Delta\text{FC}_{\text{IFGN-AG}}$: $r = 0.51$, $p < 0.02$; fluency- $\Delta\text{FC}_{\text{IFGN-AG}}$: $r = 0.49$, $p < 0.03$; originality- $\Delta\text{FC}_{\text{IFGN-IPL}}$: $r = 0.63$, $p < 0.001$) in the high-creative group, but the correlations in the low-creative group were non-significant. The Steiger's Z-test presented significant group differences in the correlation coefficients of originality- $\Delta\text{FC}_{\text{IFGN-AG}}$ ($p < 0.001$), fluency- $\Delta\text{FC}_{\text{IFGN-AG}}$ ($p < 0.004$) and originality- $\Delta\text{FC}_{\text{IFGN-IPL}}$ ($p < 0.001$). **Supplementary Table 2** summarizes all correlations between RSFC and creativity assessments [log(CAQ), AUT fluency and originality].

DISCUSSION

We demonstrated that the divergent thinking process (e.g., AUT) could modulate the RSFC of DMN and IFGN, which were associated with creativity performance. The results indicated that the baseline RSFC_{DMN} networks (Pre-AUT) were similar between the two groups, but the baseline RSFC_{IFGN} could reflect the group difference. Then, the IFGN presented significant network reorganization while the DMN did not reorganize after AUT engagements. Intriguingly, such IFGN reconfiguration was associated with AUT performances (originality and fluency, **Figure 3**). These findings supported that the dynamic changes of intrinsic network connections could reflect the cognitive performances of divergent thinking. However, the high- and

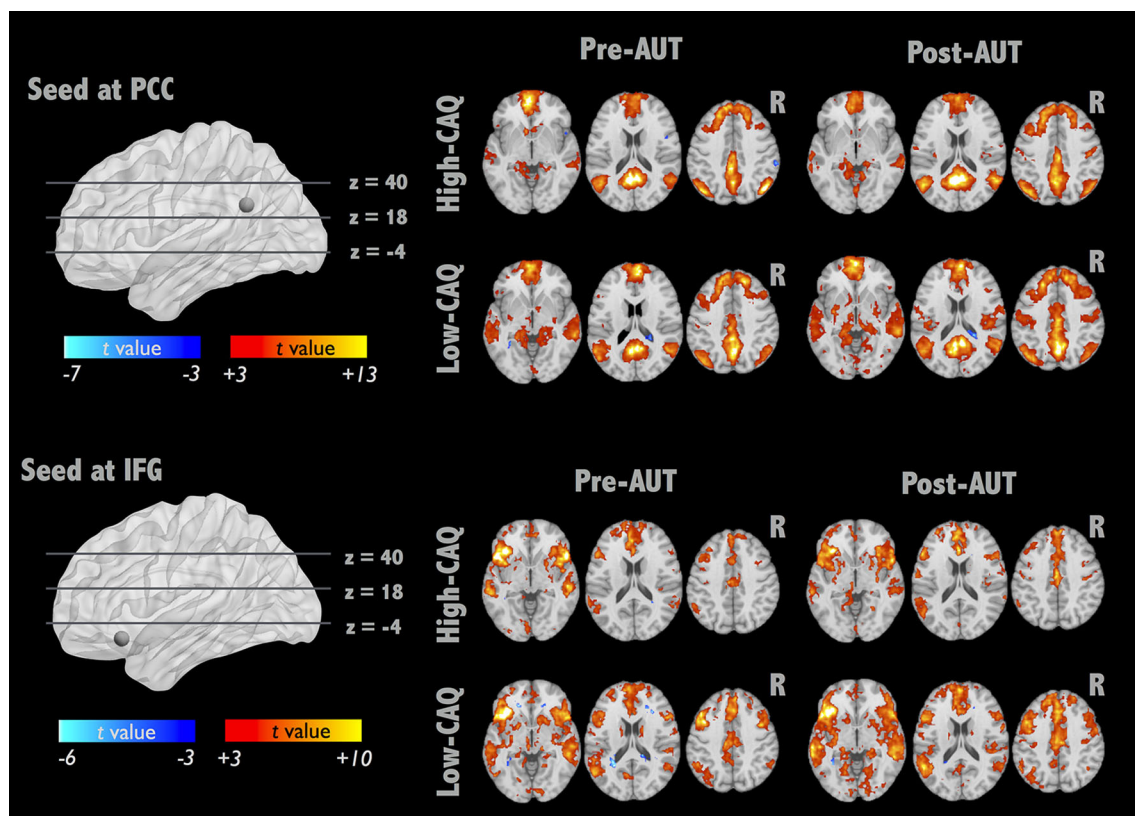


FIGURE 2 | IFGN and DMN RSFC patterns before and after AUT. For each network, the top panel displays the one-sample RSFC map of Pre-AUT and Post-AUT for the high-creative group, and the lower panel refers to the low-creative group.

low-creative groups demonstrated distinct network—creativity associations as follows: positive relationships were evident between originality and RSFC changes in the high-creative group but the low-creative group showed non-significant correlations (**Supplementary Table 2**), further indicating the dissimilar functional architectures involved in divergent thinking among individuals with diverse creativity achievements. These findings suggest that the RSFC_{IFGN} can change dynamically in accordance with the divergent thinking process, and the populations with various creativity levels may employ distinct reconfigurations after divergent thinking.

Network Reconfiguration After Divergent Thinking

The integrity of brain network can be modulated by the previous cognitive engagements within minutes, and the functional reorganization may be important to the cognitive performances. Previous research investigated the correlation between the IFG connectivity and the subsequent memory (Stevens et al., 2010), disclosing the connectivity facilitates the recognition accuracy. In the current study, we focused on the dynamic RSFC changes of the two networks, DMN and IFGN, corresponded to the networks related with high-creative thinking ability reported in literature (Beatty et al., 2018). Interestingly, the same study also demonstrated the evidence of dynamic RSFC alterations

following divergent thinking. In their supporting information, Beatty et al. presented fairly similar correlations between network strengths and creativity scores based on a resting-state dataset ($r = 0.13$ for high-creative networks and $r = 0.11$ for low-creative networks). However, the RSFC–creativity correlation became elevated for high-creative networks when performing the AUT ($r = 0.35$ and 0.28 for two separate datasets), but the correlation turned out non-significant for low-creative networks during AUT engagements ($r = 0.02$ to -0.04 for two separate datasets). The evidence implies that dynamic network reorganizations in AUT can be associated with the creativity performances. In this study, assuming that the short-term network reorganizations induced by AUT could sustain after the task cessation (the second resting state in Post-AUT), we demonstrated that dynamic RSFC alterations were associated with the AUT performance (originality and fluency). Specifically, we prescribed the ROIs from both the creativity trait effect (High-creative > Low-creative defined by CAQ in **Table 1**) and the creativity state effect (Post-AUT > Pre-AUT), because the CAQ (trait) and AUT performances (state) were highly correlated within each group (**Supplementary Table 2**). Our results showed that DMN demonstrated the trait difference between groups; however, no prominent state effect was found in DMN. Relatively, IFGN contained both trait and state effect (**Table 1**), indicating that the IFGN indeed presented

TABLE 1 | Group comparison of IFGN and DMN RSFC before and after AUT.

Brain area	t-value	Voxel size	Peak x	Peak y	Peak z
DMN (SEED AT LEFT INFERIOR PARIETAL GYRUS [$\pm 3, -53, 26$])					
Pre-AUT: High-creative > Low-creative					
--					
Post-AUT: High-creative > Low-creative					
Postcentral gyrus (right)	-0.31	41	54	-26	20
Inferior frontal gyrus (right)	-0.34	30	46	34	2
Thalamus (right)	0.24	15	4	-36	10
Middle temporal gyrus (left)	-0.26	12	-46	0	-24
Superior temporal gyrus (left)	-0.31	12	-44	4	-8
High-creative group: Post-AUT > Pre-AUT					
--					
Low-creative group: Post-AUT > Pre-AUT					
--					
IFGN (SEED AT LEFT INFERIOR PARIETAL GYRUS [$-32, 24, -14$])					
Pre-AUT: High-creative > Low-Creative					
Angular gyrus (right)	-0.28	71	42	-78	28
Inferior parietal lobule (right)	-0.28	56	40	-62	40
Post-AUT: High-creative > Low-creative					
Middle frontal gyrus (right)	-0.26	76	52	2	44
Anterior cingulate cortex (left)	-0.21	25	-10	22	-10
High-creative group: Post-AUT > Pre-AUT					
Inferior parietal lobule (left)	0.18	11	-40	-66	38
Low-creative group: Post-AUT > Pre-AUT					
Inferior occipital gyrus (right)	0.18	28	36	-80	-8
Parahippocampal gyrus (left)	0.2	9	-6	-36	0

the network reconfiguration after AUT engagements. The concept of network reconfiguration was supported by Wei et al. (2014), in which they demonstrated that the RSFC strengths was modulated in general after the performance of a cognitive stimulation task. The dynamic alterations of RSFC following cognitive tasks can be regarded as a short-term functional reconfiguration of brain circuits to facilitate associated cognitive tasks, which fits the neurophysiological perspective of the dynamo framework (Kopell et al., 2014). For example, Wang et al. presented the dynamic reorganizations of DMN during and after a word-picture matching task, so as the corresponding whole-brain small-world topology (Wang et al., 2012). However, previous studies did not specifically present the relationship between the network reconfiguration and the cognitive performances. In the current work, **Figure 3** illustrates that dynamic IFGN reconfigurations were in accordance with the creativity performances, supporting the short-term reconfiguration and sustenance of network integrity underlying the mental process of divergent thinking.

Regional Functionality in Divergent Thinking

Three brain regions in the left hemisphere were reported to be involved in divergent thinking—the IFG, pre-/post-central gyri, and the MTG (Boccia et al., 2015), and these are similar to our results of brain activity in AUT. Generally, the lateralization

of brain function in the left hemisphere in AUT is considered to be the verbal processing and semantic memory (Thompson-Schill, 2003). In addition, the left hemisphere executes the functions of interpreting and reasoning about the sentences and the causal relationships (Marinsek et al., 2014). By presenting the CAQ-based group difference of brain activation in AUT (**Figure 1**), the left IFG was highlighted in the performance of creativity, regarded as the functional localizer for the following RSFC analysis.

Figure 3 illustrates the relationship between originality and the corresponding RSFC changes after AUTs; the RSFC changes are prominent between the left IFG, right AG, and right IPL. Previously, the SPL/IPL was the core for top-down attention or abstract thinking (Shomstein, 2012), and the IPL was determined to usually play roles involving the voluntary orienting of attention to a location (Corbetta et al., 2000). In addition, the left IFG and right IPL involved the verbal working memory (LaBar et al., 1999), and the strength of RSFC between the left IFG and right IPL decreased in high working memory loading (Liu et al., 2017). Furthermore, a previous study reported that the low working memory loading task could help the divergent thinking and (Baird et al., 2012). For the multifunction of AG, the right side was determined to play the role of spatial cognition for perceptual learning and shifting attention to relevant information (Seghier, 2013). Therefore, our result demonstrated that the high-creative group leveraged from the frontoparietal reorganizations for

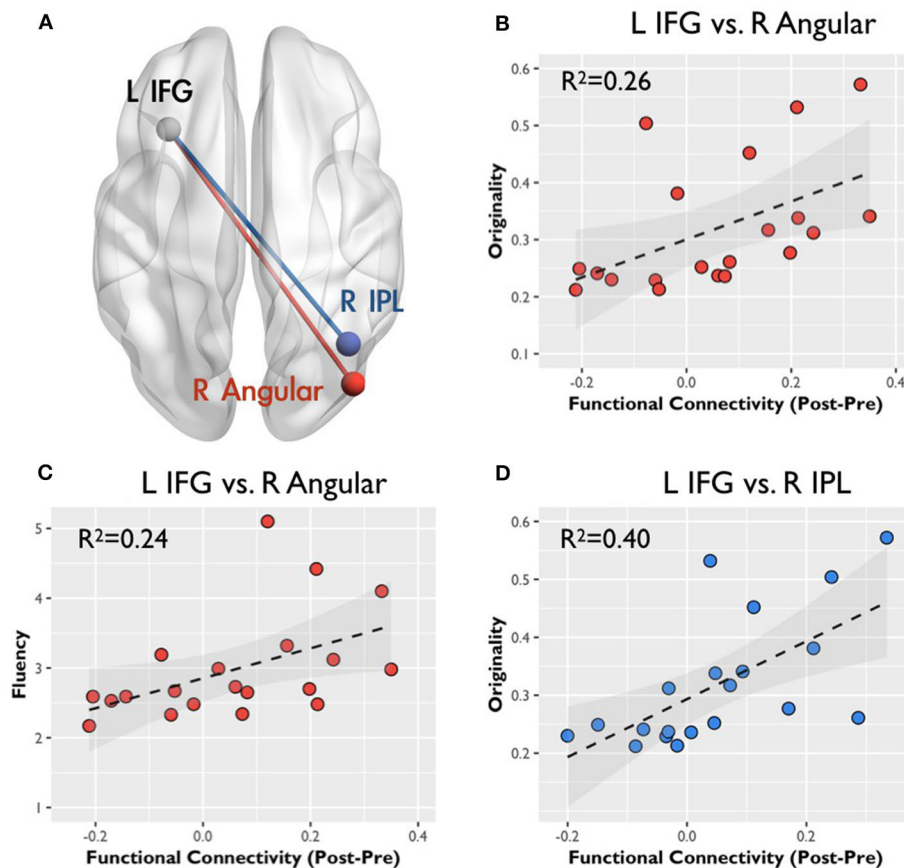


FIGURE 3 | Correlation between the functional connectivity and creative behavior: color regions of two connectivity indicate spatial locations. **(A)** The gray spheres indicate the seeds for functional connectivity, the red line represents the connectivity between the left inferior frontal gyrus (IFG) and the right angular gyrus (AG), and the blue line indicates the connectivity between the left IFG and the right inferior parietal lobule (IPL). All scatter plots display the correlation between RSFC changes (in z-value) and CAQ, originality, and fluency scores in r squares. **(B,C)** Represent the relation with the connectivity change between the left IFG and the right AG, and **(D)** correspond to the relation with the connectivity change between the left IFG and the right IPL.

elevating the abstract thinking, verbal working memory or spatial cognition involved in AUT. In contrast, the changes of RSFC in the low-creative group showed no relationship associated with the creative indexes. Following the finding, our results revealed the possibility that the low-creative group might adopt distinct strategies with respect to the semantic or sensory processing in AUT.

As evident in **Supplementary Table 1**, the majority of AUT-induced negative activations were located in the right hemisphere (the superior temporal gyrus, MTG, medial frontal gyrus, precuneus, precentral gyrus, and superior frontal gyrus) in both groups. Previously, the mechanism of the negative activation may have originated from the neural inhibition hypothesis, the affected neurotransmitter caused the reduction of local cerebral blood flow or the elevation of the cerebral metabolic rate of oxygen consumption in the cerebrovasculature (Sten et al., 2017). Interestingly, Benedek et al. also presented similar lateralized negative activation during idea generation (Benedek et al., 2014). To further verify the role of negative AUT activations in creativity, the negative brain activations were associated with

the log-transform CAQ, originality, and fluency in both the high- and low-creative groups. We noted significantly positive correlations between the log(CAQ) and the right MTG [52, −10, −16] activity ($r = 0.52$, $p < 0.02$) in the high-creative group, as well as a significantly positive relationship between originality and the right MTG activity ($r = 0.51$, $p < 0.02$). However, the association between the negative activation and creativity performances is beyond the scope of the current work. Future studies are warranted to discern the underlying mechanism of divergent thinking.

Disparity Between High- and Low-Creative Groups

About grouping of CAQ scores, we chose the median split because of a positively-skewed distribution of CAQ in Carson's report. Additionally, the average CAQ score of Carson's report was as high as 14.4, the cut-off line would be around 8 to 9 when we adopted the median split (Carson et al., 2005). Therefore, the cut-off threshold of 8 points was in agreement with Carson's report. Although we used the CAQ (the creative

traits) for splitting participants into two groups, we confirmed the functional brain distinction between the high- and low-creative groups when performing AUT (**Figure 1**), revealing the active role of the left IFG in AUT engagements. By definition, the participants with high CAQ scores were believed to possess the capability of openness to experience (Carson et al., 2005), in part explaining that the high-creative group uses the brain systems of abstract thinking to achieve superior originality. Moreover, the between-group comparison disclosed that the DMN connectivity did not exhibit significant differentiation at the baseline level before AUT engagements (**Table 1**). In IFGN, the baseline connectivity strength between the left IFG and right IPL in the high-creative group was higher than that in the low-creative group at the Pre-AUT condition. This finding was contradictory to the previous report (Beatty et al., 2014). The disparity might be originated from the cultural difference, because previous study showed that frontal-parietal attentional control network involves in the visual and attention task and these functions exhibits cultural disparity (Han and Northoff, 2008; Hedden et al., 2008). Then, the connection from IFG to IPL involved the lexical judgment in Chinese studies (Deng et al., 2012). Future studies are warranted to prove such conjecture. In addition, when comparing the RSFC_{IFGN} changes between the two groups in **Table 1**, the low-creative group demonstrated more Pre-Post connectivity differences in the posterior brain, whereas the high-creative group mostly remained unchanged. Wei et al. observed similar between-group differences in RSFC between MTG and MPFC (Wei et al., 2014). They suggested that the group with higher TTCT scores exhibited fewer RSFC differences after creativity training, though the RSFC changes were not associated with the creativity performances. Collectively, it is speculated that each individual may possess specific brain-network plasticity to facilitate the performances of divergent thinking, where such brain reconfigurations are distinct between the high- and low-creative groups.

Limitation

This study has several limitations. First, the sample size of the current study was above the request of 36 samples for sufficient fMRI replicability (Turner et al., 2018); however, the statistical power might be reduced after splitting the samples into two equal-size groups. Compared with literature, the AUT-induced brain activities were in well-agreement with previous findings, but the RSFC might be inconsistent with other studies. The sample size higher than 20 in each group is recommended for future neuroimaging studies with AUT engagements. Second, we observed only the two most addressed functional networks in AUT, namely DMN and IFGN, by which we might miss additional AUT-related brain reconfigurations in other brain networks, such as the reported global architectures involved in divergent thinking (Beatty et al., 2018). Future studies are warranted to test the whole-brain functional changes other than DMN and IFGN. Third, the group separation (based on CAQ) and their creative performances (i.e., originality and fluency) were all defined by the subjective preferences from the participants or the raters. However, creativity judgment is a relatively challenging task due to the lack of objective definitions,

as aforementioned. Therefore, we used the AUT activity to verify the effectiveness of group separation and employed three raters to minimize the subjectivity involved in the creativity scoring. Fourth, we did not include any other cognitive tasks to test whether such network reconfiguration is specific to the divergent thinking. It was unanswered because the assumption that the network plasticity is subject to distinct cognitive tasks has yet been studied systematically. The post-cognition network plasticity is pending for further investigations in the near future. At last, the causal relation between the network reorganization and creativity performance remains limited, because the findings in this study was built upon the assumption that the AUT-associated network reconfigurations can sustain in a short period of time after the cessation of tasks. The short-term sustenance of functional organizations in brain circuits shall be further tested before confirming the causal relationship between network plasticity and creativity.

CONCLUSION

Using multiple AUTs for creativity engagements, we presented the prolonged changes of RSFCs (DMN and IFGN) correlated with the performances of divergent thinking. Furthermore, individuals with different creativity level (high- and low-creative groups) might present diverse alterations of RSFC changes. Before divergent thinking, both high- and low-creative group did not exhibit significant difference of DMN connectivity, but the group difference was highlighted after AUT engagements. Meanwhile, the IFGN indeed presented the network reconfiguration after divergent thinking. Furthermore, the ΔFC_{IFG-AG} and $\Delta FC_{IFG-IPL}$ positively contributed to the AUT performances in the high-creative group, but no prominent brain-behavior relation was found in the low-creative group. These findings indicated that divergent-thinking performances could be modulated by distinct creativity traits and diverse brain-network reconfigurations.

DATA AVAILABILITY STATEMENT

The datasets generated and analyzed in the current study are not publicly available due to IRB but are available from the corresponding author on reasonable request. Requests to access the datasets should be directed to Changwei W. Wu, sleepbrain@tmu.edu.tw.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Research Ethics Committee of National Taiwan University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

H-YW: data collection, analysis, and writing. B-CK and C-MH: experiment design and writing. P-JT and A-LH: analysis and

writing. L-MH: analysis. C-YL: data collection. J-HC: suggestion of writing. CW: the main idea, experimental design, and writing. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by Ministry of Science and Technology, Taiwan (MOST 104-2420-H-008-002 and MOST 108-2321-B-038-005-MY2) and Taipei Medical

University (TMU105-AE1-B11) to CW. We sincerely thank the technical supports of Imaging Center for Integrated Body, Mind and Culture Research, National Taiwan University.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnhum.2020.571118/full#supplementary-material>

REFERENCES

- Abraham, A. (2013). The promises and perils of the neuroscience of creativity. *Front. Hum. Neurosci.* 7:246. doi: 10.3389/fnhum.2013.00246
- Abraham, A. (2014). Creative thinking as orchestrated by semantic processing vs. cognitive control brain networks. *Front. Hum. Neurosci.* 8:95. doi: 10.3389/fnhum.2014.00095
- Abraham, A., Pieritz, K., Thybusch, K., Rutter, B., Kröger, S., Schweckendiek, J., et al. (2012). Creativity and the brain: Uncovering the neural signature of conceptual expansion. *Neuropsychologia* 50, 1906–1917. doi: 10.1016/j.neuropsychologia.2012.04.015
- Baird, B., Smallwood, J., Mrazek, M. D., Kam, J. W. Y., Franklin, M. S., and Schooler, J. W. (2012). Inspired by distraction: mind wandering facilitates creative incubation. *Psychol. Sci.* 23, 1117–1122. doi: 10.1177/0956797612446024
- Beaty, R. E., Benedek, M., Silvia, P. J., and Schacter, D. L. (2016). Creative cognition and brain network dynamics. *Trends Cogn. Sci.* 20, 87–95. doi: 10.1016/j.tics.2015.10.004
- Beaty, R. E., Benedek, M., Wilkins, R. W., Jauk, E., Fink, A., Silvia, P. J., et al. (2014). Creativity and the default network: a functional connectivity analysis of the creative brain at rest. *Neuropsychologia* 64, 92–98. doi: 10.1016/j.neuropsychologia.2014.09.019
- Beaty, R. E., Kenett, Y. N., Christensen, A. P., Rosenberg, M. D., Benedek, M., Chen, Q., et al. (2018). Robust prediction of individual creative ability from brain functional connectivity. *Proc. Natl. Acad. Sci. U.S.A.* 115, 1087–1092. doi: 10.1073/pnas.1713532115
- Benedek, M., Jauk, E., Fink, A., Koschutnig, K., Reishofer, G., Ebner, F., et al. (2014). To create or to recall? Neural mechanisms underlying the generation of creative new ideas. *NeuroImage* 88, 125–133. doi: 10.1016/j.neuroimage.2013.11.021
- Boccia, M., Piccardi, L., Palermo, L., Nori, R., and Palmiero, M. (2015). Where do bright ideas occur in our brain? Meta-analytic evidence from neuroimaging studies of domain-specific creativity. *Front. Psychol.* 6:1195. doi: 10.3389/fpsyg.2015.01195
- Carson, S. H., Peterson, J. B., and Higgins, D. M. (2005). Reliability, validity, and factor structure of the creative achievement questionnaire. *Creat. Res. J.* 17, 37–50. doi: 10.1207/s15326934crj1701_4
- Christoff, K., Irving, Z. C., Fox, K. C. R., Spreng, R. N., and Andrews-Hanna, J. R. (2016). Mind-wandering as spontaneous thought: a dynamic framework. *Nat. Rev. Neurosci.* 17, 718–731. doi: 10.1038/nrn.2016.113
- Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., and Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nat. Neurosci.* 3, 292–297. doi: 10.1038/73009
- Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29, 162–173. doi: 10.1006/cbmr.1996.0014
- Deng, Y., Guo, R., Ding, G., and Peng, D. (2012). Top-down modulations from dorsal stream in lexical recognition: an effective connectivity fMRI study. *PLoS ONE* 7:e33337. doi: 10.1371/journal.pone.0033337
- Ellamil, M., Dobson, C., Beeman, M., and Christoff, K. (2012). Evaluative and generative modes of thought during the creative process. *NeuroImage* 59, 1783–1794. doi: 10.1016/j.neuroimage.2011.08.008
- Fink, A., Grabner, R. H., Gebauer, D., Reishofer, G., Koschutnig, K., and Ebner, F. (2010). Enhancing creativity by means of cognitive stimulation: evidence from an fMRI study. *NeuroImage* 52, 1687–1695. doi: 10.1016/j.neuroimage.2010.05.072
- Form, S., and Kaernbach, C. (2018). More is not always better: the differentiated influence of empathy on different magnitudes of creativity. *Eur. J. Psychol.* 14, 54–65. doi: 10.5964/ejop.v14i1.1432
- Gonen-Yaacovi, G., de Souza, L. C., Levy, R., Urbanski, M., Josse, G., and Volle, E. (2013). Rostral and caudal prefrontal contribution to creativity: a meta-analysis of functional imaging data. *Front. Hum. Neurosci.* 7:465. doi: 10.3389/fnhum.2013.00465
- Guilford, J. P. (1967). Creativity: yesterday, today and tomorrow. *J. Creat. Behav.* 1, 3–14. doi: 10.1002/j.2162-6057.1967.tb00002.x
- Han, S., and Northoff, G. (2008). Culture-sensitive neural substrates of human cognition: a transcultural neuroimaging approach. *Nat. Rev. Neurosci.* 9, 646–654. doi: 10.1038/nrn2456
- Hao, N., Wu, M., Runco, M. A., and Pina, J. (2015). More mind wandering, fewer original ideas: be not distracted during creative idea generation. *Acta Psychol.* 161, 110–116. doi: 10.1016/j.actpsy.2015.09.001
- Hedden, T., Ketay, S., Aron, A., Markus, H. R., and Gabrieli, J. D. E. (2008). Cultural influences on neural substrates of attentional control. *Psychol. Sci.* 19, 12–17. doi: 10.1111/j.1467-9280.2008.02038.x
- Jung, R. E., Mead, B. S., Carrasco, J., and Flores, R. A. (2013). The structure of creative cognition in the human brain. *Front. Hum. Neurosci.* 7:330. doi: 10.3389/fnhum.2013.00330
- Jung, R. E., Segall, J. M., Jeremy Bockholt, H., Flores, R. A., Smith, S. M., Chavez, R. S., et al. (2010). Neuroanatomy of creativity. *Hum. Brain Mapp.* 31, 398–409. doi: 10.1002/hbm.20874
- Kopell, N. J., Gritton, H. J., Whittington, M. A., and Kramer, M. A. (2014). Beyond the connectome: the dynome. *Neuron* 83, 1319–1328. doi: 10.1016/j.neuron.2014.08.016
- LaBar, K. S., Gitelman, D. R., Parrish, T. B., and Mesulam, M. (1999). Neuroanatomic overlap of working memory and spatial attention networks: a functional MRI comparison within subjects. *NeuroImage* 10, 695–704. doi: 10.1006/nimg.1999.0503
- Liu, H., Yu, H., Li, Y., Qin, W., Xu, L., Yu, C., et al. (2017). An energy-efficient intrinsic functional organization of human working memory: a resting-state functional connectivity study. *Behav. Brain Res.* 316, 66–73. doi: 10.1016/j.bbr.2016.08.046
- Marinsek, N., Turner, B. O., Gazzaniga, M., and Miller, M. B. (2014). Divergent hemispheric reasoning strategies: reducing uncertainty versus resolving inconsistency. *Front. Hum. Neurosci.* 8:839. doi: 10.3389/fnhum.2014.00839
- Plucker, J. A. (1999). Is the proof in the pudding? reanalyses of Torrance's (1958 to Present) longitudinal data. *Creat. Res. J.* 12, 103–114. doi: 10.1207/s15326934crj1202_3
- Runco, M. A., and Jaeger, G. J. (2012). The standard definition of creativity. *Creat. Res. J.* 24, 92–96. doi: 10.1080/10400419.2012.650092
- Seghier, M. L. (2013). The angular gyrus: multiple functions and multiple subdivisions. *Neuroscientist* 19, 43–61. doi: 10.1177/1073858412440596
- Shi, L., Sun, J., Xia, Y., Ren, Z., Chen, Q., Wei, D., et al. (2018). Large-scale brain network connectivity underlying creativity in resting-state and task fMRI: cooperation between default network and frontal-parietal network. *Biol. Psychol.* 135, 102–111. doi: 10.1016/j.biopsycho.2018.03.005

- Shomstein, S. (2012). Cognitive functions of the posterior parietal cortex: top-down and bottom-up attentional control. *Front. Integr. Neurosci.* 6:38. doi: 10.3389/fnint.2012.00038
- Sowden, P. T., Pringle, A., and Gabora, L. (2015). The shifting sands of creative thinking: Connections to dual-process theory. *Think. Reason.* 21, 40–60. doi: 10.1080/13546783.2014.885464
- Sten, S., Lundengård, K., Witt, S. T., Cedersund, G., Elinder, F., and Engström, M. (2017). Neural inhibition can explain negative BOLD responses: a mechanistic modelling and fMRI study. *NeuroImage* 158, 219–231. doi: 10.1016/j.neuroimage.2017.07.002
- Sternberg, R. J. (1999). “Handbook of creativity,” in *Psychology, Cognition, Educational Psychology*. ed R. J. Sternberg (Cambridge: Cambridge University Press).
- Sternberg, R. J., and Lubart, T. I. (1996). Investing in creativity. *Am. Psychol.* 51, 677–688. doi: 10.1037/0003-066X.51.7.677
- Stevens, W. D., Buckner, R. L., and Schacter, D. L. (2010). Correlated low-frequency BOLD fluctuations in the resting human brain are modulated by recent experience in category-preferential visual regions. *Cereb. Cortex* 20, 1997–2006. doi: 10.1093/cercor/bhp270
- Takeuchi, H., Taki, Y., Hashizume, H., Sassa, Y., Nagase, T., Nouchi, R., et al. (2012). The association between resting functional connectivity and creativity. *Cereb. Cortex* 22, 2921–2929. doi: 10.1093/cercor/bhr371
- Thompson-Schill, S. L. (2003). Neuroimaging studies of semantic memory: Inferring “how” from “where.” *Neuropsychologia* 41, 280–292. doi: 10.1016/S0028-3932(02)00161-6
- Turner, B. O., Paul, E. J., Miller, M. B., and Barbey, A. K. (2018). Small sample sizes reduce the replicability of task-based fMRI studies. *Commun. Biol.* 1, 62–10. doi: 10.1038/s42003-018-0073-z
- Van Dijk, K. R. A., Hedden, T., Venkataraman, A., Evans, K. C., Lazar, S. W., and Buckner, R. L. (2010). Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. *J. Neurophysiol.* 103, 297–321. doi: 10.1152/jn.00783.2009
- Wang, Z., Liu, J., Zhong, N., Qin, Y., Zhou, H., and Li, K. (2012). Changes in the brain intrinsic organization in both on-task state and post-task resting state. *NeuroImage* 62, 394–407. doi: 10.1016/j.neuroimage.2012.04.051
- Ward, T. B. (1994). Structured imagination: the role of category structure in exemplar generation. *Cogn. Psychol.* 27, 1–40. doi: 10.1006/cogp.1994.1010
- Wei, D., Yang, J., Li, W., Wang, K., Zhang, Q., and Qiu, J. (2014). Increased resting functional connectivity of the medial prefrontal cortex in creativity by means of cognitive stimulation. *Cortex* 51, 92–102. doi: 10.1016/j.cortex.2013.09.004
- Yan, F.-X., Wu, C. W., Cheng, S.-Y., Lim, K.-E., Hsu, Y.-Y., and Liu, H.-L. (2013). Resting-state functional magnetic resonance imaging analysis with seed definition constrained by regional homogeneity. *Brain Connect.* 3, 438–449. doi: 10.1089/brain.2013.0164

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

Copyright © 2020 Wu, Kuo, Huang, Tsai, Hsu, Hsu, Liu, Chen and Wu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Characterization of the Stages of Creative Writing With Mobile EEG Using Generalized Partial Directed Coherence

Jesus G. Cruz-Garza^{1*}, Akshay Sujatha Ravindran¹, Anastasiya E. Kopteva¹, Cristina Rivera Garza^{1,2} and Jose L. Contreras-Vidal¹

¹ Laboratory for Non-Invasive Brain-Machine Interface Systems, NSF IUCRC BRAIN, University of Houston, Houston, TX, United States, ² Department of Hispanic Studies, University of Houston, Houston, TX, United States

OPEN ACCESS

Edited by:

Vasil Kolev,
Bulgarian Academy of Sciences
(BAS), Bulgaria

Reviewed by:

Danilo Mandic,
Imperial College London,
United Kingdom
Luiz Antonio Baccala,
University of São Paulo, Brazil

*Correspondence:

Jesus G. Cruz-Garza
jgcruzgarza@gmail.com

Specialty section:

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

Received: 29 June 2020

Accepted: 10 November 2020

Published: 07 December 2020

Citation:

Cruz-Garza JG, Sujatha Ravindran A,
Kopteva AE, Rivera Garza C and
Contreras-Vidal JL (2020)
Characterization of the Stages of
Creative Writing With Mobile EEG
Using Generalized Partial Directed
Coherence.
Front. Hum. Neurosci. 14:577651.
doi: 10.3389/fnhum.2020.577651

Two stages of the creative writing process were characterized through mobile scalp electroencephalography (EEG) in a 16-week creative writing workshop. Portable dry EEG systems (four channels: TP09, AF07, AF08, TP10) with synchronized head acceleration, video recordings, and journal entries, recorded mobile brain-body activity of Spanish heritage students. Each student's brain-body activity was recorded as they experienced spaces in Houston, Texas ("Preparation" stage), and while they worked on their creative texts ("Generation" stage). We used Generalized Partial Directed Coherence (gPDC) to compare the functional connectivity among both stages. There was a trend of higher gPDC in the Preparation stage from right temporo-parietal (TP10) to left anterior-frontal (AF07) brain scalp areas within 1–50 Hz, not reaching statistical significance. The opposite directionality was found for the Generation stage, with statistical significant differences ($p < 0.05$) restricted to the delta band (1–4 Hz). There was statistically higher gPDC observed for the inter-hemispheric connections AF07–AF08 in the delta and theta bands (1–8 Hz), and AF08 to TP09 in the alpha and beta (8–30 Hz) bands. The left anterior-frontal (AF07) recordings showed higher power localized to the gamma band (32–50 Hz) for the Generation stage. An ancillary analysis of Sample Entropy did not show significant difference. The information transfer from anterior-frontal to temporal-parietal areas of the scalp may reflect multisensory interpretation during the Preparation stage, while brain signals originating at temporal-parietal toward frontal locations during the Generation stage may reflect the final decision making process to translate the multisensory experience into a creative text.

Keywords: creative writing, creativity, EEG, MoBI, generalized partial directed coherence

1. INTRODUCTION

Creative writing involves embodied practices that physically and emotionally connect us with our surroundings (Rivera Garza, 2013). We investigated creative writing as a bodily experience, in which the author's interaction with the world around them (physically, verbally, perceptually, and emotionally) informs the preparation and elaboration of their written work. In this way, as authors actively seek and engage in experiences in the world around them through their body and mind, these experiences affect the aesthetic and semantic components of their creative output.

Previous studies have used scalp electroencephalography (EEG) to analyze neural correlates of writing. For example, in a study using a custom test to evaluate reading and writing achievement to assess educational grade requirements, Harmony et al. (1990) found that high correlation in the delta and theta bands in frontal and temporal electrodes was related to poor writing performance, while high correlation in the alpha band in occipital areas is related to high writing performance in a group of 81 children. The same research group (Marosi et al., 1995) reported high coherence in delta (1–4 Hz), theta (4–8 Hz), and beta (12–24 Hz) bands associated with poor performance in reading and writing in 84 children. Conversely, high coherence in the alpha (8–12 Hz) band was related to proficient reading and writing. Coherence metrics between pairs of EEG electrodes in different scalp areas have been shown to have a positive correlation with an individual's creativity level, in short creative verbal and visual tasks (Petche et al., 1997), and in the Torrance Test of creative thinking (Jaušovec and Jaušovec, 2000a). These findings suggest that coherence has a more intense relationship with creative performance than EEG frequency-band-power metrics.

Neuroimaging studies with functional magnetic resonance imaging (fMRI) have been used to measure functional connectivity (FC) of participants at rest, the resting state FC (rFC). Lotze et al. (2014) found decreased rFC between inter-hemispheric areas BA 44, and left area BA 44 with the left temporal lobe for individuals who scored higher in a verbal creativity index test. These brain areas are part of Broca's area, a region located in the frontal part of the left hemisphere of the brain that is active in semantic tasks, such as semantic decision tasks (determination of whether a word represents an abstract or a concrete entity) and generation tasks (generation of a verb associated with a noun).

In studies where participants generated original text compositions, fMRI studies have also analyzed the human creative process through its distinct stages of preparation, generation, and revision. Shah et al. (2013) studied the Preparation and Generation stages, by the implementation of an actual creative writing paradigm. They found distinct cortical networks associated with each stage in a fronto-parieto-temporal network. The Preparation stage, "brainstorming," was found to activate the premotor cortex (involved in the cognitive-motor preparation to write), language processing areas in the bilateral IFG and left temporal areas, and left lateral orbito-frontal regions for higher order cognitive processing. Creative writing activated areas associated with handwriting (primary motor cortex and somatosensory areas), and cognitive processing such as episodic memory retrieval and semantic integration in bilateral hippocampi and temporal poles, with right-lateralized activation in posterior and anterior lobes. The same research group found that expert writers had higher activation in medial prefrontal cortex (mPFC) and basal ganglia areas (Erhard et al., 2014). Liu et al. (2015) studied the generation and revision stages and reported that the mPFC was active during both stages and the responses in dorsolateral prefrontal cortex (DLPFC) and Intraparietal sulcus (IPS) were deactivated during the Generation stage.

Differences in brain activity for the distinct stages of the creative process however, remain mostly unexplored in the EEG domain; particularly for creative writing tasks. As fMRI studies represent indirect evidence about cortical dynamics in cognitive-motor tasks such as creative writing, it is important to use time-resolved, direct methods that assess brain dynamics in action and in context in natural settings. In this regard, mobile EEG allows for the collection of brain activity data in more natural settings, where the users have freedom of motion and can freely walk around their surroundings (Cruz-Garza et al., 2019).

To better understand the brain dynamics of writers working 'in action and in context' in both preparatory and generative stages of creative writing, we integrated wearable MoBI technology into a creative writing course in Spanish at the University of Houston. The course was designed and led by Prof. Cristina Rivera Garza at the University of Houston. The experiment was designed together with the aims of the course to provide an equal consideration in the experimental design and evaluation process to best assess the creative process in an authentic creative writing experience.

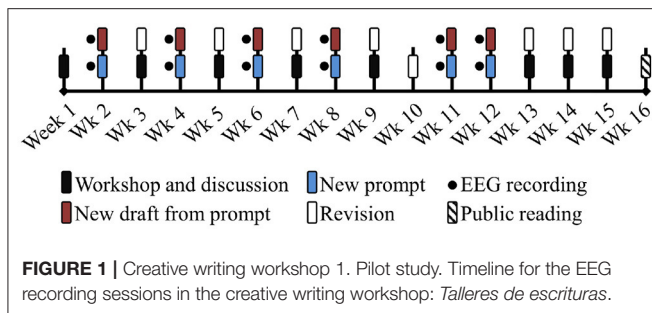
Specifically, we studied the process of creative writing with eighteen Spanish heritage speakers, as they engaged in the Preparation and Generation stages of their writing. The students were asked to walk through different areas of the City of Houston to experience a variety of pre-selected contextual environmental settings (we called them "prompts") chosen by the instructor, and use the experience to provide aesthetic and semantic content in their narratives. This study aimed to use scalp electroencephalography (EEG) to identify brain features or neural markers related to the different stages of creative writing where free behaving participants were able to move, explore their surroundings to inform their creative texts (Preparation stage), and write at their own time (Generation stage).

This study investigates the neural features associated with creative writing using quantitative and mobile EEG, through the characterization of the Preparation and Generation of creative texts in students that participated in a creative writing workshop. The students developed their writing skills throughout the workshop, and physically interacted with space and their communities during the Preparation stage. The Generation stage consisted of creating a first draft for each of their assignments.

2. METHODS

2.1. Human Participants

Eighteen heritage Spanish speaking undergraduate students participated in a Spanish language creative writing upper-division undergraduate workshop (SPAN 3308 YOUR BRAIN ON WRITING: Writing, Body, and Neuroaesthetics) at the University of Houston. The participants provided Anonymous Informed Consent, approved by the University of Houston Institutional Review Board, at the beginning of the workshop. The students received training to set up their own EEG headsets and body-mounted video cameras for the experiment. The students were responsible for the collection of EEG data, video, and to keep a diary with notes on each recording session.



2.2. Experimental Task

Through readings and writing prompts, participants were asked to experience and acknowledge the physicality of the writing process and to relate it to the materiality of language. Prompts, designed by the instructor, encouraged students to develop and record a series of specific writing preparation tasks that emphasized the physicality of the participants' bodily experience (e.g., walking, running, climbing in different locations of Houston) as part of the required assignments for the workshop. The writing prompts are provided in the **Supplementary Materials**.

The participants were then asked to utilize these prompted experiences to generate their creative texts in a 3–5 page suggested draft length (double space, 11 point font). The students were instructed to use the EEG and video cameras during both their prompted physical activities and writing time. There could be more than one session of walking and writing EEG recording sessions per prompt.

A timeline schematic of the creative writing workshop is represented in **Figure 1**. The workshop was 16 weeks long, with six creative writing exercises distributed in weeks 2, 4, 6, 8, 11, and 12. In the weeks in between the creative writing sessions, the students would discuss their peers' texts and improve their own previous drafts. At the end of the workshop, there was a public reading of the finalized creative texts.

This experimental setup produced data in two stages of the creative process: the Preparation stage and the Generation stage. The Preparation stage involved tasks such as walking, active observation of their environment, taking notes, and ideation. For the Generation stage, the task involved reviewing their notes and creating a creative narrative, with iterative revisions and modifications.

2.3. Equipment

EEG and head acceleration data were captured using Muse headsets (Interaxon, Toronto, Ontario, Canada). The headset has seven sensors, two out of these seven sensors are positioned at the frontal region (AF07 and AF08), two at temporal-parietal region (TP09 and TP10), and the remaining three sensors served as electrical reference located at the center of the forehead (Fpz). The 61 g headset has a built-in accelerometer that was used to measure the head acceleration. EEG data for each channel are measured in microvolts (μV) with a sampling rate of 220 Hz at 10-bit resolution.

The acceleration data was recorded at 50 Hz. Additionally, the data recordings contain a vector indicating contact quality (sampled at 10 Hz) for each electrode, rating contact quality as “indicator = 1: good,” “indicator = 2: acceptable,” “indicator \geq 3: bad.” (see <http://developer.choosemuse.com/hardware-firmware/hardware-specifications> for full technical specifications). Previous experiments have shown the capacity of the Muse headsets for the collection of mobile EEG data outside a laboratory setting (Ravindran et al., 2019).

The participants set up their own headset with a custom application given to them in a personal tablet, which recorded EEG and head acceleration data and labeled the participant identification number and date/time for the recording session automatically. The data recording setup is illustrated in **Figure 2**. Additionally, the participants set up body-cameras (Conbrov, ShenZhen, China) to record their exploration (Preparation) and writing (Generation) sessions. The camera recorded 720 HD video on a 75° wide-angle lens.

2.4. Data Collection

The students were asked to make five writing exercises and collect their brain activity as they walked and observed their environment (Preparation stage), and created their texts (Generation stage). Only writing assignments that were submitted and accompanied by both video and EEG data were considered for the analysis. From the 18 participants, data from eleven students was discarded due to incomplete data (video or EEG missing) or assignments not submitted on time. This represents a yield of 39% of the total data collected for further analysis.

The Preparation and Generation stages for each writing exercise were done in several distinct recording sessions as each stage could take several recording sessions to complete during the semester. We kept each data recording as a separate session to analyze. Recording sessions were considered for analysis when all four electrodes had a “good” contact indicator for at least one continuous minute of data.

2.5. Pre-processing

Data recordings with both video (providing contextual cues) and EEG were considered for this analysis. An online notch filter was applied on the EEG data to remove the 60 Hz power line noise (available as a user preset for the headband). We applied an offline 4th order, zero-stage Butterworth band-pass filter from 1 to 100 Hz. Artifact Subspace Reconstruction (ASR) (Mullen et al., 2013) was used for the removal of short-time high-amplitude artifacts in the continuous data as in Ravindran et al. (2019). Calibration data for ASR for each student was computed from the entire length of the trial using automated methods. A cut off threshold of ten standard deviations was used for the identification of corrupted subspaces, and a window length of 500 ms with a step size of 250 ms was used for the ASR. Among the segments, channels having corrupt PC loading to be >0.75 were removed. The remaining segments were then inspected automatically to remove data from any electrode disconnections from the scalp (tracked by the headband status data), any abrupt change of voltage $>100 \mu V$, or EEG data collected while there

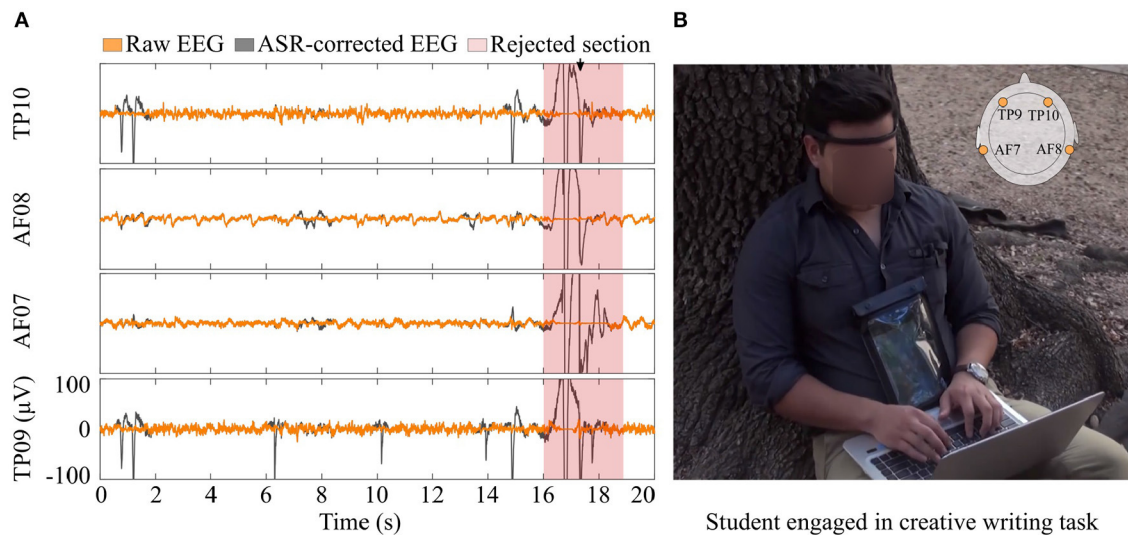


FIGURE 2 | Equipment setup and EEG data pre-processing. **(A)** Raw (black) and pre-processed (orange) EEG data. Shaded areas indicate rejected intervals. **(B)** A student wearing the EEG headset during the Generation stage.

was an absolute acceleration magnitude larger than 1 ms^{-2} (Ravindran et al., 2019). After removing noisy segments of data with our pre-processing methods, 77% of the data was kept for analysis.

2.6. Feature Extraction

2.6.1. Functional Connectivity, FC

Functional Connectivity is defined as the statistical association among two or more anatomically distinct time-series and can be assessed with EEG coherence measures or fMRI (Friston et al., 1993). FC analysis was performed upon the EEG channels by computing the generalized partial directed coherence (gPDC) measure (Baccala et al., 2007) over 6 s time segments with 50% overlap, a short-time based stationary approach (Omidvarnia et al., 2011). Partial coherence measures have been found to perform well with low-density EEG (Barzegaran and Knyazeva, 2017), making it a useful tool for our EEG dataset of four channels. PDC is a frequency-domain metric that provides information about directionality in the interaction between two signals among a larger number of signals (Baccalá and Sameshima, 2001), but it is dependent on the scale of the individual inputs (Blinowska, 2011). To account for this, Baccala et al. (2007) introduced a normalization term based on the variance of the signals and denoted this modified measure as gPDC, which was used in this study with the FieldTrip implementation [ft_connectivityanalysis()] (Oostenveld, 2020); with the BSMART toolbox (Cui et al., 2008). We chose a 6 s epoch due to observed stability at epoch lengths of 6 s or more for FC measures (Fraschini et al., 2016). gPDC was estimated using a multivariate autoregressive model (MVAR) using all four electrodes. We used an MVAR model order of 12 (54 ms), which was obtained by using the ARFIT algorithm (Schneider and Neumaier, 2001) and evaluating the SBC criterion, which is least affected by noise (Porcaro et al., 2009). The observed gPDC

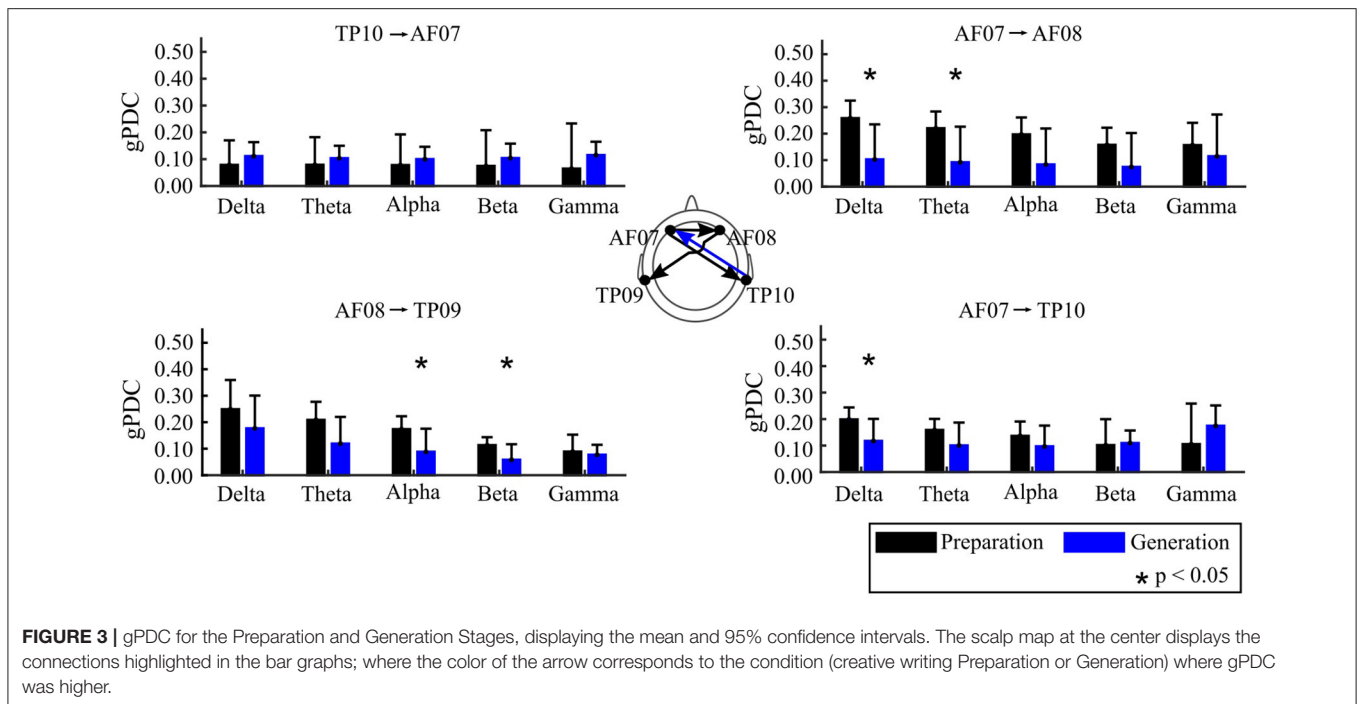
estimates were plotted for all pairs of electrodes in the frequency bands: delta [1–4 Hz], theta [4–8 Hz], alpha [8–12 Hz], beta [12–30 Hz], gamma [30–50 Hz].

2.6.2. Power Spectral Density, PSD

The power spectral density (PSD) was computed for each data window using Thomson's multitaper PSD estimate, with 4,096 frequency bins [1–50 Hz] and half-bandwidth product $nw = 4$. The mean PSD was obtained for each recorded session. Those recorded sessions corresponding to the Preparation and the Generation stage were compared for each of the four electrodes separately.

2.6.3. Sample Entropy, SampEn

Complexity measures has been used in different studies to measure levels of creativity before (Jaušovec and Jaušovec, 2000b; Shourie et al., 2014). Approximate Entropy is a measure of signal regularity which that explores the time ordering of data points by calculating the log likelihood that runs of pattern which are close remain close for incremental comparison (Pincus, 1995). Lower value of Approximate Entropy indicates that the signal is more regular or predictable. However, many studies have reported reliability issues using Approximate Entropy due to the self-match involved in Approximate Entropy computation leading to a bias (Richman and Moorman, 2000; Chen et al., 2006). A new metric called Sample Entropy (SampEn) (Richman and Moorman, 2000) was proposed aimed at reducing the bias of Approximate Entropy. The parameters remained the same for both Approximate Entropy and SampEn: the “filter factor” r , length of sequences being compared m and the signal length N . SampEn has shown to be relatively less dependent on the signal length and shows better stability for wider range of parameters (Richman and Moorman, 2000; Chen et al., 2006; Boskovic et al., 2012). Earlier studies showed that SampEn gives better statistical



validity for $m = 2$ and the r in the range of 0.1–0.25 standard deviations (Richman and Moorman, 2000; Bruce et al., 2009; Zarjam et al., 2012). In this study, we used $m = 2$ and $r = 0.2$ standard deviations of the signal window, and $N = 1,320$ (6 s of data sampled at 220 Hz). As with the PSD analysis, the SampEn means were computed for each recorded data collection session, and those corresponding to the Preparation and Generation Stages were compared for the four electrodes separately.

2.6.4. Statistical Analyses

Statistical significance comparing the Preparation and Generation Stages was obtained using a two-tailed unpaired T -test with a significance level of $p < 0.05$. All comparisons show the mean, confidence intervals at the $p < 0.05$ significance level, and an indication for when the difference was statistically significant at that statistical level. The Fisher-Snedecor F -test was performed to assess if the variances were equal. When the variances were different (Cardillo, 2020), the Satterthwaite approximate T -test was performed (Satterthwaite, 1946).

3. RESULTS

The main finding of this study is that the Preparation and Generation stages of creative writing were characterized differentially in terms of the functional connectivity among the scalp locations examined. Specifically, the gPDC between the Preparation and Generation stages showed the opposite directionality between right temporal and left anterior frontal areas. **Figure 3** shows significant differences in FC between electrode pairs, using gPDC, during the two stages of the creative writing process analyzed.

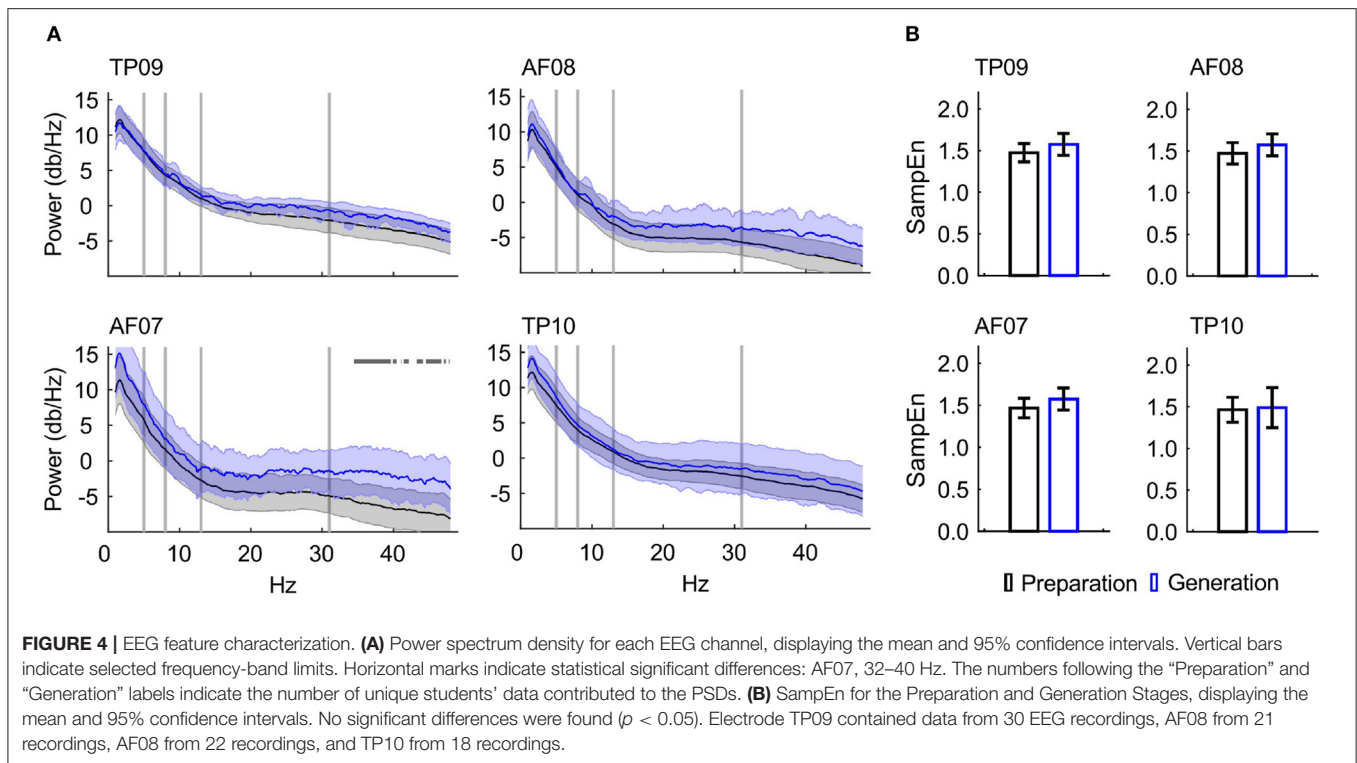
Preparation stage: There was higher gPDC in the Preparation stage originating from anterior-frontal electrodes toward temporal-parietal electrodes. The gPDC difference between the Preparation and the Generation stages showed statistically broadband significance, at a confidence level of $p < 0.05$, for three connections. AF07 to TP10 showed significant differences in the delta (1–4 Hz) frequency range. There was statistically higher gPDC observed for the connections AF07 to AF08 in the delta and theta (1–8 Hz) frequency bands, and AF08 to TP09 from alpha (8–12 Hz) to beta (12–30 Hz). **Figure 3** shows the gPDC values, bounded between 0 and 1, for those connections where statistical differences were found between conditions.

Generation stage: There was higher Partial Directed Coherence in the Generation stage originating from TP10 toward AF07. Although there was no significant differences in the TP10 to AF07 comparison (**Figure 3**) it was the only connection that showed higher gPDC for the Generation stage in a clear trend across frequency bands.

The statistical difference in gPDC and its opposite directionality when comparing the Preparation and the Generation Stages indicates that there was a strong functional relation between the left anterior frontal with the right temporal-parietal areas when the students engaged in the tasks.

Frequency band-power analysis showed a statistically significant difference across writing stages within the gamma (32–40 Hz) frequency range for the AF07 electrode only (**Figure 4A**).

The SampEn was higher during Generation stage compared to Preparation stage, although this difference did not reach statistical significance (**Figure 4B**).



4. DISCUSSION

The higher functional connectivity from anterior frontal scalp areas toward temporal-parietal areas during the Preparation stage suggests a functional relationship between areas involved in the processing of multisensory inputs (Perrodin et al., 2014; Schapiro et al., 2014) and episodic emotional memory integration (Dolcos et al., 2004, 2005; Lech and Suchan, 2013) in the temporal lobe as participants explore their surroundings actively engaging the frontal cortex in integrating the experience. The opposite directionality between the same electrodes (Figure 3) at the Generation stage reinforces this hypothesis in which sensory input are reprocessed in the frontal areas to produce a draft of a creative text based on those experiences.

Our results, although constrained to frontal and temporal recording locations, relate to previous findings in EEG and fMRI studies analyzing the different stages of the creative process. It furthers the suggestions from Petche et al. (1997) and Jaušovec and Jaušovec (2000a) that EEG correlates of creative performance are more pronounced in functional connections between brain areas than localized frequency-band power. In the last decade, distinct cortical networks have been associated with each stage of the human creative writing process. Shah et al. (2013) identified ventrolateral prefrontal cortex activation during the Preparation stage, and central-parietal areas involved in the Generation stage. “Brainstorming” engaged cognitive, linguistic, and creative functions represented in a parieto-frontal-temporal network, while “Creative writing” activated motor, visual, a cognitive and linguistic areas mainly over central and parietal networks (Shah et al., 2013). Liu et al. (2015) found that the mPFC was

active during the generation and revision stages. They observed deactivation of the dorsolateral prefrontal cortex (DLPFC) and inferior parietal sulcus (IPS) during the Generation stage.

Our results show a trend of higher gPDC values from the right temporal toward the left anterior frontal electrode during the Generation stage of creative writing, for all frequency bands analyzed (1–50 Hz), albeit without reaching statistical significance; and the opposite directionality for the Preparation stage with statistical significance at 1–4 Hz. The Preparation stage also showed higher connectivity compared to the Generation stage with connections originating from anterior-frontal electrodes: AF07 to AF08 (significance at 1–8 Hz), and AF08 to TP09 (significance at 8–30 Hz). We did not find statistical differences between the Preparation and the Generation Stages for Sample Entropy; and frequency band-power showed differences only in the left anterior frontal electrode in the gamma band. In future studies, Entropy analysis could benefit from multi-scale (Gao et al., 2015) and multivariate (Ahmed and Mandic, 2011) entropy models to account for the varying and complex nature of physiological signals (Costa et al., 2002).

Overall, these findings suggest that ideation, exploration, and observation during the Preparation stage of a creative writing task can be characterized by a state of long-range cortico-cortical communication between multisensory integration brain areas (temporal regions) and high-order execution and planning areas of the brain (prefrontal regions), perhaps leading to selective storage of ideas, concepts or observations candidate for creating writing during the Generation stage. We hypothesize this focal activity may be related to working memory, sequence

production, and processing of filtered information from the Preparation stage.

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 University of Houston's Institutional Review Board (IRB). The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JC-G performed the data analysis and wrote the manuscript. AS prepared and pre-processed the data, preformed preliminary analysis, and contributed with data-processing code. AK assisted students on a weekly basis on data collection, and compiled

the multimodal data in a working dataset. JC-G, AK, and CR planned the experiment. CR conducted the workshop. JC-V and CR conceived the research and edited the manuscript. All authors reviewed the manuscript.

FUNDING

This research was funded in part by NSF award #BCS1533691, NSF IUCRC BRAIN Award CNS1650536, the Cullen College of Engineering at the University of Houston, and the SeFAC grant from the Hewlett Packard Enterprise Data Science Institute at the University of Houston.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnhum.2020.577651/full#supplementary-material>

The gPDC distributions for the data from all participants and creative writing exercises used to generate **Figure 3** can be found as **Supplementary Materials**.

REFERENCES

- Ahmed, M. U., and Mandic, D. P. (2011). Multivariate multiscale entropy: a tool for complexity analysis of multichannel data. *Phys. Rev. E* 84:061918. doi: 10.1103/PhysRevE.84.061918
- Baccalá, L. A., and Sameshima, K. (2001). Partial directed coherence: a new concept in neural structure determination. *Biol. Cybern.* 84, 463–474. doi: 10.1007/PL00007990
- Baccalá, L. A., Sameshima, K., and Takahashi, D. (2007). "Generalized partial directed coherence," in *2007 15th International Conference on Digital Signal Processing* (Cardiff), 163–166. doi: 10.1109/ICDSP.2007.4288544
- Barzegaran, E., and Knyazeva, M. G. (2017). Functional connectivity analysis in EEG source space: the choice of method. *PLoS ONE* 12:e0181105. doi: 10.1371/journal.pone.0181105
- Blinowska, K. J. (2011). Review of the methods of determination of directed connectivity from multichannel data. *Med. Biol. Eng. Comput.* 49, 521–529. doi: 10.1007/s11517-011-0739-x
- Boskovic, A., Loncar-Turukalo, T., Sarenac, O., Japundzic-Zigon, N., and Bajic, D. (2012). Unbiased entropy estimates in stress: a parameter study. *Comput. Biol. Med.* 42, 667–679. doi: 10.1016/j.combiomed.2012.03.003
- Bruce, E. N., Bruce, M. C., and Vennelaganti, S. (2009). Sample entropy tracks changes in EEG power spectrum with sleep state and aging. *J. Clin. Neurophysiol.* 26:257. doi: 10.1097/WNP.0b013e3181b2f1e3
- Cardillo, G. (2020). *Student t-Test for Unpaired or Paired Samples*. GitHub. Available online at: <https://github.com/dnafinder/testt> (accessed May 20, 2020).
- Chen, X., Solomon, I. C., and Chon, K. H. (2006). "Comparison of the use of approximate entropy and sample entropy: applications to neural respiratory signal," in *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference* (Shanghai), 4212–4215.
- Costa, M., Goldberger, A. L., and Peng, C.-K. (2002). Multiscale entropy analysis of complex physiologic time series. *Phys. Rev. Lett.* 89:068102. doi: 10.1103/PhysRevLett.89.068102
- Cruz-Garza, J. G., Kopteva, A. E., Fleischhauer, J. A., and Contreras-Vidal, J. L. (2019). "Into the mind of an artist: convergent research at the nexus of art, science, and technology," in *Mobile Brain-Body Imaging and the Neuroscience of Art, Innovation and Creativity*, eds J. L. Contreras-Vidal, D. Robleto, J. G. Cruz-Garza, J. M. Azorin, and C. Nam (Springer), 61–74. doi: 10.1007/978-3-030-24326-5_8
- Cui, J., Xu, L., Bressler, S. L., Ding, M., and Liang, H. (2008). BSMAT: a Matlab/C toolbox for analysis of multichannel neural time series. *Neural Netw.* 21, 1094–1104. doi: 10.1016/j.neunet.2008.05.007
- Dolcos, F., LaBar, K. S., and Cabeza, R. (2004). Interaction between the amygdala and the medial temporal lobe memory system predicts better memory for emotional events. *Neuron* 42, 855–863. doi: 10.1016/S0896-6273(04)00289-2
- Dolcos, F., LaBar, K. S., and Cabeza, R. (2005). Remembering one year later: role of the amygdala and the medial temporal lobe memory system in retrieving emotional memories. *Proc. Natl. Acad. Sci. U.S.A.* 102, 2626–2631. doi: 10.1073/pnas.0409848102
- Erhard, K., Kessler, F., Neumann, N., Ortheil, H.-J., and Lotze, M. (2014). Professional training in creative writing is associated with enhanced fronto-striatal activity in a literary text continuation task. *NeuroImage* 100, 15–23. doi: 10.1016/j.neuroimage.2014.05.076
- Fraschini, M., Demuru, M., Crobe, A., Marrosu, F., Stam, C. J., and Hillebrand, A. (2016). The effect of epoch length on estimated EEG functional connectivity and brain network organisation. *J. Neural Eng.* 13:036015. doi: 10.1088/1741-2560/13/3/036015
- Friston, K., Frith, C., Liddle, P., and Frackowiak, R. (1993). Functional connectivity: the principal-component analysis of large (PET) data sets. *J. Cereb. Blood Flow Metab.* 13, 5–14. doi: 10.1038/jcbfm.1993.4
- Gao, J., Hu, J., Liu, F., and Cao, Y. (2015). Multiscale entropy analysis of biological signals: a fundamental bi-scaling law. *Front. Comput. Neurosci.* 9:64. doi: 10.3389/fncom.2015.00064
- Harmony, T., Hinojosa, G., Marosi, E., Becker, J., Rodriguez, M., Reyes, A., et al. (1990). Correlation between EEG spectral parameters and an educational evaluation. *Int. J. Neurosci.* 54, 147–155. doi: 10.3109/00207459008986630
- Jaušovec, N., and Jaušovec, K. (2000a). Differences in resting EEG related to ability. *Brain Topogr.* 12, 229–240. doi: 10.1023/A:1023446024923
- Jaušovec, N., and Jaušovec, K. (2000b). EEG activity during the performance of complex mental problems. *Int. J. Psychophysiol.* 36, 73–88. doi: 10.1016/S0167-8760(99)00113-0
- Lech, R. K., and Suchan, B. (2013). The medial temporal lobe: memory and beyond. *Behav. Brain Res.* 254, 45–49. doi: 10.1016/j.bbr.2013.06.009
- Liu, S., Erkinen, M. G., Healey, M. L., Xu, Y., Swett, K. E., Chow, H. M., et al. (2015). Brain activity and connectivity during poetry composition: toward a multidimensional model of the creative process. *Hum. Brain Mapp.* 36, 3351–3372. doi: 10.1002/hbm.22849
- Lotze, M., Erhard, K., Neumann, N., Eickhoff, S. B., and Langner, R. (2014). Neural correlates of verbal creativity: differences in resting-state functional connectivity associated with expertise in creative writing. *Front. Hum. Neurosci.* 8:516. doi: 10.3389/fnhum.2014.00516

- Marosi, E., Harmony, T., Becker, J., Reyes, A., Bernal, J., Fernández, T., et al. (1995). Electroencephalographic coherences discriminate between children with different pedagogical evaluation. *Int. J. Psychophysiol.* 19, 23–32. doi: 10.1016/0167-8760(94)00059-N
- Mullen, T., Kothe, C., Chi, Y. M., Ojeda, A., Kerth, T., Makeig, S., et al. (2013). “Real-time modeling and 3D visualization of source dynamics and connectivity using wearable EEG,” in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* (Osaka), 2184–2187. doi: 10.1109/EMBC.2013.6609968
- Omidvarnia, A., Mesbah, M., O’Toole, J. M., Colditz, P., and Boashash, B. (2011). “Analysis of the time-varying cortical neural connectivity in the newborn EEG: a time-frequency approach,” in *International Workshop on Systems, Signal Processing and Their Applications, WOSSPA* (Tipaza), 179–182. doi: 10.1109/WOSSPA.2011.5931445
- Oostenveld, R. (2020). *ft_connectivityanalysis, FieldTrip*. GitHub. Available online at: https://github.com/fieldtrip/fieldtrip/blob/master/ft_connectivityanalysis.m (accessed November 11, 2020).
- Perrodin, C., Kayser, C., Logothetis, N. K., and Petkov, C. I. (2014). Auditory and visual modulation of temporal lobe neurons in voice-sensitive and association cortices. *J. Neurosci.* 34, 2524–2537. doi: 10.1523/JNEUROSCI.2805-13.2014
- Petche, H., Kaplan, S., von Stein, A., and Filz, O. (1997). The possible meaning of the upper and lower alpha frequency ranges for cognitive and creative task. *Int. J. Psychophysiol.* 26, 77–97. doi: 10.1016/S0167-8760(97)00757-5
- Pincus, S. (1995). Approximate entropy (ApEn) as a complexity measure. *Chaos* 5, 110–117. doi: 10.1063/1.166092
- Porcaro, C., Zappasodi, F., Rossini, P. M., and Tecchio, F. (2009). Choice of multivariate autoregressive model order affecting real network functional connectivity estimate. *Clin. Neurophysiol.* 120, 436–448. doi: 10.1016/j.clinph.2008.11.011
- Ravindran, A. S., Mobiny, A., Cruz-Garza, J. G., Paek, A., Kopteva, A., and Vidal, J. L. C. (2019). Assaying neural activity of children during video game play in public spaces: a deep learning approach. *J. Neural Eng.* 16:036028. doi: 10.1088/1741-2552/ab1876
- Richman, J. S., and Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart Circul. Physiol.* 278, H2039–H2049. doi: 10.1152/ajpheart.2000.278.6.H2039
- Rivera Garza, C. (2013). *Los Muertos Indóviles: Necroescrituras y Desapropiación*. Grupo Planeta Spain.
- Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometr. Bull.* 2, 110–114. doi: 10.2307/3002019
- Schapiro, A. C., Gregory, E., Landau, B., McCloskey, M., and Turk-Browne, N. B. (2014). The necessity of the medial temporal lobe for statistical learning. *J. Cogn. Neurosci.* 26, 1736–1747. doi: 10.1162/jocn_a_00578
- Schneider, T., and Neumaier, A. (2001). Algorithm 808: Arfit—a matlab package for the estimation of parameters and eigenmodes of multivariate autoregressive models. *ACM Trans. Math. Softw.* 27, 58–65. doi: 10.1145/382043.382316
- Shah, C., Erhard, K., Ortheil, H.-J., Kaza, E., Kessler, C., and Lotze, M. (2013). Neural correlates of creative writing: an fMRI study. *Hum. Brain Mapp.* 34, 1088–1101. doi: 10.1002/hbm.21493
- Shourie, N., Firoozabadi, M., and Badie, K. (2014). Analysis of EEG signals related to artists and nonartists during visual perception, mental imagery, and rest using approximate entropy. *BioMed Res. Int.* 2014:764382. doi: 10.1155/2014/764382
- Zarjam, P., Epps, J., and Lovell, N. H. (2012). “Characterizing mental load in an arithmetic task using entropy-based features,” in *2012 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA)* (Montreal, QC), 199–204. doi: 10.1109/ISSPA.2012.6310545

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

Copyright © 2020 Cruz-Garza, Sujatha Ravindran, Kopteva, Rivera Garza and Contreras-Vidal. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Feasibility and Safety of Bilateral Hybrid EEG/EOG Brain/Neural–Machine Interaction

Marius Nann^{1,2}, Niels Peekhaus^{1,2}, Cornelius Angerhöfer^{1,2} and Surjo R. Soekadar^{1,2*}

¹ Clinical Neurotechnology Lab, Charité – University Medicine Berlin, Berlin, Germany, ² Applied Neurotechnology Lab, University Hospital Tübingen, Tübingen, Germany

OPEN ACCESS

Edited by:

Cuntai Guan,
Nanyang Technological University,
Singapore

Reviewed by:

Ashraf S. Gorgey,
Hunter Holmes McGuire VA Medical
Center, United States
Jianjun Meng,
Shanghai Jiao Tong University, China

*Correspondence:

Surjo R. Soekadar
surjo.soekadar@charite.de

Specialty section:

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

Received: 06 July 2020

Accepted: 09 November 2020

Published: 09 December 2020

Citation:

Nann M, Peekhaus N,
Angerhöfer C and Soekadar SR
(2020) Feasibility and Safety
of Bilateral Hybrid EEG/EOG
Brain/Neural–Machine Interaction.
Front. Hum. Neurosci. 14:580105.
doi: 10.3389/fnhum.2020.580105

Cervical spinal cord injuries (SCIs) often lead to loss of motor function in both hands and legs, limiting autonomy and quality of life. While it was shown that unilateral hand function can be restored after SCI using a hybrid electroencephalography/electrooculography (EEG/EOG) brain/neural hand exoskeleton (B/NHE), it remained unclear whether such hybrid paradigm also could be used for operating two hand exoskeletons, e.g., in the context of bimanual tasks such as eating with fork and knife. To test whether EEG/EOG signals allow for fluent and reliable as well as safe and user-friendly bilateral B/NHE control, eight healthy participants (six females, mean age 24.1 ± 3.2 years) as well as four chronic tetraplegics (four males, mean age 51.8 ± 15.2 years) performed a complex sequence of EEG-controlled bilateral grasping and EOG-controlled releasing motions of two exoskeletons visually presented on a screen. A novel EOG command performed by prolonged horizontal eye movements (>1 s) to the left or right was introduced as a reliable switch to activate either the left or right exoskeleton. Fluent EEG control was defined as average “time to initialize” (TTI) grasping motions below 3 s. Reliable EEG control was assumed when classification accuracy exceeded 80%. Safety was defined as “time to stop” (TTS) all unintended grasping motions within 2 s. After the experiment, tetraplegics were asked to rate the user-friendliness of bilateral B/NHE control using Likert scales. Average TTI and accuracy of EEG-controlled operations ranged at 2.14 ± 0.66 s and $85.89 \pm 15.81\%$ across healthy participants and at 1.90 ± 0.97 s and $81.25 \pm 16.99\%$ across tetraplegics. Except for one tetraplegic, all participants met the safety requirements. With $88 \pm 11\%$ of the maximum achievable score, tetraplegics rated the control paradigm as user-friendly and reliable. These results suggest that hybrid EEG/EOG B/NHE control of two assistive devices is feasible and safe, paving the way to test this paradigm in larger clinical trials performing bimanual tasks in everyday life environments.

Keywords: bilateral exoskeleton control, bimanual tasks, EEG, EOG, brain-computer interface, BCI, brain-machine (computer) interface

INTRODUCTION

Cervical spinal cord injuries (SCIs) often result in loss of motor function in all four extremities. According to the National Spinal Cord Injury Statistical Center (NSCISC), 41.1% of all SCIs lead to complete or incomplete tetraplegia (National Spinal Cord Injury Statistical Center, 2019). While the inability to walk is usually sufficiently compensated by use of a wheelchair (Rushton et al., 2010),

restoration of hand and arm function is still insufficiently solved. Therefore, restoration of hand and arm function is of highest priority in this patient population (Anderson, 2004; Snoek et al., 2004; Lo et al., 2016). Depending on the SCI's location, the degree of impairment and related motor inabilities can vary substantially. In particular, injuries between the spinal motion sections C5 and C7 are characterized by some remaining motor function in the shoulder and arm but absence of movements in the wrist and fingers (Ahuja et al., 2017). For these cases, restoration of hand function would be an important goal to regain autonomy and to improve quality of life (Campbell et al., 1999).

To date, the most common methods for restoration of upper limb motor function are surgical interventions (Bunketorp-Käll et al., 2017). To a certain degree, upper limb reconstructive surgeries, such as tendon transfers or tenodesis (Bednar and Woodside, 2018), can restore arm and hand function in SCI. However, besides the risks associated with surgery, tendon transfer strongly depends on the availability and quality of tendons and muscles suitable for transfer. While tenodesis enables tetraplegics to passively grasp objects through extension of the wrist (termed tenodesis grasp), the resulting grasping force is often insufficient to perform basal activities of daily living (ADLs), e.g., lifting up a water bottle, zipping a jacket, or reliably holding cutlery for eating (Dunn et al., 2016).

As an alternative to surgical interventions, recent advancements in neurotechnology and robotics opened up new possibilities to restore hand and arm function after cervical SCI (Soekadar et al., 2016) or stroke (Soekadar et al., 2008, 2015a; Nann et al., 2020). It was shown that exoskeletons or functional electrical stimulation (FES) of paralyzed muscles can enhance grasping force and improve hand function in tetraplegics (Ragnarsson, 2008; Ho et al., 2014; Yun et al., 2017; Cappello et al., 2018). A very intuitive way to control such assistive devices can be achieved by using a brain-computer interface (BCI; Wolpaw et al., 2002; Collinger et al., 2013a). BCIs translate electric, magnetic, or metabolic brain activity, e.g., associated with motor imagery (MI) or the attempt to move the paralyzed fingers, into control signals of digital devices, e.g., a robotic arm (Hochberg et al., 2012; Collinger et al., 2013b), exoskeleton (Soekadar et al., 2016; Tang et al., 2016; Frolov et al., 2017; Benabid et al., 2019), or FES device (Osuagwu et al., 2016; Vidaurre et al., 2016). Besides providing assistance, it was shown that repeated BCI use following SCI can also trigger neural recovery (Donati et al., 2016). Several studies showed that BCI-controlled FES can restore hand movement (Bouton et al., 2016; Vidaurre et al., 2016; Ajiboye et al., 2017). However, it is noteworthy that persons with SCI can develop upper extremity spasticity (Holtz et al., 2017; Gohritz and Fridén, 2018). In such cases, effective restoration of hand function *via* FES may not be successful due to increased muscle tone and tendon contractures. In contrast, a BCI-controlled hand exoskeleton, which actively opens and closes the affected hand, can overcome such limitations and may, thus, be superior to BCI-controlled FES. Within the last years, several robotic devices have entered the commercial market including three exoskeletons that were specifically designed for SCI patients (Mekki et al., 2018). Although still rather cost-intensive, new

3D-printed designs may yield low-cost hand exoskeletons in the near future (Yoo et al., 2019).

The most common approach for non-invasive brain/neural control of an exoskeleton uses modulation of sensorimotor rhythms (SMRs, 8–12 Hz) quantified as event-related desynchronization (ERD; SMR-ERD; Pfurtscheller and da Silva, 1999; Soekadar et al., 2011). SMR-ERD modulations related to MI or attempted finger movements are most prominent over the hand knob area of the contralateral primary motor cortex. Using electroencephalography (EEG), the optimal position to record SMR-ERD is typically at electrode positions C3 or C4 (according to the international 10/20 system; Neuper et al., 2006). Recently, it was demonstrated that a SMR-based brain/neural hand exoskeleton (B/NHE) can fully restore unilateral hand function in tetraplegics in an everyday life environment, e.g., to eat and drink in an outside restaurant (Soekadar et al., 2016). To deal with the inherent low signal-to-noise ratio of EEG recordings in everyday life environments, a hybrid EEG/electrooculography (EEG/EOG) brain/neural-machine interaction (B/NMI) system has been successfully introduced (Soekadar et al., 2015b, 2016; Crea et al., 2018; Nann et al., 2020). To enhance BCI control in everyday life environments, maximal horizontal oculoversion (HOVs) assessed by EOG were integrated as an additional control signal to reduce false classifications (Witkowski et al., 2014; Soekadar et al., 2015b). While exoskeleton closing motions were controlled by SMR-ERD related to intended grasping movements, HOVs were translated into opening motions or veto commands to interrupt unintended closing motions.

To date, the majority of studies in clinical settings have mainly focused on the restoration of *unilateral* motor function (Alam et al., 2016; Carvalho et al., 2019; Coscia et al., 2019). Most ADLs, however, involve *bilateral* motor function, e.g., eating with fork and knife, opening a water bottle, or a bag of potato chips. While, for example, a unilateral B/NHE might be sufficient to restore bimanual ADLs in hemiplegic stroke patients, patients suffering from tetraplegia depend on mobilization of both hands and arms to execute bimanual tasks. Therefore, a reliable and safe control paradigm allowing intuitive operation of bilateral hand exoskeletons would be very desirable.

The goal of such a bilateral control paradigm is to reliably detect the user's attempt to operate either the left or right exoskeleton, both exoskeletons simultaneously, or none of them. This results in a four-class classification problem. The simplest approach to deal with such a multiclass problem is to implement a single classifier that differentiates between left and right hemispheric SMR-ERD (Meng et al., 2016; León, 2017; Lotte et al., 2018). Although Meng et al. (2016) demonstrated that this kind of classification method is feasible in principle, it requires sufficient lateralization of SMR-ERD to C3 and C4. Given that chronic tetraplegics often do not show such lateralization (Osuagwu et al., 2016; Dahlberg et al., 2018), such approach may not be suitable for reliable exoskeleton control in SCI. A possible solution to overcome the lack of lateralization in SCI patients is to introduce a reliable switch to activate either the left or right exoskeleton.

Here, we introduce a novel EOG command performed by *prolonged* HOV (>1 s; **Figure 3**) to the left or right and

tested whether user use of such new command allows for reliable control of two hand exoskeletons. The prolonged HOV is not in conflict with the already established hybrid EEG/EOG paradigm according to Soekadar et al. (2016), where a *short* HOV (< 1 s; **Figure 3**) is used to veto an ongoing exoskeleton opening or closing. To test the feasibility and safety of such novel bilateral EEG/EOG-based B/NMI control, eight healthy participants as well as four chronic tetraplegics performed a neurofeedback paradigm consisting of a complex sequence of bilateral grasping and releasing motions of two exoskeletons visually presented on a screen. In the following work, feasibility was defined as fluency and accuracy of bilateral EEG/EOG B/NHE control. While fluent control was defined as “time to initialize” (TTI) EEG-controlled operations in average below 3 s (i.e., valid SMR-ERDs were detected in average within 3 s; Crea et al., 2018), reliable control was defined as average classification accuracy above 80%, following the recommendation of Vidaurre and Blankertz (2010) and Ortner et al. (2015), e.g., when benchmarking common spatial patterns (CSPs). Safety requirements were met when all unintended closing motions were interrupted by using short HOV before the exoskeleton was fully closed. This means the “time to stop” (TTS) all unintended closing motions ranged within 2 s, the time of a full exoskeleton closing motion. Moreover, user-friendliness of bilateral control was assessed among tetraplegics by using a Likert scale.

MATERIALS AND METHODS

Participants

Eight BCI-naïve healthy participants (six females, mean age 24.1 ± 3.2 years) and four BCI-naïve chronic tetraplegics (four males, mean age 51.8 ± 15.2 years, time since injury > 2 years) with complete ($n = 2$; American Spinal Injury Association (ASIA), grade A) and incomplete ($n = 2$, ASIA grades B and C) SCI (injury location between C5 and C7) were invited to a single-session experiment at the University Hospital of Tübingen, Germany. Before entering the study, all participants provided written informed consent. The study protocol complied with the Declaration of Helsinki and was approved by University of Tübingen's local ethics committee (registration code of ethical approval: 201/2018BO1).

Experimental Setup and Biosignal Online Processing

Electroencephalography was recorded from nine conventional recording sites (F3, T3, C3, P3, F4, T4, C4, P4, and Cz according to the international 10/20 system; **Figure 1**). Two additional EOG electrodes were placed laterally to the outer canthi of the left and right eye to assess HOVs (**Figures 1, 2**; Heide et al., 1999). A reference electrode was symmetrically placed over the sagittal midline at FCz to avoid biased electrical potentials toward one hemisphere (**Figure 1**). The ground electrode was located at Fpz (**Figure 1**). All biosignals were sampled at 1 kHz and amplified by a wireless active-electrode EEG system (actiCAP®, LiveAmp®, Brain Products GmbH, Gilching, Germany; **Figure 1**). To ensure high signal quality, all impedances were kept below 25 k Ω .

For online processing and classification, the BCI2000 software platform was used (Schalk et al., 2004). In order to attenuate eye blinks and other bihemispheric artifacts, bipolar EOG signal was calculated by subtracting left from right EOG. To remove low-frequency drifts as well as high-frequency noise, the bipolar EOG signal was then band-pass filtered with a first-order Butterworth filter at 0.02–3 Hz. To reduce the relatively long settling time that the low high-pass corner frequency at 0.02 Hz would have caused (> 50 s), the band-pass filter was initialized with the mean value of the first processed sample block of the bipolar EOG signal. Such filter initialization drastically reduced the settling time to be applicable in online settings. The very low frequency content in the EOG signal allows to extract the quasi-rectangular curve shapes resulting from HOVs and thus ensures reliable detection of prolonged HOVs (i.e., threshold was exceeded for > 1 s; **Figure 3**). EEG signals were first band-pass filtered with a first-order Butterworth filter at 1–30 Hz to remove baseline drifts and high-frequency noise. Afterward, surface Laplacian filters were applied to increase signal-to-noise ratio of the target electrodes at C3 and C4, respectively, (McFarland, 2015). A surface Laplacian filter was shown to be effective in detecting motor-specific SMR-ERD especially in online settings while suppressing distant sources (e.g., eye blinks) without the need for complex models, e.g., accounting for volume conduction. Subsequently, the power spectra of Laplace-filtered C3 and C4 EEG signals were estimated online from 500 ms moving

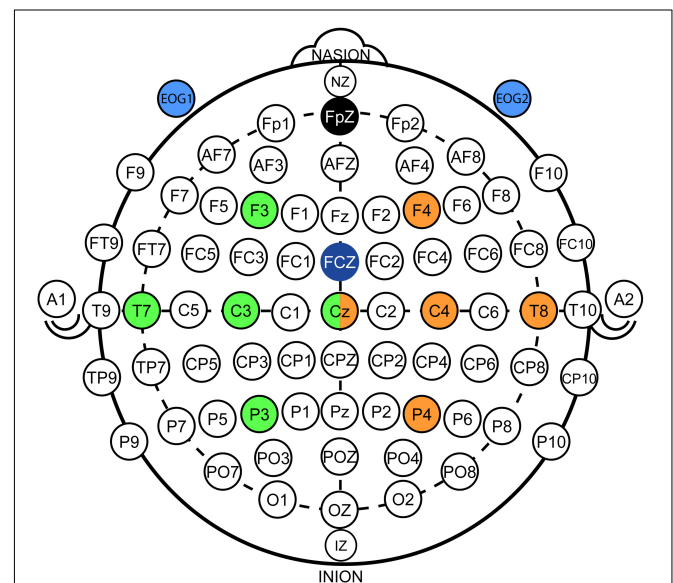
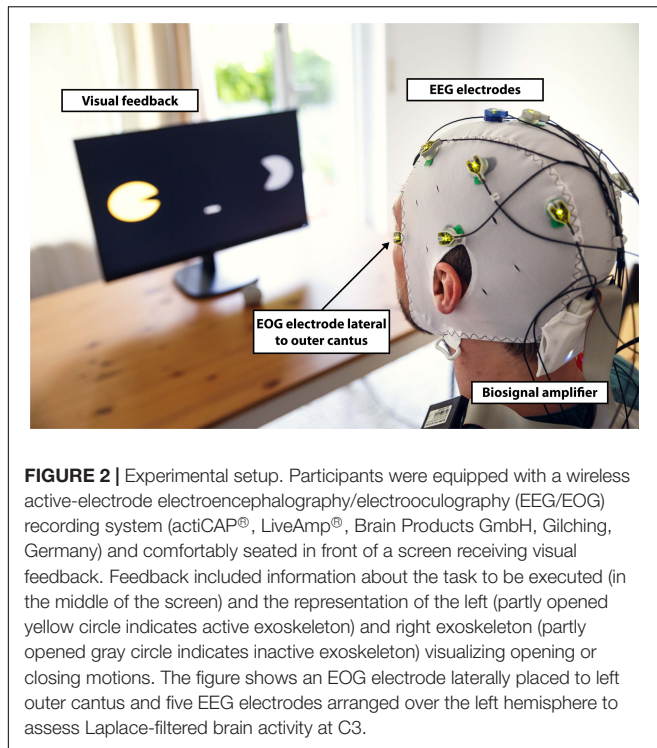


FIGURE 1 | Electroencephalography/electrooculography (EEG/EOG) electrode setup. EEG setup: Nine conventional EEG recording sites were used in accordance to the international 10/20 system. Five electrodes on each hemisphere were applied that were centered around C3 (green color coding) and C4 (orange color coding). Signals from Cz were used for both hemispheres. EOG setup: Two EOG electrodes (light blue color coding) were placed laterally of the outer canthi of the left and right eye to assess horizontal oculoversions (HOVs) based on the bipolar EOG signal (i.e., difference between EOG1 and EOG2). Ground and reference electrodes were placed at Fpz (black color coding) and FCz (dark blue color coding), respectively.



windows based on an autoregressive model of order 100 (Burg algorithm; Soekadar et al., 2011). Dependent on the optimal SMR frequency showing the largest modulation between 8 and 13 Hz during motor imagination/attempted finger movements vs. rest, the accumulated power of a 3-Hz bin around that modulation frequency [frequency of interest (FOI) ± 1.5 Hz] was extracted. Lastly, SMR-ERD related to imagined or attempted right- or left-hand movements was computed according to the power method described by Pfurtscheller and Aranibar (1979):

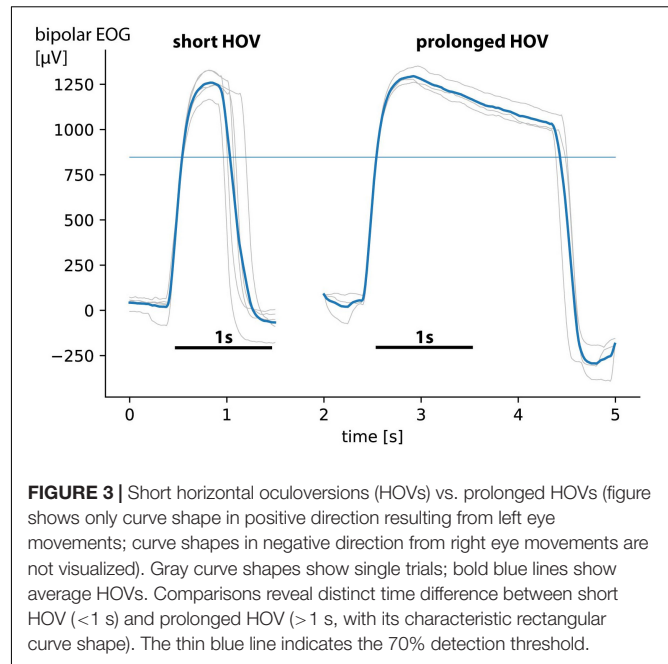
$$RV = \frac{1}{|T_{ref}|} \sum_{t \in T_{ref}} P_t \quad (1)$$

$$ERD(t) = \frac{P_t - RV}{RV} \times 100 \% \quad (2)$$

where P_t is the estimated power of the 3-Hz-wide bin at every sample block t . RV is the reference value to normalize power P_t to receive the instantaneous $ERD(t)$ at every sample block t . Notably, to receive ERD related to Laplace-filtered C3 (C3-ERD) and C4 (C4-ERD) EEG signals, two identical SMR-ERD processing pipelines were implemented in parallel for online calculation.

Brain-Computer Interface Calibration and Familiarization

To calibrate HOV detection thresholds for each side, participants were instructed to perform 5 *short* as well as 3 *prolonged* HOVs to each side, respectively. HOV detection thresholds were set at $\pm 70\%$ of median single-trial EOG maxima and minima (median was selected to receive a more robust



estimation; Figure 3). To determine the C3- as well as C4-ERD detection thresholds, two calibration runs were conducted. During the first run, participants were instructed to either imagine (healthy participants)/attempt (tetraplegics) left or right finger movements (active phases) or to relax (rest phases) according to 20 externally paced randomized visual cues lasting 5 s each. After each active or rest phase, an intertrial interval (ITI) with a randomized length of 4–6 s followed. After the first run, FOI was set to the optimal SMR frequency, and RVs for C3 and C4 were determined as average power of the entire run including all active and rest phases as well as all ITIs. During the second run, which consisted the same 20 visual cues, participants received online visual feedback based on their elicited SMR-ERD at C3 and C4. Finally, individual SMR-ERD detection thresholds were set to the average C3- and C4-ERD elicited within all active phases, respectively. After successful calibration, several familiarization runs were performed until the participant felt comfortable with all control commands.

Electroencephalography/ Electrooculography-Based Bilateral Control Paradigm

The EEG/EOG-based bilateral control paradigm was implemented as a hierarchical classifier with two sequential binary classification stages. This is a common approach to decompose the multiclass classification problem into several binary classification problems (Lotte et al., 2018). At the first stage, a linear classifier detected *prolonged* HOVs either to the left or to the right to activate the respective exoskeleton. As soon as the HOV detection threshold was exceeded for longer than 1 s, the classifier recognized this as a volitional laterality switch and enabled the specific classifier at the second stage. Dependent on the selected exoskeleton, either C3- or C4-ERD was then continuously analyzed and translated

into closing motions as long as the laterality-specific ERD detection threshold was exceeded. The principle of this two-stage EEG/EOG-based hierarchical classifier is illustrated in **Figure 4A**. To open the closed exoskeleton or to interrupt (veto) an unintended closing motion, a *short* HOV to any direction reset the exoskeleton again. A *short* HOV was classified when HOV detection threshold was exceeded less than 1 s (see **Figure 3** for differences in HOV type). Such hybrid *short* EOG/EEG-based paradigm was already successfully applied in tetraplegics during unilateral hand exoskeleton control

(Soekadar et al., 2016). To ensure safety, *short* HOV commands had the highest priority to veto any ongoing action in case two EEG/EOG-based features were detected at the same time (see priority order in **Table 1**).

Study Protocol and Audiovisual Online Feedback

To test for feasibility and safety of the novel EEG/EOG-based bilateral control paradigm, healthy participants as well

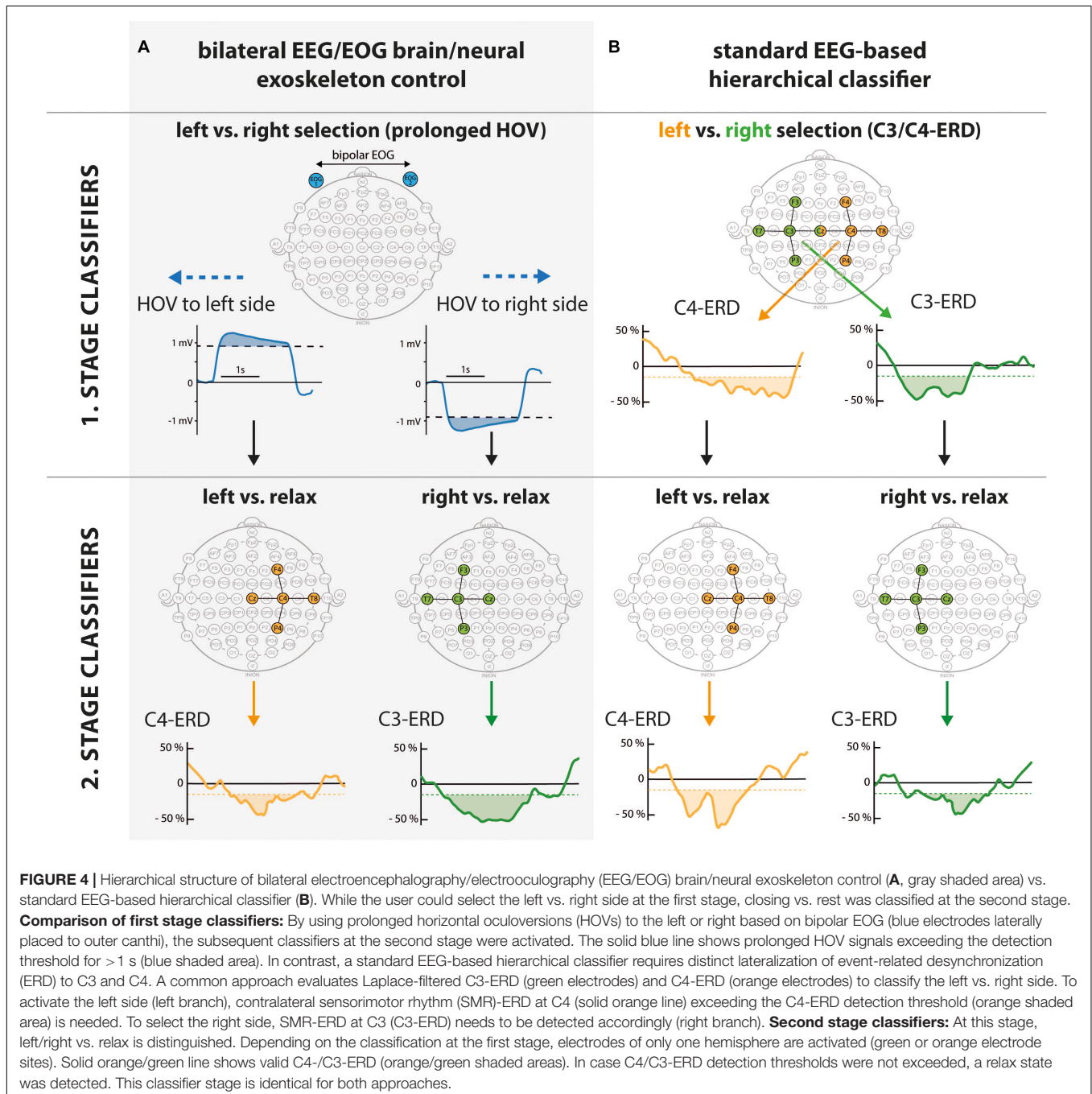
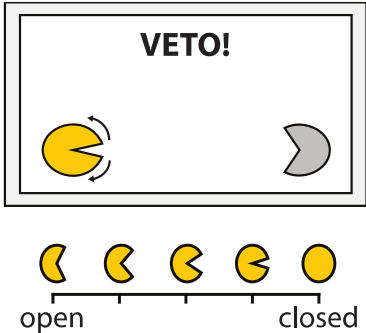
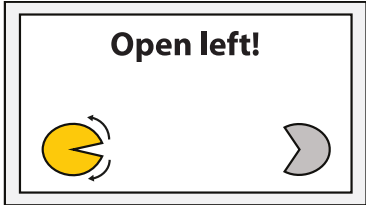
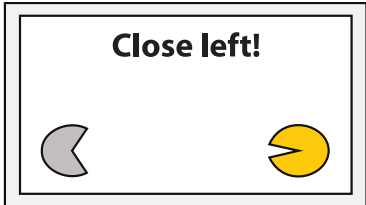
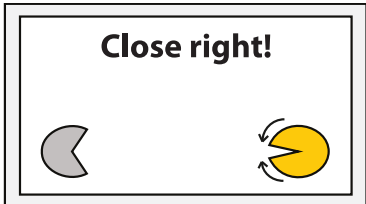
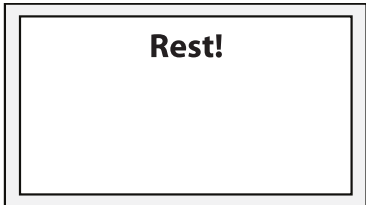


TABLE 1 | Overview of brain/neural-machine interface (B/NMI) control commands.

B/NMI control command	EEG/EOG-based feature	Respective task instruction with visual feedback
Interrupt closing motion	Short HOV toward any side Example: Short HOV to any side to interrupt (veto) the left ongoing exoskeleton motion	
Open exoskeleton	Short HOV toward the direction of activated exoskeleton Example: Short HOV to the left to open left closed exoskeleton	
Switch active exoskeleton	Prolonged HOV (> 1 s) toward desired hand exoskeleton Example: Before execute task instruction "Close left!," prolonged HOV to the left is required to activate left exoskeleton	
Close exoskeleton	SMR-ERD of contralateral motor cortex (C3- or C4-ERD) Example: C3-ERD required to close right exoskeleton	
Rest	No action required.	

The first column lists all possible commands for controlling each side. Importantly, the order of control commands listed in the table defines priority in case that two electroencephalography/electrooculography (EEG/EOG)-based features are detected at the same time starting with the highest priority at the top. The second column shows EEG/EOG-based features including examples for specific task instructions depicted in the third column. In the last column, visual feedback including exoskeleton motions for specific task instructions is illustrated. Yellow color coding indicates the active exoskeleton; gray color coding, the inactive exoskeleton. Only one possible instruction is illustrated for each control command. ERD, event-related desynchronization; HOV, horizontal oculoversion; SMR, sensorimotor rhythm.

as tetraplegics performed a pseudo-randomized sequence of $2 \times$ approximately 40 subtasks consisting of all B/NMI control commands required for bimanual operation of the two visual exoskeletons (Table 1). The sequence included subtasks to close one of the exoskeletons (requiring C4- or C3-ERD), to open them again, or to stop (*veto*) an ongoing closing motion as fast as possible to simulate for unintended hand exoskeleton motions or unexpected incidents (the latter two required both

short HOVs). In case a subtask required to close an exoskeleton, which had not been activated yet, participants first had to perform a *prolonged* HOV to the respective side before closing of the exoskeleton could be performed. To test for false positives, intervals to rest were randomly built in, in which the participants were instructed to avoid any action. A detailed overview on the bilateral B/NMI control commands, their corresponding EEG/EOG-based features, and their respective visual feedback

are summarized in **Table 1**. To enhance reliable distinction of *short* vs. *prolonged* HOV, an auditory feedback with two different sounds was provided to confirm successful HOV execution. The time between subtasks varied randomly between 5 and 7 s. Each sequence lasted approximately 5 min. In case no SMR-ERD was elicited, subtasks were aborted after 10 s. The total number of HOV-based subtasks being executed slightly varied depending on the users' previous SMR-ERD performance. For example, in case the user was not able to elicit ERD during a closing task, there was no need to reopen the exoskeleton again and was thus not requested. At the end of the session, tetraplegics rated user-friendliness of B/NMI control by using a five-level Likert-scale questionnaire. To account for the special needs of the tetraplegics, study protocols slightly differed between healthy participants and the patients. To reduce the overall session length, only six instead of eight rest phases were included. Moreover, the veto instructions were not randomly interspersed within the main study protocol but evaluated in a preceding pure EOG-based sequence. This was done to not overstrain the capabilities of the tetraplegic participants, since it was just a one-session study without any additional training day.

Outcome Measures and Offline Data Analysis

Feasibility and safety of the novel EEG/EOG-based control paradigm were assessed according to the following outcome measures. Feasibility was defined as *fluency* and *accuracy* of EEG-controlled operations. Fluency of control was evaluated as time from appearance of task instruction until exceedance of the SMR-ERD detection threshold. In case a laterality switch was required, timer count started just after successful activation of the exoskeleton (by performing a *prolonged* HOV). Fluent control was assumed when the average TTI such EEG-controlled operations ranged below 3 s (Crea et al., 2018). To assess the accuracy of bilateral control, the two-stage classifier performance was evaluated. At the first stage, exoskeleton selection was considered valid when successful *prolonged* HOV was performed. At the second stage, a trial was counted as successful when a full exoskeleton closing motion was conducted requiring the side-specific SMR-ERD detection threshold to be exceeded by a minimum of 2 s in total. Accurate bilateral control was assumed when the accuracy of all classifiers exceeded 80% in average. Due to the fact that the sequence can contain different numbers of subtasks, the balanced accuracy was applied to account for a potential bias toward the more frequent class (Brodersen et al., 2010). The balanced accuracy is given by $\frac{1}{2} \left(\frac{TP}{P} + \frac{TN}{N} \right)$ weighting the true-positive and true-negative rate equally. Since classification stages were built up as binary classifiers, chance level ranged at 50%. To compare the presented hybrid EEG/EOG-based classifier accuracy with an implementation, which was built up with EEG-based binary classifiers only, an offline data analysis was performed. The different implementation methods at the first stage are illustrated in **Figure 4**. Unlike the online implementation, in which prolonged HOV (first stage) and side-specific ERD (second stage) were used, offline classification was only based on the recorded

side-specific ERD (second stage of online paradigm) for both stages, since this was the classification while imagined/attempted finger movements were performed. This allowed comparison of the two approaches without the need to conduct two separate online sessions. Consequently, side-specific C3- and C4-ERDs were both classified depending on the instructed task. In case a left side closing was instructed, closing motions >2 s of the right or both exoskeletons or no movement were classified as false-negative events, whereas closing motion >2 s of the left exoskeleton was classified as a true-positive event. For the instruction to close the right side, the opposite events were classified: Movement of the right exoskeleton was classified as a true-negative event, while all other events were considered as false positives. To test for differences in average classification accuracy, a mixed-design analysis with "group" (healthy participants, tetraplegics) as between-group variable and "classification approach" (hybrid EEG/EOG brain/neural control, standard EEG-based hierarchical classifier) as repeated-measures variable was performed. To account for the limited number of data samples, bootstrapping was applied (Wilcox, 2011). Significance level was defined at $p < 0.05$. Safety was assumed when the TTS an unintended closing motion was interrupted within 2 s, meaning that closing motions were aborted before the exoskeleton was fully closed. Moreover, user-friendliness was met when the majority of tetraplegics rated EEG/EOG-based bilateral control as comfortable and easy to apply.

RESULTS

Feasibility

Average TTI [mean TTI \pm standard deviation (SD)] all EEG-controlled visual closing motions ranged at 2.14 ± 0.66 s across healthy participants and at 1.90 ± 0.97 s across tetraplegics, documenting *fluent* bilateral B/NMI control. **Figures 5A, 6A** show the individual TTI distribution for each participant. Only one healthy participant exceeded the fluency criterion (P04: 3.25 ± 2.65 s).

Average accuracy (mean \pm SD) for bilateral EEG/EOG brain/neural exoskeleton control ranged across all classifiers (i.e., including 1. stage classifier: prolonged HOV, and 2. stage classifier: C3-/C4-ERD) at $85.89 \pm 9.47\%$ across healthy participants and at $81.25 \pm 5.84\%$ across tetraplegics (**Figure 4A**). For the standard EEG-based hierarchical classifier, average accuracy declined across all classifiers to $71.33 \pm 17.21\%$ among healthy participants and to $58.68 \pm 10.62\%$ among tetraplegics (**Figure 4B**). There was a significant main effect of "classification approach" ($\Psi = -17.23, p < 0.001$), confirming superiority of the novel bilateral EEG/EOG brain/neural control for both healthy participants as well as tetraplegics. There was no main effect of "group" ($\Psi = 6.04, p = 0.419$) and no interaction between "classification approach" and "group" ($\Psi = 4.88, p = 0.449$). **Tables 2, 3** list individual accuracy rates for each healthy participant and tetraplegic as well as present accuracy results of all classifiers at every hierarchical classification stage. Chance level of binary classifiers ranged at 50%. Importantly, due to the novel implementation (compare **Figure 4A**), *prolonged* HOVs to

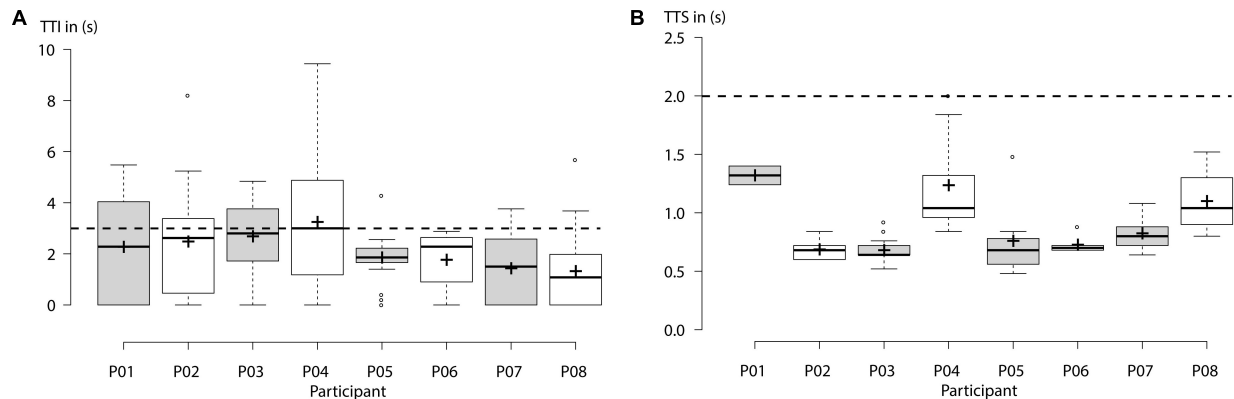


FIGURE 5 | Healthy participants: **(A)** “Time to initialize” (TTI) electroencephalography (EEG)-controlled closing motions of the left- or right-hand exoskeleton for each participant. Horizontal dashed line indicates the threshold for fluency criterion set at 3 s. Average TTI across all subjects ranged below 3 s, documenting fluent control. **(B)** “Time to stop” (TTS) an ongoing closing motion by using *short* horizontal oculoversions (HOVs). Horizontal dashed line indicates the threshold for safety criterion set at 2 s. Centerlines of boxplot show the median, while crosses show the mean. Box limits indicate the 25th and 75th percentiles.

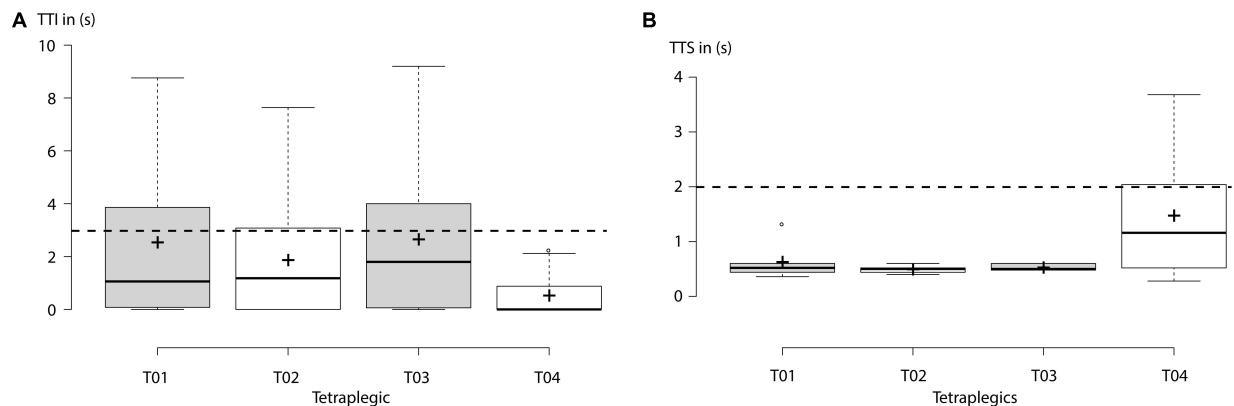


FIGURE 6 | Tetraplegics: **(A)** “Time to initialize” (TTI) electroencephalography (EEG)-controlled closing motions of the left- or right-hand exoskeleton for each participant. Horizontal dashed line indicates the threshold for the fluency criterion set at 3 s. Average TTI across all subjects ranged below 3 s, documenting fluent control. **(B)** “Time to stop” (TTS) an ongoing closing motion by using *short* horizontal oculoversions (HOVs). Horizontal dashed line indicates the threshold for the safety criterion set at 2 s. Only tetraplegic T04 exceeded the threshold of safety criterion. Centerlines of boxplot show the median, while crosses show the mean. Box limits indicate the 25th and 75th percentiles.

activate either the right or left exoskeleton at the first stage were classified in 100% of the cases.

Safety

Average TTS (mean TTS \pm SD) ongoing closing motions using *short* HOVs ranged at 0.92 ± 0.26 s across healthy participants and at 0.78 ± 0.46 s across tetraplegics. **Figures 5B, 6B** show the individual TTS distribution for each participant. Only one tetraplegic did not meet safety requirements while requiring more than 2 s to stop ongoing closing motions in some of the trials (T04: average TTS \pm SD ranged at 1.47 ± 1.24 s; **Figure 6B**).

User-Friendliness

With $88 \pm 11\%$ (mean \pm SD) of the maximum achievable score, tetraplegics rated the novel bilateral EEG/EOG-based control paradigm as user-friendly and reliable. More specifically, all tetraplegics answered that they did not experience any side effects or discomfort, that the calibration/control instructions

were easy to follow, and that the overall control was reliable and practical. Notably, all tetraplegics stated that the novel HOV-based control was easy to learn and that HOV control was comfortable. Importantly, three out of four tetraplegics would use the presented control to operate real hand exoskeletons bilaterally (**Figure 7**).

DISCUSSION

The presented study demonstrates feasibility and safety of a novel EEG/EOG-based B/NMI control paradigm for operating two hand exoskeletons. While feasibility was defined as fluency and accuracy of operation, safety was assumed when unintended closing motions could be aborted. We showed that eight healthy participants as well as four chronic tetraplegics were able to perform a complex sequence of subtasks mimicking bimanual tasks in daily life using four EEG/EOG-based control

TABLE 2 | Accuracy of bilateral electroencephalography/electrooculography (EEG/EOG) brain/neural exoskeleton control.

	Healthy participants				Tetraplegics			
	1. stage	2. stage		Total	1. stage	2. stage		Total
	Left/Right	Left/Rest	Right/Rest		Left/Right	Left/Rest	Right/Rest	
No.								
1	100.00	65.60	67.90	77.83	100.00	58.30	66.70	75.00
2	100.00	78.40	74.00	84.13	100.00	70.80	75.00	81.93
3	100.00	65.20	80.00	81.73	100.00	83.30	54.20	79.17
4	100.00	42.00	68.30	70.10	100.00	91.70	75.00	88.90
5	100.00	89.70	91.90	93.87				
6	100.00	94.40	100.00	98.13				
7	100.00	96.90	87.50	94.80				
8	100.00	78.40	81.20	86.53				
Mean	100.00	76.33	81.35	85.89	100.00	76.03	67.73	81.25
SD	0.00	18.34	11.36	9.47	0.00	14.61	9.83	5.84

Mean values with standard deviation (SD) are provided in bold.

TABLE 3 | Accuracy of standard electroencephalography (EEG)-based hierarchical classifier.

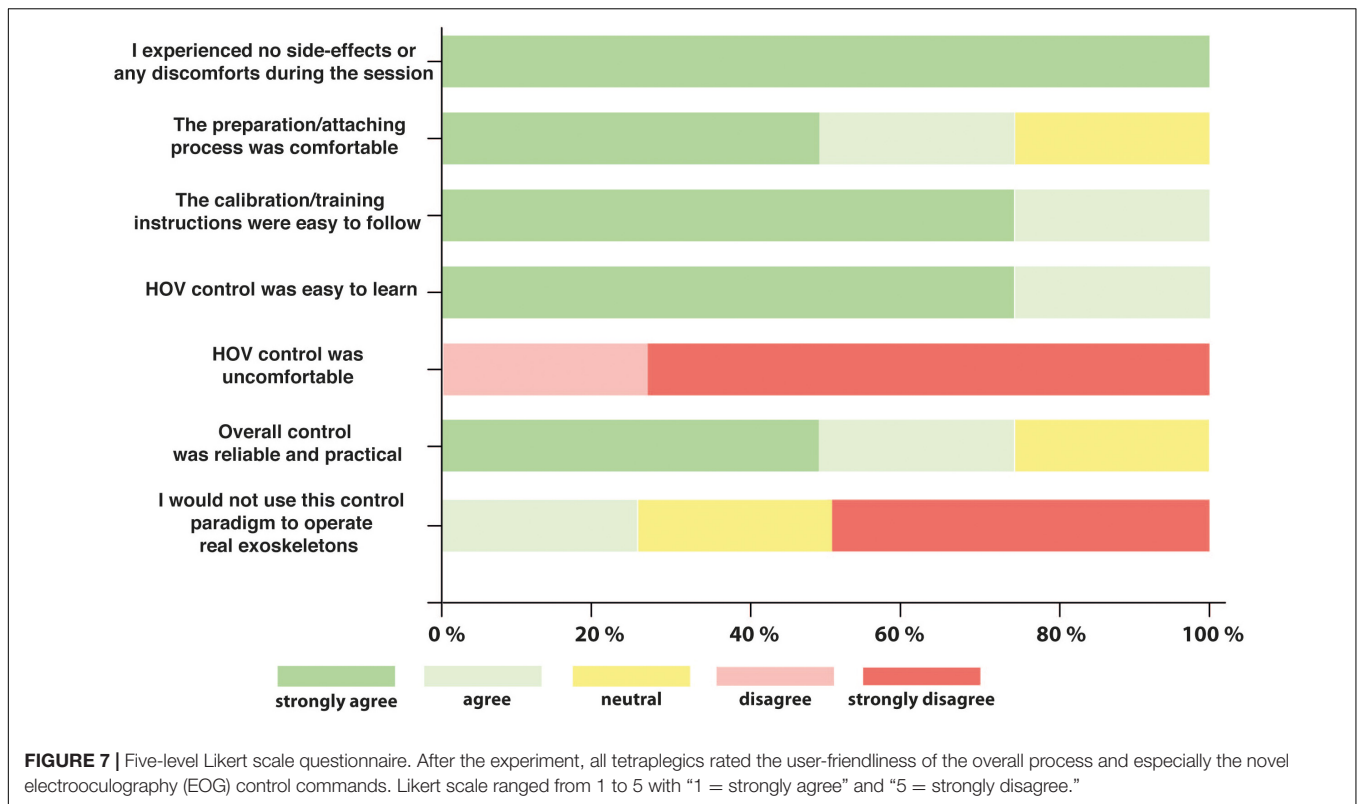
	Healthy participants				Tetraplegics			
	1. stage	2. stage		Total	1. stage	2. stage		Total
	Left/Right	Left/Rest	Right/Rest		Left/Right	Left/Rest	Right/Rest	
No.								
1	59.80	53.10	55.40	56.10	33.30	50.00	45.80	43.03
2	73.20	69.00	74.00	72.07	54.20	66.70	66.70	62.53
3	75.70	65.20	80.00	73.63	54.20	83.30	50.00	62.50
4	42.20	38.90	68.30	49.80	62.50	75.00	62.50	66.67
5	92.90	89.70	91.90	91.50				
6	83.30	94.40	100.00	92.57				
7	77.30	96.90	75.00	83.07				
8	27.30	59.70	68.80	51.93				
Mean	66.46	70.86	76.68	71.33	51.05	68.75	56.25	58.68
SD	22.03	20.99	14.05	17.21	12.46	14.22	9.94	10.62

Mean values with standard deviation (SD) are provided in bold.

commands [i.e., side-specific SMR-ERD at C3 or C4, as well as *prolonged* (>1 s) and *short* HOVs; **Table 1**]. Fluent control was documented by an average TTI EEG-controlled operations below 3 s (2.14 ± 0.66 s across healthy participants and 1.90 ± 0.97 s across tetraplegics). These results are comparable to those of previous studies, in which a unilateral whole-arm exoskeleton was controlled by healthy participants (Crea et al., 2018) or stroke survivors (Nann et al., 2020). Accurate control was confirmed by an average classification accuracy exceeding 80% ($85.89 \pm 15.81\%$ across healthy participants and $81.25 \pm 16.99\%$ across tetraplegics). Except for one tetraplegic, the TTS all ongoing motions were below 2 s (in average 0.92 ± 0.26 s across healthy participants and 0.78 ± 0.46 s across tetraplegics) underlining the system's safety. Finally, user-friendliness among tetraplegics was proven by stating no discomfort and ease of use in controlling the B/NMI system for bilateral operation with $88 \pm 11\%$ of the maximal achievable scores.

These results demonstrate for the first time that the presented hybrid EEG/EOG-based B/NMI control paradigm can be used for reliable and safe operation of two hand exoskeletons, e.g., to perform bimanual tasks.

Control of two exoskeletons requires classification of more than two classes (multiclass classification). This problem can be solved either by directly applying multiclass methods, such as naive Bayesian classifiers (Suk and Lee, 2012; Zhang et al., 2015) or multilayer perceptrons (Balakrishnan and Puthusserypady, 2005), or, as more commonly used, by decomposing the problem into several binary classifications (Lotte et al., 2018). There are different possible decomposition methods, e.g., pairwise classification (Vuckovic et al., 2018) or by hierarchical classification (Dong et al., 2017; Gundelakh et al., 2018). However, all studies have relatively low binary classification accuracies in common ranging from 50 to 70%. To achieve a higher control accuracy, fusion of EOG- and EEG-based features



was suggested and implemented in the presented bilateral control paradigm. A decisive step was to use a highly reliable EOG-based feature at safety-critical positions in the hierarchical classifier structure (Figure 4).

Fusing EEG with other biosignals like EOG is a well-established approach in the BCI field (Pfurtscheller et al., 2010). Soekadar et al. (2016) showed that such a hybrid EEG/EOG-based B/NHE fully restored hand function after SCI. Tetraplegics could eat and drink in a noisy outside restaurant by opening up the exoskeleton with *short* HOVs. This principle was now extended toward bilateral hand exoskeleton control introducing *prolonged* HOV. The advantage of this implementation was shown in the comparative offline analysis, where classification accuracy declined by 14.6% in healthy participants and by 22.6% in tetraplegics. The substantial decline in classification accuracy in tetraplegics compared to healthy participants underlines the need to compensate for the lack of lateralization in SCI by a reliable EOG-based switch between the two actuators.

One healthy participant (P04) did not meet the fluency criterion by 0.25 s in average, and one tetraplegic (T04) exceeded the safety criterion in some of the trials. However, in both cases, the unusually large SDs of 2.65 s for P04 and 1.24 s of T04 indicate that either the calibration threshold was not optimal or the participant did not attend to the task. Moreover, T04 was the only participant who stated that he would not want to use this paradigm in real life underpinning the previous assumptions.

Since EEG-based B/NMI control is generally more effortful than using other biosignals, e.g., electromyography (EMG) or HOV, one could argue that all exoskeleton movements could

be controlled by HOV. However, contrary to eye movements, EEG-based control was shown to be more intuitive since exoskeleton closing motions are directly linked to imagining or attempting to move the paralyzed fingers (Soekadar et al., 2016).

Moreover, there is increasing evidence that repeated brain/neural control of exoskeletons can trigger neural recovery (Donati et al., 2016; Wagner et al., 2018). Therefore, a combination of both operational purposes, i.e., *assistive* and *restorative* use, was suggested (Soekadar et al., 2019; Soekadar and Nann, 2020). Here, the assistive neural exoskeleton is used as a technical aid for the physiotherapist to train the patient in performing ADLs. This hybrid approach promises to facilitate generalization of learned skills to real-life environments and may increase the impact of the rehabilitation treatment. The proposed B/NMI control paradigm paves the way toward implementation of such hybrid approach for restoration of bimanual ADLs.

Besides extending the existing EEG/EOG B/NMI control paradigm toward bilateral hand exoskeleton control, minimizing electrode biosignal recording sites constitutes another important step for everyday life applicability (Cavallo et al., 2020). Moreover, considering that the high classification accuracy (>80%) was achieved with a minimalistic setup of only nine EEG recording sites, this opens up new opportunities for an easy applicable EEG headset system without the need for time-consuming whole-head EEG recordings, which is usually needed for advanced CSP algorithms, achieving comparable classification results.

To reliably detect prolonged HOVs (>1 s), bipolar EOG signals have to contain low-frequency information. Therefore, a high-pass filter (lower cutoff frequency at 0.02 Hz) has to be used.

As low-frequency bands are prone to be susceptible to movement artifacts, e.g., related to head movements, it needs to be tested whether the proposed approach for bilateral brain/neural exoskeleton control can be applied under less controlled and very noisy conditions (e.g., in an outside restaurant). Here, using other EOG signal features that are less dependent on information in the lower frequency bands could overcome this issue.

Larger clinical studies are needed to investigate whether these results can be generalized toward a broader spectrum of SCI patients. While all participants rated the brain/neural control paradigm as fluent, further increasing fluency would be desirable. In this context, taking advantage of lateralized brain activity [e.g., in the form of lateralized potential shifts preceding voluntary movements, the so-called Bereitschaftspotential or BP (Nann et al., 2019), or movement-related cortical potentials (MRCPs; Schwarz et al., 2020)] may contribute toward such aim. Since it was shown that SMR-ERDs are more pronounced over the contralateral hemisphere (Nikulin et al., 2008), it might be possible using advanced signal-processing tools to determine the side of the intended movement by assessing such lateralized activity only.

DATA AVAILABILITY STATEMENT

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

REFERENCES

- Ahuja, C. S., Wilson, J. R., Nori, S., Kotter, M. R. N., Druschel, C., Curt, A., et al. (2017). Traumatic spinal cord injury. *Nat. Rev. Dis. Primers* 3:17018. doi: 10.1038/nrdp.2017.18
- Ajiboye, A. B., Willett, F. R., Young, D. R., Memberg, W. D., Murphy, B. A., Miller, J. P., et al. (2017). Restoration of reaching and grasping movements through brain-controlled muscle stimulation in a person with tetraplegia: a proof-of-concept demonstration. *Lancet* 389, 1821–1830. doi: 10.1016/S0140-6736(17)30601-3
- Alam, M., Rodrigues, W., Pham, B. N., and Thakor, N. V. (2016). Brain-machine interface facilitated neurorehabilitation via spinal stimulation after spinal cord injury: recent progress and future perspectives. *Brain Res.* 1646, 25–33. doi: 10.1016/j.brainres.2016.05.039
- Anderson, K. D. (2004). Targeting recovery: priorities of the spinal cord-injured population. *J. Neurotrauma* 21, 1371–1383. doi: 10.1089/neu.2004.21.1371
- Balakrishnan, D., and Puthusserypady, S. (2005). “Multilayer perceptrons for the classification of brain computer interface data,” in *Proceedings of the IEEE 31st Annual Northeast Bioengineering Conference, 2005 IEEE*, Hoboken, NJ, 118–119.
- Bednar, M. S., and Woodside, J. C. (2018). Management of upper extremities in tetraplegia: current concepts. *JAAOS J. Am. Acad. Orthopaedic Surg.* 26, e333–e341. doi: 10.5435/jaaos-d-15-00465
- Benabid, A. L., Costecalde, T., Eliseyev, A., Charvet, G., Verney, A., Karakas, S., et al. (2019). An exoskeleton controlled by an epidural wireless brain-machine interface in a tetraplegic patient: a proof-of-concept demonstration. *Lancet Neurol.* 18, 1112–1122. doi: 10.1016/s1474-4422(19)30321-7
- Bouton, C. E., Shaikhouni, A., Annetta, N. V., Bockbrader, M. A., Friedenber, D. A., Nielson, D. M., et al. (2016). Restoring cortical control of functional movement in a human with quadriplegia. *Nature* 533, 247–250. doi: 10.1038/nature17435
- Brodersen, K. H., Ong, C. S., Stephan, K. E., and Buhmann, J. M. (2010). “The balanced accuracy and its posterior distribution,” in *Proceedings of the 2010 20th International Conference on Pattern Recognition: IEEE*, Istanbul, 3121–3124.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Commission at the Medical Faculty of the Eberhard Karls University and the University Hospital Tübingen. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

MN and NP designed the study. NP collected the data of the healthy participants. MN, NP, and CA collected the data of the tetraplegics. MN and NP analyzed the data. MN, NP, CA, and SS interpreted the data and performed the literature search. MN, NP, CA, and SS wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the Baden-Württemberg Stiftung (NEU007/1), the European Research Council (ERC-2017-STG-759370), and the Einstein Foundation Berlin. We acknowledge support from the German Research Foundation (DFG) and the Open Access Publication Fund of Charité–Universitätsmedizin Berlin.

- Bunketorp-Käll, L., Wangdell, J., Reinholdt, C., and Fridén, J. (2017). Satisfaction with upper limb reconstructive surgery in individuals with tetraplegia: the development and reliability of a Swedish self-reported satisfaction questionnaire. *Spinal Cord* 55, 664–671. doi: 10.1038/sc.2017.12
- Campbell, M. L., Sheets, D., and Strong, P. S. (1999). Secondary health conditions among middle-aged individuals with chronic physical disabilities: implications for unmet needs for services. *Assist. Technol.* 11, 105–122. doi: 10.1080/10400435.1999.10131995
- Cappello, L., Meyer, J. T., Galloway, K. C., Peisner, J. D., Granberry, R., Wagner, D. A., et al. (2018). Assisting hand function after spinal cord injury with a fabric-based soft robotic glove. *J. Neuroeng. Rehabil.* 15, 59–59. doi: 10.1186/s12984-018-0391-x
- Carvalho, R., Dias, N., and Cerqueira, J. J. (2019). Brain-machine interface of upper limb recovery in stroke patients rehabilitation: a systematic review. *Physiother. Res. Int.* 24:e1764. doi: 10.1002/pri.1764
- Cavallo, A., Roth, V., Haslacher, D., Nann, M., and Soekadar, S. R. (2020). Minimizing biosignal recording sites for noninvasive hybrid brain/neural control. *IEEE Syst. J.* 1–7. doi: 10.1109/JSYST.2020.3021751
- Collinger, J. L., Boninger, M. L., Bruns, T. M., Curley, K., Wang, W., and Weber, D. J. (2013a). Functional priorities, assistive technology, and brain-computer interfaces after spinal cord injury. *J. Rehabil. Res. Dev.* 50:145. doi: 10.1682/jrrd.2011.11.0213
- Collinger, J. L., Wodlinger, B., Downey, J. E., Wang, W., Tyler-Kabara, E. C., Weber, D. J., et al. (2013b). High-performance neuroprosthetic control by an individual with tetraplegia. *Lancet* 381, 557–564. doi: 10.1016/S0140-6736(12)61816-9
- Coscia, M., Wessel, M. J., Chaudary, U., Millán, J. D. R., Micera, S., Guggisberg, A., et al. (2019). Neurotechnology-aided interventions for upper limb motor rehabilitation in severe chronic stroke. *Brain* 142, 2182–2197. doi: 10.1093/brain/awz181
- Crea, S., Nann, M., Trigili, E., Cordella, F., Baldoni, A., Badesa, F. J., et al. (2018). Feasibility and safety of shared EEG/EOG and vision-guided autonomous whole-arm exoskeleton control to perform activities of daily living. *Sci. Rep.* 8:10823. doi: 10.1038/s41598-018-29091-5

- Dahlberg, L. S., Becerra, L., Borsook, D., and Linnman, C. (2018). Brain changes after spinal cord injury, a quantitative meta-analysis and review. *Neurosci. Biobehav. Rev.* 90, 272–293. doi: 10.1016/j.neubiorev.2018.04.018
- Donati, A. R., Shokur, S., Morya, E., Campos, D. S., Moiola, R. C., Gitti, C. M., et al. (2016). Long-term training with a brain-machine interface-based gait protocol induces partial neurological recovery in paraplegic patients. *Sci. Rep.* 6:30383. doi: 10.1038/srep30383
- Dong, E., Li, C., Li, L., Du, S., Belkacem, A. N., and Chen, C. (2017). Classification of multi-class motor imagery with a novel hierarchical SVM algorithm for brain-computer interfaces. *Med. Biol. Eng. Comput.* 55, 1809–1818. doi: 10.1007/s11517-017-1611-4
- Dunn, J. A., Sinnott, K. A., Rothwell, A. G., Mohammed, K. D., and Simcock, J. W. (2016). Tendon transfer surgery for people with tetraplegia: an overview. *Arch. Phys. Med. Rehabil.* 97, S75–S80.
- Frolov, A. A., Mokienko, O., Lyukmanov, R., Biryukova, E., Kotov, S., Turbina, L., et al. (2017). Post-stroke rehabilitation training with a motor-imagery-based brain-computer interface (BCI)-controlled hand exoskeleton: a randomized controlled multicenter trial. *Front. Neurosci.* 11:400. doi: 10.3389/fnins.2017.00400
- Gohritz, A., and Fridén, J. (2018). Management of spinal cord injury-induced upper extremity spasticity. *Hand Clin.* 34, 555–565. doi: 10.1016/j.hcl.2018.07.001
- Gundelakh, F., Stankevich, L., and Sonkin, K. (2018). “Mobile robot control based on noninvasive brain-computer interface using hierarchical classifier of imagined motor commands,” in *Proceedings of the MATEC Web of Conferences*, (Les Ulis: EDP Sciences), 03003. doi: 10.1051/mateconf/201816103003
- Heide, W., Koenig, E., Trillenber, P., Kompf, D., and Zee, D. S. (1999). Electrooculography: technical standards and applications. The International Federation of Clinical Neurophysiology. *Electroencephalogr. Clin. Neurophysiol. Suppl.* 52, 223–240.
- Ho, C. H., Triolo, R. J., Elias, A. L., Kilgore, K. L., Dimarco, A. F., Bogie, K., et al. (2014). Functional electrical stimulation and spinal cord injury. *Phys. Med. Rehabil. Clin.* 25, 631–654.
- Hochberg, L. R., Bacher, D., Jarosiewicz, B., Masse, N. Y., Simeral, J. D., Vogel, J., et al. (2012). Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature* 485, 372–375. doi: 10.1038/nature11076
- Holtz, K. A., Lipson, R., Noonan, V. K., Kwon, B. K., and Mills, P. B. (2017). Prevalence and effect of problematic spasticity after traumatic spinal cord injury. *Arch. Phys. Med. Rehabil.* 98, 1132–1138. doi: 10.1016/j.apmr.2016.09.124
- León, C. L. (2017). *Multilabel Classification of EEG-based Combined Motor Imageries Implemented for the 3D Control of a Robotic Arm*. Nancy: University of Lorraine.
- Lo, C., Tran, Y., Anderson, K., Craig, A., and Middleton, J. (2016). Functional priorities in persons with spinal cord injury: using discrete choice experiments to determine preferences. *J. Neurotrauma* 33, 1958–1968. doi: 10.1089/neu.2016.4423
- Lotte, F., Bougrain, L., Cichocki, A., Clerc, M., Congedo, M., Rakotomamonjy, A., et al. (2018). A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update. *J. Neural Eng.* 15:031005. doi: 10.1088/1741-2552/aab2f2
- McFarland, D. J. (2015). The advantages of the surface Laplacian in brain-computer interface research. *Int. J. Psychophysiol.* 97, 271–276. doi: 10.1016/j.ijpsycho.2014.07.009
- Mekki, M., Delgado, A. D., Fry, A., Putrino, D., and Huang, V. (2018). Robotic rehabilitation and spinal cord injury: a narrative review. *Neurotherapeutics* 15, 604–617.
- Meng, J., Zhang, S., Bekyo, A., Olsoe, J., Baxter, B., and He, B. (2016). Noninvasive electroencephalogram based control of a robotic arm for reach and grasp tasks. *Sci. Rep.* 6:38565.
- Nann, M., Cohen, L. G., Deecke, L., and Soekadar, S. R. (2019). To jump or not to jump—the Bereitschaftspotential required to jump into 192-meter abyss. *Sci. Rep.* 9:2243. doi: 10.1038/s41598-018-38447-w
- Nann, M., Cordella, F., Trigili, E., Lauretti, C., Bravi, M., Miccinilli, S., et al. (2020). Restoring activities of daily living using an EEG/EOG-controlled semi-autonomous and mobile whole-arm exoskeleton in chronic stroke. *IEEE Syst. J.* 1–8. doi: 10.1109/JSYST.2020.3021485
- National Spinal Cord Injury Statistical Center (2019). *National Spinal Cord Injury Statistical Center Annual Statistical Report*. Birmingham, AL: University of Alabama at Birmingham.
- Neuper, C., Wortz, M., and Pfurtscheller, G. (2006). ERD/ERS patterns reflecting sensorimotor activation and deactivation. *Prog. Brain Res.* 159, 211–222. doi: 10.1016/S0079-6123(06)59014-4
- Nikulin, V. V., Hohlefeld, F. U., Jacobs, A. M., and Curio, G. (2008). Quasi-movements: a novel motor-cognitive phenomenon. *Neuropsychologia* 46, 727–742. doi: 10.1016/j.neuropsychologia.2007.10.008
- Ortner, R., Scharinger, J., Lechner, A., and Guger, C. (2015). “How many people can control a motor imagery based BCI using common spatial patterns?” in *Proceedings of the 2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*: IEEE, Montpellier, 202–205.
- Osugwu, B. C., Wallace, L., Fraser, M., and Vuckovic, A. (2016). Rehabilitation of hand in subacute tetraplegic patients based on brain computer interface and functional electrical stimulation: a randomised pilot study. *J. Neural Eng.* 13:065002. doi: 10.1088/1741-2560/13/6/065002
- Pfurtscheller, G., Allison, B. Z., Brunner, C., Bauernfeind, G., Solis-Escalante, T., Scherer, R., et al. (2010). The hybrid BCI. *Front. Neurosci.* 4:30. doi: 10.3389/fnpro.2010.00003
- Pfurtscheller, G., and Aranibar, A. (1979). Evaluation of event-related desynchronization (ERD) preceding and following voluntary self-paced movement. *Electroencephalogr. Clin. Neurophysiol.* 46, 138–146. doi: 10.1016/0013-4694(79)90063-4
- Pfurtscheller, G., and da Silva, F. H. L. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin. Neurophysiol.* 110, 1842–1857. doi: 10.1016/S1388-2457(99)00141-8
- Ragnarsson, K. (2008). Functional electrical stimulation after spinal cord injury: current use, therapeutic effects and future directions. *Spinal Cord* 46, 255–274. doi: 10.1038/sj.sc.3102091
- Rushton, P. W., Miller, W. C., Mortenson, W. B., and Garden, J. (2010). Satisfaction with participation using a manual wheelchair among individuals with spinal cord injury. *Spinal Cord* 48, 691–696. doi: 10.1038/sc.2009.197
- Schalk, G., Mcfarland, D. J., Hinterberger, T., Birbaumer, N., and Wolpaw, J. R. (2004). BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Trans. Biomed. Eng.* 51, 1034–1043. doi: 10.1109/TBME.2004.827072
- Schwarz, A., Pereira, J., Kobler, R., and Muller-Putz, G. R. (2020). Unimanual and bimanual reach-and-grasp actions can be decoded from human EEG. *IEEE Trans. Biomed. Eng.* 67, 1684–1695. doi: 10.1109/TBME.2019.2942974
- Snoek, G. J., Mj, I. J., Hermens, H. J., Maxwell, D., and Biering-Sorensen, F. (2004). Survey of the needs of patients with spinal cord injury: impact and priority for improvement in hand function in tetraplegics. *Spinal Cord* 42, 526–532. doi: 10.1038/sj.sc.3101638
- Soekadar, S. R., Cohen, L. G., and Birbaumer, N. (2015a). “Clinical brain-machine interfaces,” in *Cognitive Plasticity in Neurologic Disorders*, eds J. I. Tracy and B. M. Hampstead (Oxford: Oxford University Press), 347–363.
- Soekadar, S. R., Haagen, K., and Birbaumer, N. (2008). “Brain-computer interfaces (BCI): restoration of movement and thought from neuroelectric and metabolic brain activity,” in *Coordination: Neural, Behavioral and Social Dynamics*, eds A. Fuchs and V. K. Jirsa (Berlin: Springer). doi: 10.1007/978-3-540-74479-5_11
- Soekadar, S. R., and Nann, M. (2020). “Neural-gesteuerte robotik für assistenz und rehabilitation im alltag,” in *Mensch-Roboter-Kollaboration*, ed. H. J. Buxbaum (Wiesbaden: Springer), 117–131. doi: 10.1007/978-3-658-28307-0_8
- Soekadar, S. R., Nann, M., Crea, S., Trigili, E., Gómez, C., Opisso, E., et al. (2019). “Restoration of finger and arm movements using hybrid brain/neural assistive technology in everyday life environments,” in *Brain-Computer Interface Research, A State-of-the-Art Summary* 7, eds N. M.-K. Christoph Guger and B. Z. Allison (Berlin: Springer International Publishing), 53–61. doi: 10.1007/978-3-030-05668-1_5
- Soekadar, S. R., Witkowski, M., Gómez, C., Opisso, E., Medina, J., Cortese, M., et al. (2016). Hybrid EEG/EOG-based brain/neural hand exoskeleton restores fully independent daily living activities after quadriplegia. *Sci. Robot.* 1:eag3296. doi: 10.1126/scirobotics.aag3296
- Soekadar, S. R., Witkowski, M., Mellinger, J., Ramos, A., Birbaumer, N., and Cohen, L. G. (2011). ERD-based online brain-machine interfaces (BMI) in the context of neurorehabilitation: optimizing BMI learning and performance. *IEEE Trans. Neural Syst. Rehabil. Eng.* 19, 542–549. doi: 10.1109/TNSRE.2011.2166809

- Soekadar, S. R., Witkowski, M., Vitiello, N., and Birbaumer, N. (2015b). An EEG/EOG-based hybrid brain-neural computer interaction (BNCI) system to control an exoskeleton for the paralyzed hand. *Biomed. Tech.* 60, 199–205. doi: 10.1515/bmt-2014-0126
- Suk, H.-I., and Lee, S.-W. (2012). A novel Bayesian framework for discriminative feature extraction in brain-computer interfaces. *IEEE Trans. Pattern Anal. Machine Intell.* 35, 286–299. doi: 10.1109/tpami.2012.69
- Tang, Z., Sun, S., Zhang, S., Chen, Y., Li, C., and Chen, S. (2016). A brain-machine interface based on ERD/ERS for an upper-limb exoskeleton control. *Sensors* 16:2050. doi: 10.3390/s16122050
- Vidaurre, C., and Blankertz, B. (2010). Towards a cure for BCI illiteracy. *Brain Topogr.* 23, 194–198. doi: 10.1007/s10548-009-0121-6
- Vidaurre, C., Klauer, C., Schauer, T., Ramos-Murguialday, A., and Müller, K.-R. (2016). EEG-based BCI for the linear control of an upper-limb neuroprosthesis. *Med. Eng. Phys.* 38, 1195–1204. doi: 10.1016/j.medengphy.2016.06.010
- Vuckovic, A., Pangaro, S., and Finda, P. (2018). Unimanual versus bimanual motor imagery classifiers for assistive and rehabilitative brain computer interfaces. *IEEE Trans. Neural Syst. Rehabil. Eng.* 26, 2407–2415. doi: 10.1109/tnsre.2018.2877620
- Wagner, F. B., Mignardot, J. B., Le Goff-Mignardot, C. G., Demesmaeker, R., Komi, S., Capogrosso, M., et al. (2018). Targeted neurotechnology restores walking in humans with spinal cord injury. *Nature* 563, 65–71. doi: 10.1038/s41586-018-0649-2
- Wilcox, R. R. (2011). *Introduction to Robust Estimation and Hypothesis Testing*. Cambridge, MA: Academic press.
- Witkowski, M., Cortese, M., Cempini, M., Mellinger, J., Vitiello, N., and Soekadar, S. R. (2014). Enhancing brain-machine interface (BMI) control of a hand exoskeleton using electrooculography (EOG). *J. Neuroeng. Rehabil.* 11:165. doi: 10.1186/1743-0003-11-165
- Wolpaw, J. R., Birbaumer, N., Mcfarland, D. J., Pfurtscheller, G., and Vaughan, T. M. (2002). Brain-computer interfaces for communication and control. *Clin. Neurophysiol.* 113, 767–791. doi: 10.1016/s1388-2457(02)00057-3
- Yoo, H.-J., Lee, S., Kim, J., Park, C., and Lee, B. (2019). Development of 3D-printed myoelectric hand orthosis for patients with spinal cord injury. *J. Neuro Eng. Rehabil.* 16:162.
- Yun, Y., Dancausse, S., Esmatloo, P., Serrato, A., Merring, C. A., Agarwal, P., et al. (2017). “Maestro: an EMG-driven assistive hand exoskeleton for spinal cord injury patients,” in *Proceedings of the 2017 IEEE International Conference on Robotics and Automation (ICRA): IEEE*, Singapore, 2904–2910.
- Zhang, Y., Zhou, G., Jin, J., Zhao, Q., Wang, X., and Cichocki, A. (2015). Sparse Bayesian classification of EEG for brain-computer interface. *IEEE Trans. Neural Netw. Learn. Syst.* 27, 2256–2267. doi: 10.1109/tnnls.2015.2476656

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

Copyright © 2020 Nann, Peekhaus, Angerhöfer and Soekadar. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Using Posterior EEG Theta Band to Assess the Effects of Architectural Designs on Landmark Recognition in an Urban Setting

James D. Rounds¹, Jesus Gabriel Cruz-Garza² and Saleh Kalantari^{2*}

¹ Human Development, Cornell University, Ithaca, NY, United States, ² Department of Design and Environmental Analysis, Cornell University, Ithaca, NY, United States

OPEN ACCESS

Edited by:

Surjo R. Soekadar,
Charité – Universitätsmedizin
Berlin, Germany

Reviewed by:

Eleonora Vecchio,
University of Bari Aldo Moro, Italy
Gustavo Pacheco-Lopez,
Autonomous Metropolitan
University, Mexico

*Correspondence:

Saleh Kalantari
sk3268@cornell.edu

Specialty section:

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

Received: 17 July 2020

Accepted: 13 November 2020

Published: 11 December 2020

Citation:

Rounds JD, Cruz-Garza JG and
Kalantari S (2020) Using Posterior
EEG Theta Band to Assess the Effects
of Architectural Designs on Landmark
Recognition in an Urban Setting.
Front. Hum. Neurosci. 14:584385.
doi: 10.3389/fnhum.2020.584385

The process of urban landmark-based navigation has proven to be difficult to study in a rigorous fashion, primarily due to confounding variables and the problem of obtaining reliable data in real-world contexts. The development of high-resolution, immersive virtual reality technologies has opened exciting new possibilities for gathering data on human wayfinding that could not otherwise be readily obtained. We developed a research platform using a virtual environment and electroencephalography (EEG) to better understand the neural processes associated with landmark usage and recognition during urban navigation tasks. By adjusting the architectural parameters of different buildings in this virtual environment, we isolated and tested specific design features to determine whether or not they served as a target for landmarking. EEG theta band (4–7 Hz) event-related synchronization/desynchronization over posterior scalp areas was evaluated at the time when participants observed each target building along a predetermined self-paced route. A multi-level linear model was used to investigate the effects of salient architectural features on posterior scalp areas. Our results support the conclusion that highly salient architectural features—those that contrast sharply with the surrounding environment—are more likely to attract visual attention, remain in short-term memory, and activate brain regions associated with wayfinding compared with non-salient buildings. After establishing this main aggregate effect, we evaluated specific salient architectural features and neural correlates of navigation processing. The buildings that most strongly associated extended gaze time, location recall accuracy, and changes in theta-band neural patterns with landmarking in our study were those that incorporated rotational twist designs and natural elements such as trees and gardens. Other building features, such as unusual façade patterns or building heights, were to a lesser extent also associated with landmarking.

Keywords: EEG, virtual reality, architectural design, landmark recognition, wayfinding

INTRODUCTION AND BACKGROUND

More than 55% of the world's population currently resides in urban environments, and that percentage is expected to increase in upcoming decades (United Nations, 2018). The architectural design of these built environments can have a significant impact on those who live there (Kalantari and Shepley, 2020). Numerous cognitive and health-related effects associated with urban design have been documented in recent years (e.g., Nasar, 1994; Heft and Nasar, 2000; Knöll et al., 2018; Kondo et al., 2018). However, in some topical areas it has been difficult to obtain rigorous empirical data about the effects of urban design. Wayfinding is one of these problematic areas. There has long been speculation and anecdotal evidence that the design of cities can make wayfinding much easier or harder for people who live in and visit these environments—Tolman (1948) introduced the concept of “cognitive maps” to discuss neural wayfinding processes, a concept that was later used by Kevin Lynch, in *The Image of the City* (1960), to explain human wayfinding in the urban environment. More recently, researchers have adjusted and refined this analysis to discuss the specific mental processes that are involved in learning routes and establishing orientation in cities (Blades et al., 2002; Julian and Epstein, 2013), as well as the ways in which people make use of navigation tools such as signs and maps (He et al., 2015). Gradually, this research is beginning to take on a more rigorous character, with a growing interest in collecting both behavioral and neurological data.

One crucial aspect of wayfinding that has been gaining increasing scholarly attention is the use of landmarks. May et al. (2003) found that landmarks are the most common type of cue used for navigation by pedestrians in urban centers. Lee et al. (2006) demonstrated that the availability of landmarks greatly reduces the number of mistakes made during wayfinding, and that landmarks also reduce the stress of wayfinding decisions. Specific studies have identified landmark-use in a wide array of navigation contexts, including city exteriors and the internal navigation of hospitals, airports, train stations, and other large urban structures (Fewings, 2001; Joseph, 2006; Epstein and Vass, 2014; Chang and Zheng, 2016; Sharma et al., 2017). Several studies have found that landmarks facilitate the development of spatial orientation skills in children (Howard and Templeton, 1966; Acredolo, 1977; Sadalla et al., 1980) and that they are particularly important wayfinding tools for older adults (Burns, 1997; Bradley and Dunlop, 2005). Other researchers have noted that landmarks are frequently used when one person attempts to communicate a route, convey a mental model of an environment, or give directions to another individual (Siegel and White, 1975; Egenhofer and Mark, 1995; Maass and Schmauks, 1998; Lovelace et al., 1999; Michon and Denis, 2001; Tenbrink, 2008).

Saliency, which refers to the contrast between a landmark and its surrounding terrain, is widely recognized as a central component of landmark selection (Clark, 1996; Sorrows and Hirtle, 1999; Raubal and Winter, 2002; Caduff and Timpf, 2008). Landmarks generally need to stand out from their surroundings in order to be easily recognized and used as

spatial reference points in memory and in communication. If individuals are required to use landmarks with low saliency, then the wayfinding process will be more difficult and time-consuming, with commensurate increases in cognitive burdens and a greater likelihood of wayfinding errors (Sorrows and Hirtle, 1999). Various kinds of landmark saliency have been proposed by researchers, including distinctive visual patterns or colors, structural/geometric anomalies, and memorable cultural associations (Sorrows and Hirtle, 1999; Raubal and Winter, 2002; Caduff and Timpf, 2008; Grabler et al., 2008). In the design field, there have been numerous attempts to quantify and categorize the specific saliency factors that can promote landmark identification and recognition, with the goal of improving the design of built environments and making wayfinding easier (Raubal and Winter, 2002; Nothegger et al., 2004; Klippel and Winter, 2005; Aziz and Mertsching, 2007; Claramunt and Winter, 2007; Duckham et al., 2010; Zhang et al., 2014). Some researchers have even developed algorithmic and data-mining methods to automatically identify potential salient landmarks in virtual spaces or in architectural design plans (Elias, 2003; Peters et al., 2010; Winter et al., 2008).

Despite all of this research on landmark saliency, very little scholarly consensus has emerged regarding what features in architectural design are most likely to produce readily identifiable landmarks for use by diverse human populations. Part of the reason for the conflicting and inconclusive nature of these studies may be methodological. Most of the previous studies on urban landmarking have been based on showing participants pictures of urban scenes and asking them to find or identify landmarks. This approach is limited due to the static, two-dimensional impression given by the pictures, which may not reflect the same saliency factors that are present when people move through complex, dynamic, and immersive environments.

Another concern in the existing landmarking design literature is that these analyses have relied mostly on behavioral and/or self-reported data (Cornwell et al., 2008; Pu et al., 2017; Sharma et al., 2017). Less attention has been given to the possibility of collecting neurological data (such as EEG signals) as a more robust scientific basis to measure covert processes and to distinguish cause and effect in human responses to design features. EEG methods are a promising tool to characterize the interplay of neural states and information processing (Banaschewski and Brandeis, 2007). The goal of the current study was to address these concerns by developing an assessment method for landmark recognition using spectral analyses of scalp EEG electrodes over neural regions that have been implicated in spatial awareness and spatial memory. The use of an immersive, three-dimensional virtual environment (as opposed to static pictures) allowed the researchers to differentiate experimental conditions by making targeted adjustments to specific landmark design variables. It also enabled the collection of cleaner real-time neural data while participants completed active navigational and recall tasks (compared to real-world navigation that would generate motion artifacts in the EEG).

Brain Activity Related to Landmark Recognition

In an influential study, Epstein and Vass (2014) identified four cognitive mechanisms related to landmark-based navigation. These include: (a) the landmark-recognition mechanism, which clarifies “what am I looking at?”; (b) the localization/orientation process, which clarifies “where am I and which direction am I pointing?”; (c) the encoding and retrieval system for spatial knowledge, which clarifies “where are other specific places in relation to my position?”; and (d) the route-planning process, which solves the problem of “how do I get from here to another specific place?” In the present study, participants were tasked with passively observing buildings along a simulated urban route, and then asked to later recall where several “target” buildings were along the route. Therefore, of these mechanisms we hypothesized that more salient building features would increase activation of brain regions and networks associated with “(b)”, localization/reorientation, and “(c)”, encoding and retrieval of spatial knowledge.

Recent EEG and fMRI work has demonstrated that the task demands of integrating allocentric and egocentric information during spatial navigation are strongly associated with activity in the retrosplenial cortex (RSC; Auger et al., 2012; Lin et al., 2015; Fischer et al., 2020), as part of network of spatial awareness and memory centers including the hippocampus [HC], occipital place area [OPA], and the parahippocampal cortex [PHC], which includes the parahippocampal place area [PPA]; (Wolbers and Büchel, 2005; Hanslmayr et al., 2011; Julian et al., 2018). The RSC is generally located below the inferior and superior posterior scalp areas (e.g., EEG sites P3, P1, CP3, CP1 on the left, and P2, P4, CP2, CP4 on the right), and is highly amenable to EEG measurement, although its activity can be hard to disentangle from that of the PHC and posterior parietal cortex in general. And certainly, EEG data is limited in its spatial resolution, for several reasons. But carefully designed experiments with rodents and humans have revealed that the RSC seems uniquely involved in integrating “egocentric” spatial awareness—the participant’s sense of their current spatial position and direction in relation to previous personal positions and actions—with incoming “allocentric” information about the relative position of external objects in the environment during movement (Buzsáki, 2005; Dhindsa et al., 2014; Lin et al., 2015).

Multimodal evidence supports the role of theta band EEG, and the connectivity between the RSC and other areas, in this path integration process. In a tractography-based study, Ramanoël et al. (2019) found correlations between spatial memory skills and age-associated deficits in resting state structural connectivity between the left RSC and Hippocampus. And in a recent study involving patients with mild cognitive impairment and prodromal Alzheimer’s, virtual navigation skills were significantly negatively correlated with levels of disease biomarkers in RSC and HC (Howett et al., 2019). Support for the mechanistic role of theta activity among these regions comes from simultaneous EEG-fMRI experiments that show a negative correlation between Default-Mode Network activation (turning attention inward) and theta power in the superior posterior

parietal and frontal regions (Scheeringa et al., 2009; however, see Zumer et al., 2014), and that increased theta power during encoding (across several regions anti-correlated with the default mode network) predicts which encoded information is later remembered (White et al., 2013).

It follows naturally that researchers would look to the RSC and related areas when measuring the effect of different navigation strategies, as well as the effect of different environment features. Auger et al. (2012) found that RSC activity while viewing landmarks correlated with an individual’s skills as a navigator, and specifically when viewing buildings deemed more “permanent,” suggesting this brain area’s role is to help tag especially meaningful spatial cues during memory formation. And Lin et al. (2015) found that a participant’s natural navigation strategy preference (egocentric vs. allocentric) in a virtual navigation task was associated with several EEG markers, such as theta and alpha synchronization and desynchronization in sources localized to the retrosplenial and posterior parietal cortices during turning and new-scene encoding (among other findings). Previous research has also shown a scalp-wide increase in theta power after making a turn decision in a virtual navigation task (Bischof and Boulanger, 2003), and during decision-making phases in navigational tasks (Jaiswal et al., 2010). The theta band signals are thought to act as a mechanism by which different neuronal groups and regions synchronize with each other in order to accomplish objectives (Buzsáki, 2005). In both human and animal studies, theta activation in posterior parietal regions has been observed to feature prominently in navigational tasks (White et al., 2011; Belchior et al., 2014) and goal-directed environmental information processing (Cornwell et al., 2008; Pu et al., 2017), as well as in memory formation and recall (Paller and Wagner, 2002; Jaiswal et al., 2010; Vaidya and Johnston, 2013; Koike et al., 2017; Scholz et al., 2017). Based on these prior findings, in our study we focused on measuring theta activity in the medial and lateral superior posterior scalp locations while participants gazed at the target buildings of different hypothesized saliency levels.

We measured theta activation by calculating event-related desynchronization/ synchronization (ERD/ERS) averaged across the 4–7 Hz EEG frequency band, during periods (minimum 1.5 s) when participants gazed at the target buildings. A 1 min resting period was used as a baseline for comparing task-related theta-power changes, following Pfurtscheller and Lopes da Silva (1999), and Sharma et al. (2017).

The Use of Virtual Reality in Navigational Behavior Research

Today’s high-resolution virtual environments are becoming astonishingly lifelike, opening new opportunities to study various types of human behavior in a controlled context. The use of virtual reality (VR) is already widespread in behavioral (Makransky et al., 2019), cognitive (Wolbers and Büchel, 2005), medical (Plancher et al., 2012; Clay et al., 2020), and design research (Kalantari and Neo, 2020). This technology allows researchers to isolate and adjust environmental variables in a way that would not be possible in the real world. For example,

researchers can easily add or remove windows from an otherwise identical building, or change the color or pattern of a building's façade, without incurring any construction expenses. These types of design studies have the potential to provide an enormous wealth of empirical data (Jeffery, 2019). We must always remain aware that there are possible discrepancies between virtual and real-world results, but research using virtual environments is in many ways superior to prior studies that relied on static images and/or were rife with confounding variables. In many cases virtual platforms can provide an important way to obtain feedback about design questions that otherwise could not be rigorously tested at all.

With the groundwork for the VR largely in place, the use of this technology as a design research tool mostly requires the technological know-how to make targeted modifications in the virtual environment—along with any coding that is necessary to implement other desired research features, such as presenting real-time interactive questionnaires to study participants within the environment. Previous researchers have demonstrated important concordances between human behavior in real-world navigation and virtual navigation (Werner and Schindler, 2004; Jansen-Osmann et al., 2007; Jiang and Li, 2009; Tang et al., 2009; Kuliga et al., 2015; Slone et al., 2015; De Tommaso et al., 2016). Many of these studies used virtual contexts to develop hypotheses about behavior that were later confirmed in real-world environments. Of particular note for the current research are studies by Gazova et al. (2013) and Nys et al. (2015), which used VR as a research tool for evaluating the relative importance of landmark use during wayfinding. Similarly, Röser et al. (2012) used a virtual environment to evaluate the ideal position of a landmark at an intersection.

In the current study participants were immersed in a high-resolution VR urban environment and asked to complete various navigational tasks. In contrast with physical environment navigation, this allowed us to use neurophysiological sensors to better understand the impact of different architectural landmark designs during navigational memory formation. The use of the virtual environment made it possible to collect robust neurological data while reducing motion or sweat-related artifacts in the EEG signals. In addition, it allowed the researchers to carefully and precisely create the desired urban design testing environment and to readily move or substitute individual buildings and features during preparation, or condition creation, thus helping to isolate specific design variables. Our primary focus was to triangulate behavioral and neurological responses to different architectural designs during navigation and to determine if certain aspects of these designs may assist or hinder in the identification of landmarks for navigation and recall.

Hypotheses

Our broadest hypothesis was that buildings useful as landmarks would have a relatively high “saliency” factor, which is defined in navigational studies as a striking feature that stands out from the surrounding information terrain (Sorrows and Hirtle, 1999; Raubal and Winter, 2002; Caduff and Timpf, 2008). The term “saliency” has several related, yet distinct, meanings. In

perception and cognition studies, for example, a red flower against a background of green foliage has a high perceptual saliency, as does a loud noise in a quiet room, or an object in motion in an otherwise still environment (Koch and Ullman, 1985; Kerzel et al., 2011). It is important to note that the landmark saliency of architectural design is relative to the surrounding urban environment. In a city where nearly all of the buildings were geodesic domes, for example, an ordinary square apartment building might stand out as a striking landmark. If an object in a scene were also especially novel, the initial effect would be similar. For the purposes of this study we used currently existing Western metropolitan architecture as the urban background, and categorized buildings as salient or non-salient based on variations from this environmental norm along several different design features: relative height, footprint-shape, twist, and façade design. We aimed to control for the previously discussed overlapping influences of saliency and novelty during the creation of the virtual environment by: (a) ensuring that all the target buildings were similar to each other regarding how distinct they were from the surrounding background in major low-level features, e.g. luminance, color, general style; and (b) making sure that all the target buildings were generally plausible, and not exceptionally unusual.

Previous studies have noted in an anecdotal fashion that striking architectural design features in a building can be related to people's tendency to regard that building as a landmark (Lynch, 1960; Darken and Peterson, 2001). We expected that the concept of environmental/landmark saliency can help to explain this correlation between specific design features and the prominence of a building in human visual memory. During viewing of these more salient buildings, we would therefore expect to see increased posterior superior scalp EEG theta power, (i.e., synchronization) as one indicator of increased processing in response to relevant navigational cues. To support this overall perspective, we tested the following specific hypotheses:

Hypothesis H: Mean recall accuracy, user interaction (gaze duration), and neural signatures of spatial awareness (superior posterior theta power), will be heightened when participants gaze at salient buildings as compared to non-salient buildings.

- Hypothesis H1: Recall accuracy, user interaction (gaze duration), and neural signatures of spatial awareness (superior posterior theta power) will be heightened as a linear function of how salient a viewed building is, e.g. strongly salient, weakly salient, weakly non-salient, or strongly non-salient.

Upon testing for this aggregate main effect (Hypothesis H), as well as dose-response model of increasing salience (H1), we then drilled down to see if the same outcome measures tested above will be heightened in response to viewing buildings with specific salient architectural features compared with buildings that lack those specific architectural features.

- Hypothesis H2: Recall accuracy, user interaction (gaze duration), and neural signatures of spatial awareness (superior posterior theta power) will be heightened when participants gaze at buildings that are *taller than surrounding buildings*.

- Hypothesis H3: Recall accuracy, user interaction (gaze duration), and neural signatures of spatial awareness (superior posterior theta power) will be heightened when participants gaze at buildings containing *biologically inspired and natural elements*.
- Hypothesis H4: Recall accuracy, user interaction (gaze duration), and neural signatures of spatial awareness (superior posterior theta power) will be heightened when participants gaze at buildings containing a *twisted architectural façade or unusual footprints*.

MATERIALS AND METHODS

Drawing from the previous theoretical literature on landmark identification (Raubal and Winter, 2002; Klippel and Winter, 2005; Duckham et al., 2010; Peters et al., 2010), and focusing on the concept of visual saliency, we designed twelve different buildings to be tested in the virtual urban environment. These buildings were created using a parametric modeling algorithm in the Grasshopper software platform (www.grasshopper3d.com) so that their design features could be readily adjusted. The overall building contours were defined by their height, footprint-shape, and various aspects of their façade design. Color saliency was used as a way to make all of these buildings stand out to some extent from their background (Aziz and Mertsching, 2007). The buildings under investigation were presented using white colors, while all other background buildings were presented in gray and opaque colors. The background buildings were also designed to unobtrusively mirror typical Western metropolitan environments, thereby helping to reduce their structural saliency (Raubal and Winter, 2002). Designing the target buildings like this ensured that they shared numerous low-level perceptual features with each other, regardless of their level of architectural “saliency,” especially in contrast with the background urban scenery and architecture. This helped ensure that none of the target buildings would attract attention due to the “novelty” of low-level features, more so than the others. One important distinction to make is between saliency and novelty. In one sense, this can mean that new items in an environment can gain saliency over familiar items simply as a result of being new, somewhat regardless of their other features (Stirk and Underwood, 2007). In another sense, it can mean that an object with a very unique feature, e.g., a shape or pattern, that a person has never experienced before may make it highly salient to that person, even if the object’s other low-level features are not distinctive or salient in the traditional sense (Underwood et al., 2009). We attempted to control for novelty while manipulating saliency.

For the height variable, we were primarily interested in testing relative height compared to the environmental background, so we also included a contextual height difference variable, which indicates whether a building is taller or the same height compared to surrounding structures. For the footprint variable we used a square, a triangle, a pentagon, and a circle. We included a “twist” variable in some designs, meaning that the building outline had a twisted form in the upward direction (z axis). Some of the buildings included a horizontal overhang. The

study incorporated five different façade patterns, including a design with dominant horizontal lines, one with strong vertical lines, one with a grid pattern, one with a biologically inspired abstract Voronoi pattern, and one that included natural elements such as integrated gardens. Based on the overall prior evidence about the effects of building size, shape deviation, and façade eccentricity on visual saliency (Itti and Koch, 2000, 2001; Raubal and Winter, 2002; Zhang et al., 2014) we intentionally created seven buildings with salient features in height, shape, or façade patterns, and created five non-salient buildings to evaluate the psychophysiological responses across these two categories (Table 1).

Virtual Reality Development













The creation of the virtual urban environment was carried out using Epic Games’ Unreal Engine (www.epicgames.com). Most of the modeling and UV-mapping took place within Autodesk Maya. Parametric modeling for building exteriors was performed in Maya using Python and Mel scripting. Texturing was done procedurally using the Substance software platform, at a resolution of 4096×4096 . The Unreal Engine uses blueprint scripting, which can allow for a quick learning curve on the part of researchers and designers who may want to expand or replicate our work. All of the front-end interaction and user interactivity in our environment also leverages the blueprint platform.

The test buildings were integrated into an interactive and user-friendly VR environment simulating a standard North American urban exterior. We set the camera position at 1.70 m above the street (corresponding to average human eye-height). The environment was presented to study participants using an Oculus Rift head-mounted display (HMD) through a gaming laptop with a resolution of $1,920 \times 1,200$ pixels. The Oculus HMD provides a 100-degree horizontal field view with 75 Hz refresh rate, and can be adjusted for participants with different interpupillary distances. Some examples of two-dimensional screen captures from the virtual environment are shown in Figures 1A–D.

Participants

Twenty-nine healthy human adults with normal or corrected-to-normal vision were recruited for the study by word of mouth and broadly distributed email announcements within various departments of the academic institution. After exclusion, (see “Data Exclusion” below), data from twenty-one participants were analyzed. The participants’ ages ranged from 18 to 55 years ($M = 27.65$, $SD = 10.04$). The majority of the participants were university students ($n = 19$), and the rest were academic faculty members. All of the participants were associated with the University of Houston, representing the departments in Engineering, Biological Sciences, Humanities, Economics, and Computer Sciences. All participants gave informed written consent before participating in the experiment and were compensated with a \$25 gift card at the end of the study. The participants represented various national backgrounds (U.S., India, and Mexico were most common), and had a variety of ethnic backgrounds: nine reported as Asian, seven as Latino and/or Hispanic, one as Middle Eastern, and four as White (Non-Hispanic). Eight of the participants reported as women, thirteen

TABLE 1 | Landmark buildings included in the experiment were coded and described according to a basic set of design variables including height, footprint, and façade pattern.

												
	B01	B02	B03	B04	B05	B06	B07	B08	B09	B10	B11	B12
Height	SA	SA	SA	LO	SA	HI	SA	SA	SA	SA	SA	SA
Shape	RE	RE	RE	TR	RE	CI	RE	RE	RE	RE	RE	PE
Façade	HO	HO	GR	GR	HO	GR	VE	NA	GR	TW	VO	TW
	Non-Salient Design						Salient Design					

SA, Same level with neighbor buildings; LO, Lower than neighbor buildings; HI, Higher than neighbor buildings; RE, Rectangular; TR, Triangle; CI, Circle; PE, Pentagon; HO, Horizontal lines pattern; GR, Grid lines pattern; VE, Vertical lines pattern; NA, Nature feature included in the façade; TW, Twisted façade type; VO, Voronoi pattern.



as men, and none as other. The participants were also asked to report their sleep patterns and rate their mental fatigue level ($M = 3.95$, on scale from 1, low fatigue to 10, high fatigue) at the beginning of the experiment. Thirty-three percent reported getting seven or more hours of sleep the previous night, 29% reported between 6 and 7 h of sleep, and 38% reported 4 to 6 h of sleep. Participants were free of current neurological conditions,

abstained from any psychoactive substances prior to the study (except for two who had recently consumed caffeine, within 1 h), and only two participants were actively taking psychological medication. Participants were asked in a post-experiment survey to indicate their level of stress or discomfort with the virtual reality experience and the EEG headset, (1=comfortable with no stress; 10 = uncomfortable and stressful), and to report how

realistic the virtual environment seemed (1=not similar; 10=very similar). Participants indicated having no familiarity with virtual reality technology, and they reported that the VR system ($M = 4.083$, $SD = 2.465$) and EEG headset ($M = 3.75$; $SD = 2.156$) were not uncomfortable or stressful. Participants found the VR environment realistic ($M = 6.33$; $SD = 1.66$).

Procedure and Tasks

The experiments were conducted in a laboratory setting at the University of Houston. The study protocol was approved by the University of Houston's Institutional Review Board to protect the privacy and safety of the participants. After filling out the consent form and demographic forms, participants were fitted with the biometric sensors (described in more detail below). The initial stages of the experiment took place without the use of a VR headset. To establish baseline biometric data, the participants were asked to sit quietly facing a blank computer monitor for 1 min. Once this baseline data was collected, the participants donned the VR headset and entered the virtual environment. They were given an initial 5-min period in the VR to become accustomed with the navigation tools and to explore the platform.

The participants were then asked to travel along two different routes through the virtual city (with a brief break in between), and both routes included the twelve "target" buildings but in different locations. Participants "moved" through the virtual environment by using a Microsoft Xbox controller button-press to advance at their own pace along pre-specified locations on the path. Immediately after completing each route, the participants were asked to identify the general region/location of five different target buildings (as a result, there were a total of 10 items in the memory recall assessment). At the end of the experiment, the participants filled out an exit survey to provide additional feedback about their reactions to the study setting, the virtual environment, and the experimental design.

Signal-Derived Metrics

To obtain measurements of their biometric responses, the participants were instrumented with a non-invasive electroencephalography (EEG) cap to record electrical activity in their brains, and appropriately-placed sensors to record eye muscle movements (electro-oculography, EOG), heart activity (electrocardiogram, ECG, not analyzed for this experiment), skin conductance (galvanic sensor response, GSR, not analyzed for this experiment), and head motion (tri-axial head accelerometer, not analyzed for this experiment) (Figure 2A). All of this data was recorded at 500 Hz and synchronized using the 64-channel ActiCHamp module (Brain Products GmbH, Germany) with Ag/AgCl active electrodes. A total of 63 electrodes were used (57 for EEG, 4 for EOG, and 2 for ECG). Only EEG data was analyzed in this study, although EOG data was included during ICA to help with IC-based artifact elimination. We focused our data analysis on the Left Posterior Superior scalp region (electrodes CP3, CP1, P1, P3, P5, and PO3); the Right Posterior Superior scalp region (electrodes CP2, CP4, P6, P4, P2, and PO4); the Left Parietal Inferior region (electrodes P8 and PO8); and the Right Posterior Inferior scalp region (electrodes P7 and PO7). These regions are shown in Figure 2B. The two Posterior

Inferior regions are displayed as slightly off the scalp since in their true position they are not easily viewed from a top-down viewpoint. Lab-Streaming Layer (LSL), a multi-modal data collection software, was used to synchronize all modes of data (Kothe, 2014).

EEG Data Processing

The EEG data were analyzed using the EEGLAB software package (Delorme and Makeig, 2004). Raw ".xdf" data files were imported at their original sampling rate of 500 Hz, low-pass filtered at 100 Hz, high-pass filtered at 0.1 Hz, and then run through the "Cleanline" algorithm, which selectively filters out the 60 Hz power-line noise using an adaptive frequency-domain (multi-taper) regression technique. The chronological segment of the EEG data relevant to the experimental procedure was then extracted and run through the PREP Pipeline (Bigdely-Shamlo et al., 2015), which is a robust re-referencing method that minimizes the bias introduced by referencing using noisy channels. Artifact-laden time-windows and channels were identified using manual inspection and deleted, by research team members blind to the order of the trial conditions. These trimmed datasets were then each run through Independent Component Analysis (ICA), using the extended Infomax variation as implemented in the EEGLAB package. ICA is a form of second-order blind identification, which performs spatial unmixing to identify underlying sources of signal within an under-specified space (Onton et al., 2006), such as EEG. Independent Components indicative of eye-movements and muscle activity were identified and deleted, as well as time windows with gross artifact contamination.

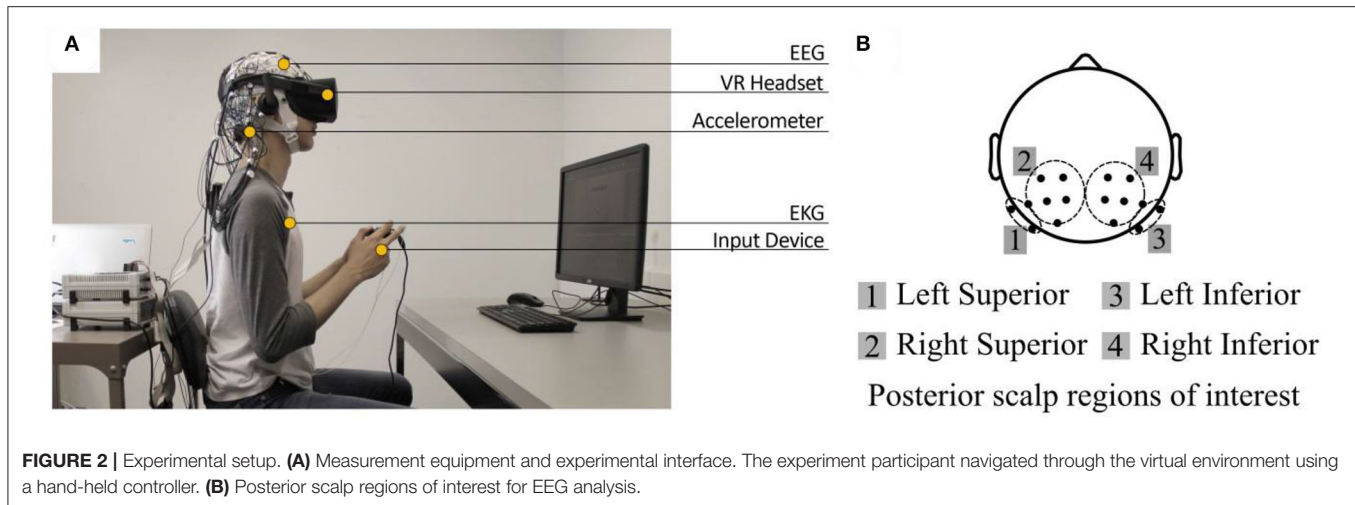
Data Exclusion

Several participants' data were excluded from final analyses. Final EEG analyses included 21 participants: two were excluded due to extremely high theta power activity (> 3 St.Dev. of the overall mean theta Event-Related Desynchronization); three were excluded due to low EEG recording quality ($> 33\%$ of channels exceeding $100 \mu V$ for $> 50\%$ of time series); and three were excluded because they did not gaze at a sufficient number of target buildings for longer than 1.5 s each (1.5 s was the *a priori* duration threshold used to include time-points for analysis).

Data Analysis

Gaze Time

We used recorded screen-capture video from the experiment to identify the periods of time when the participants observed the buildings that were being tested. The start and end of these time periods were determined based on the content of the VR display, which itself was linked to the direction of the participants' gaze. One researcher marked all of the screen data using video-editing software (Camtasia), and then another researcher reviewed the marked segments for accuracy. These gaze event markers were exported for analysis, and to be automatically imported into the EEG files as events. We analyzed the total amount of time that participants spent looking at each building (gaze time), as well as the number of instances in which their gaze returned to the building (gaze count). These measurements of gaze times and



counts were averaged across both of the routes that were tested (route 1 and route 2) (**Supplementary Table 1**).

We used linear models in R [using the `lm()` function] to determine if participants looked for longer periods of time at certain types of buildings, with Overall Building Saliency and various building design features as the predictor variables in separate tests for each hypothesis, and average Gaze Time as the outcome variable. Paired *t*-test comparisons were then computed, using a tukey method for *p*-value adjustments. Gaze times of 0 s were excluded from the statistical analyses (< 1% of all cases).

Location Recall Accuracy

After completing each route, participants were presented with a series of images of target buildings, alongside a map of the route that they had just traversed. The map was divided into four zones (A, B, C, and D).

Participants were asked to identify the zone in which each test building appeared. Only ten of the buildings were tested for visual memory (omitting the B01 and the B09 buildings). In the analysis of the visual memory test data, we used linear models to determine if participants more accurately remembered the zone locations of certain types of buildings. Overall Building Saliency and various building design features were used as the predictor variables, and accuracy of visual memory was used as the outcome variable.

EEG Theta Power ERD/ERS as Participants Observed the Buildings

The mean Event-Related Synchronization/Desynchronization (ERD/ERS) was calculated across all of the EEG electrodes in each scalp region, averaged across the theta-band frequencies (4–7 Hz), for the combined duration of the time periods during which a participant was viewing a particular test building. Welch's method of overlapped segment averaging was used as an estimator of power spectral density, as implemented in EEGLAB's "spectopo" function. EEG power spectral density values were converted to $\mu V^2/Hz$ units, so that all theta power values would

be positive. The ERD/ERS value was then calculated as:

$$\% \text{ ERD/ERS} = \frac{A - R}{R} \times 100$$

where *R* indicates baseline reference data and *A* indicates the theta power value associated with the time in which a participant viewed a particular test building (Pfurtscheller and Lopes da Silva, 1999). Since the viewing time differed for each participant and for each building, the amount of data that was fed into each ERD calculation varied. These ERD values were entered into a separate mixed multilevel model for each planned comparison, with scalp region of interest and building design features as fixed effects, participant ID as a random effect, and mean theta ERD as the dependent outcome variable. Paired *t*-test comparisons were then computed, using a Tukey method for *p*-value adjustments.

RESULTS

Each hypothesis was tested using a separate linear model, using the building categories relevant to the variables being evaluated: Non-salient vs. Salient buildings, relative height, and façade designs (Nature vs. non-nature, and Twist vs. non-twist).

Gaze Time Looking at Buildings

A high building Saliency as defined by **Table 1** was associated with longer Gaze Time in the two-category (i.e. Salient and Non-salient) comparison ($F_{(1, 390)} = 27.60, p < 0.001$). We also found associations between several architectural features and the average Gaze Time. The Vertical and Voronoi façade patterns, in particular, attracted significantly longer Gaze Times compared to other façade types (all $|t|s > |-2.84|, ps < 0.03$). We observed longer Gaze Times for the buildings with the salient feature of Twist façade design, compared with non-twist façade building designs ($F_{(1, 54)} = 26.74, p < 0.001$). Longer Gaze Times were found for the Nature façade design compared with non-Nature designs ($F_{(1, 54)} = 14.67, p < 0.001$); and longer Gaze Times for contextually the taller building compared to those with the

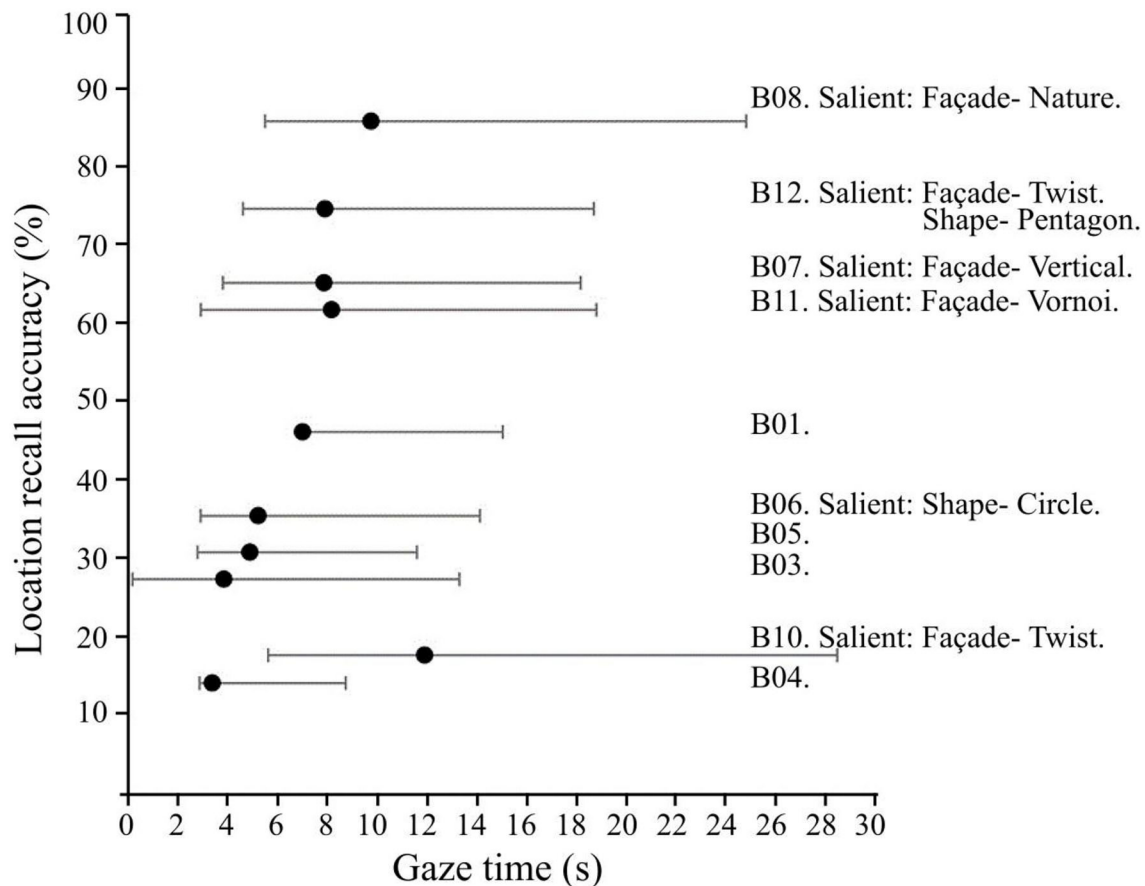


FIGURE 3 | Location recall accuracy compared with gaze time distributions, with error bars showing the 25th and 75th percentile.

same height level with the neighbor buildings ($F_{(2, 109)} = 6.74, p < 0.001$).

Location Recall Accuracy

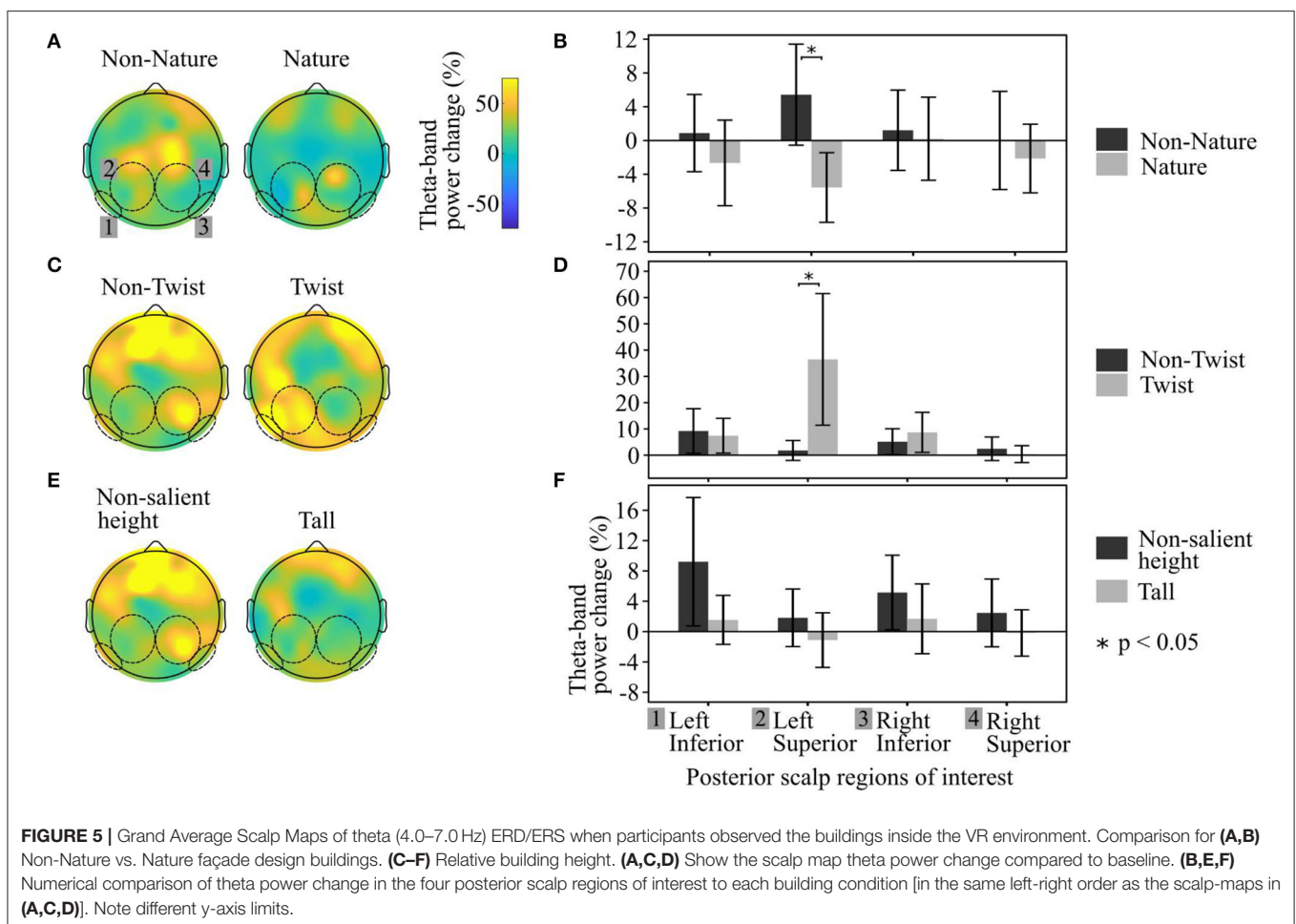
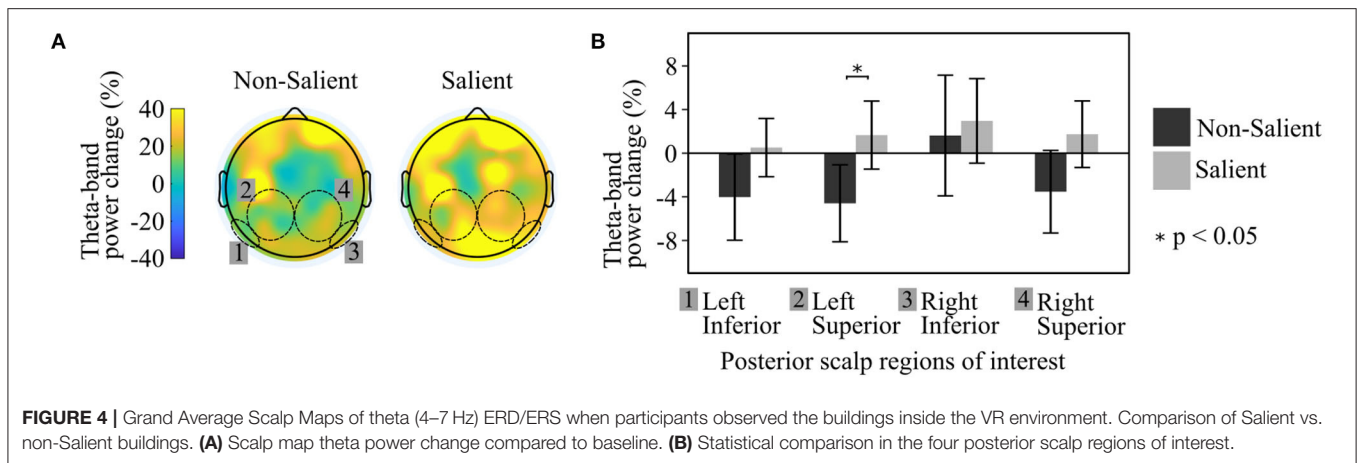
Building Saliency was associated with more accurate visual memory in the two-category comparison ($F_{(1, 283)} = 22.26, p < 0.001$) of Non-Salient vs. Salient buildings (Figure 3). Similar to the gaze-tracking results, more accurate visual memory was associated with the Vertical and Voronoi façade patterns (both $t_s > 5.2, p < 0.001$), and with the Nature façade design ($F_{(1, 56)} = 30.19, p < 0.001$), compared with their respective matched non-salient buildings. In contrast to the gaze-time results, no significant overall association was found between visual memory accuracy and the Twist façade design feature. While the building B12 (Pentagon-footprint, Twist-façade design) obtained high location recall accuracy in the visual memory tests, building B10 (Square footprint, Twist-façade design) obtained low location recall accuracy. No contextually tall buildings were tested for visual memory, so we lack the data to evaluate the association between visual memory accuracy and contextual height. Buildings B02 and B09 were not included in the recall task, simply as part of the effort to keep the experiment duration as short as possible.

EEG Theta Power ERD/ERS as Participants Observed the Buildings

The multilevel linear model for building Saliency indicated a significant association with theta activation ($F_{(1, 140)} = 8.343, p = 0.005$), i.e., Event-related synchronization (ERS), thus supporting the main effect hypothesis H (Figure 4). Individual *t*-tests comparing these conditions within each of the four scalp regions of interest revealed that theta power changes over baseline, for more salient buildings, were significantly greater in Left Posterior Superior region ($t = -2.08, p = 0.04$), and were trending significantly for the Right Posterior Superior region ($t = -1.75, p = 0.08$) (Figure 4). As a result of these findings, our hypotheses were tested specifically over theta power changes (ERD/ERS) in the Left Posterior Superior region.

Individual *t*-tests comparing Twist to Non-Twist façade design buildings within each of the four scalp regions revealed that theta power changes over baseline for the Twist-façade condition were significantly greater than for the Non-Twisted condition in only the Left Posterior Superior region ($t = 2.08, p = 0.01$) (Figures 5C,D). The findings in regard to Twisted buildings support Hypothesis H4.

The comparison of the Nature façade design with a corresponding “Non-Nature” building yielded an activation



difference in the Left Posterior Superior region ($t = -2.35$, $p = 0.02$). However, the direction of this effect was the opposite of the Twist and overall Saliency comparisons—when viewing the Nature condition (which was recalled more accurately and gazed at longer than most other buildings), participants' Left

Posterior Superior regions not only had lower theta power change than the control condition, the Left Posterior Superior theta power was actually *lower* in comparison with the rest/reference window. Conversely, during the Twist façade design and overall Saliency comparisons, the participants' Left Superior Posterior

regions had a *higher* theta power over baseline in comparison with their corresponding buildings (**Figures 5A,B**). Both findings about façades with natural elements and a twist-design façade are salient features that show statistically significant differences in theta power synchronization / desynchronization, and therefore support our proposed hypothesis H3.

The pairwise tests for the relative building-height features did not yield significant associations with the data in regard to any of the EEG regions, although the buildings with the same height as their surrounding were nearly statistically significantly greater in theta power change in the LPI region ($t = 1.93$, $p > 0.06$).

Thus, Hypotheses H2 was not supported by the EEG data.

DISCUSSION AND CONCLUSIONS

In this study behavioral data, recall accuracy, and EEG data were combined to analyze the effect of architectural features on visual landmark saliency. Our primary Hypothesis H1 was that recall accuracy, user interaction (gaze duration), and neural signatures of spatial awareness in the form of posterior lateral EEG theta power, would be heightened when participants gazed at salient buildings as compared to non-salient buildings. The EEG results support this main effect hypothesis, with the caveat that the progression in posterior theta power change, from less-salient to more-salient buildings, did not seem to act in a linear fashion. Highly salient buildings were also found to attract a greater duration of gazing time on the part of the study participants, and the spatial location of these buildings was more accurately recalled in short-term visual memory tests.

Expanding on our primary hypothesis, we aimed to show that the same behavioral and neural features mentioned above would be heightened when viewing buildings with a specific salient feature, as compared with buildings that lack the corresponding architectural feature. The support for this hypothesis was mixed. EEG results indicated significant differences in neural activation measured at left superior posterior regions of the scalp (**Figures 5A–C**) for buildings with a Twist-façade design and for buildings in which the façade designs contained natural elements. However, the results for the nature-containing buildings were opposite of what we expected, correlating with a decrease in theta activation (i.e., ERD) at the target scalp area, rather than an increase in theta activation (i.e., ERS). EEG tests for other types of design features did not show significant differences, in some cases diverging from the results found in the gaze duration and recall data. For example, participants gave seemingly more attentional resources to Vertical and Voronoi façade patterns than Horizontal or Grid patterns as evidenced by the gaze and recall results, yet the electrophysiology data did not correspond.

These results support the hypothesis that when people gaze at “interesting” buildings that stand out from the surrounding environment, scalp EEG theta power above posterior parietal cortex increases. Further experiments and analyses, including

source-localization steps and higher spatial resolution modalities (e.g., fMRI, MEG), would more conclusively test the retrosplenial cortex’s causal role in making some landmarks more effective than others. However, the current experiment provides support and justification for that work. It also suggests that effective landmarks serve as a focus of visual attention, and their location therefore persists more accurately in short-term visual memory. Consistent with the salience model proposed by Raubal and Winter (2002) and the predictive model proposed by Zhang et al. (2014), our findings show that building shape and height can influence the visual attraction of landmarks. However, several open questions remain, which should motivate future work using these and similar methods. For example, what combination of architectural design features are causing this saliency response, and what role does short-term memory play? Do more salient buildings lend themselves to successful use as landmarks simply by virtue of the increased gaze duration, i.e., as a result of spending more time “taking in” that location in the path/environment? Or do these landmarks that stand out visually recruit more attentional resources during perception? Do other regions involved in spatial navigation and memory, such as the hippocampus, show more effective connectivity with the RSC while during memory encoding, or do more salient buildings instead function better as landmarks by boosting simply recall? Or do they lead to improvements in both encoding and recall? These questions can be answered with further experiments building on this research platform.

Due to the fact that in this experiment participants did not engage in ambulatory locomotion (they were seated), and within the virtual environment they did not have the option of truly free navigation (they simply moved among pre-determined locations along the path, i.e., via “teleporting”), we can only make limited generalizations about wayfinding based on these results alone. Nevertheless, this is an important next step in understanding the psychological role of design decisions and architectural features of an environment. While the minimization of movement-related artifacts helps strengthen the quality of combined VR-EEG research, ongoing advances in mobile EEG technology will enable researchers to test the validity of virtual navigation studies. It is also important to note that locomotion may not be as necessary for visual attention as it is for wayfinding, and since this was not a wayfinding study, the absence of locomotion and vestibular input may not a major flaw. Future work would need to test whether free ambulatory locomotion would disproportionately affect visual attention and retrosplenial connectivity with other spatial memory and navigation centers more for distinctive buildings than non-distinctive buildings, perhaps by offering more gradations of perspective and longer, more dynamic periods in which the buildings are partially visible.

Regarding specific architectural features, our finding that the twisted façade was associated with significant activation in the left superior parietal areas of the scalp indicate that some architectural features inspire wayfinding-related activation (**Figures 5C,D**). This is corroborated by the finding that the

twisted buildings tended to become a strong focus of our participants' visual attention (**Figure 3**), even when the location recall accuracy differed from both buildings that contained this salient feature. Based on an evaluation of recall responses, it seems that participants often confused the two Twist-façade buildings with each other during the recall step, which is the likely cause of the discrepancy between the EEG theta results and the location recall accuracy results.

Our finding that increased recall accuracy and gaze time generally accompanied increases in Left Posterior Superior theta power change for *a-priori*-defined salient buildings, but an opposite pattern for the building with "green" or "nature elements" incorporated, suggests that different kinds of landmark saliency may act on our navigational attention systems and even our default mode network, differently. Further neural research in this domain should investigate connectivity changes in response to different landmarks, as no brain region acts alone.

The findings of this research have implications for urban planners and metropolitan authorities in their goal of developing better wayfinding systems. Previous studies have shown that landmarks are a crucial element in pedestrian and vehicle navigation (Lynch, 1960; May et al., 2003; Goodman et al., 2005; Reagan and Baldwin, 2006; Millonig and Schechtner, 2007; Stark et al., 2007; Hile et al., 2008). However, landmarks are rarely included in route descriptions and other urban wayfinding literature due to the problem of determining what environmental features should be identified and promoted as landmarks (Elias, 2003; Duckham et al., 2010; Peters et al., 2010; Winter et al., 2008). A more robust understanding of what urban features are most useful for diverse participants in landmarking—based on meaningful isolation of design variables and scientific data-collection—has the potential to help solve this dilemma and establish more reliable guidelines for the selection of urban landmarks in wayfinding communications.

DATA AVAILABILITY STATEMENT

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

REFERENCES

- Acredolo, L. P. (1977). Developmental changes in the ability to coordinate perspectives of a large-scale space. *Dev. Psychol.* 13:1. doi: 10.1037/0012-1649.13.1.1
- Auger, S. D., Mullally, S. L., and Maguire, E. A. (2012). Retrosplenial cortex codes for permanent landmarks. *PLoS ONE* 7:e43620. doi: 10.1371/journal.pone.0043620
- Aziz, M. Z., and Mertsching, B. (2007). "Color saliency and inhibition using static and dynamic scenes in region based visual attention," in *International Workshop on Attention in Cognitive Systems*, (Heidelberg: Springer), 234–250. doi: 10.1007/978-3-540-77343-6_15
- Banaschewski, T., and Brandeis, D. (2007). Annotation: what electrical brain activity tells us about brain function that other techniques cannot tell

ETHICS STATEMENT

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

AUTHOR CONTRIBUTIONS

SK developed the experimental design and oversaw the collection of data. JC-G collected the data. JC-G, JR, and SK contributed to the behavioral and physiological data analysis, the interpretation of the results, and wrote the manuscript. JR performed statistical analyses. All authors contributed to the article and approved the submitted version.

FUNDING

This research was possible with seed funding from the Gerald D. Hines College of Architecture and Design at the University of Houston, and seed funding from the Cornell Center for Social Sciences at Cornell University.

ACKNOWLEDGMENTS

The authors thank Joshua Smith for his help in developing the 3D models and finalizing the virtual reality files, as well as Pamela Banner for her assistance in data-collection and measurement device integration. The authors also thank the team at the Design and Augmented Intelligence Lab at Cornell University, including Sina Pourjabbar, Julia Kan, Mi Rae Kim, and Vidushi Tripathi for curating and annotating the behavioral and neural data. Dr. Jose Luis Contreras-Vidal assisted significantly in the research by providing all the physiological measurement devices and sensors that were used in the study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnhum.2020.584385/full#supplementary-material>

us—a child psychiatric perspective. *J. Child Psychol. Psychiatry* 48, 415–435. doi: 10.1111/j.1469-7610.2006.01681.x

- Belchior, H., Lopes-dos-Santos, V., Tort, A. B. L., and Ribeiro, S. (2014). Increase in hippocampal theta oscillations during spatial decision making. *Hippocampus* 24, 693–702. doi: 10.1002/hipo.22260
- Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.-M., and Robbins, K. A. (2015). The PREP pipeline: standardized preprocessing for large-scale EEG analysis. *Front Neuroinf.* 9:16. doi: 10.3389/fninf.2015.00016
- Bischof, W. F., and Boulanger, P. (2003). Spatial navigation in virtual reality environments: an EEG analysis. *Cyber Psychol. Behav.* 6, 487–495. doi: 10.1089/109493103769710514
- Blades, M., Lippa, Y., Golledge, R. G., Jacobson, R. D., and Kitchin, R. M. (2002). The effect of spatial tasks on visually impaired peoples' wayfinding abilities. *J. Vis. Impairment Blindness* 96, 407–419. doi: 10.1177/0145482X0209600604

- Bradley, N. A., and Dunlop, M. D. (2005). An experimental investigation into wayfinding directions for visually impaired people. *Pers. Ubiquitous Comput.* 9, 395–403. doi: 10.1007/s00779-005-0350-y
- Burns, P. C. (1997). *Navigation and the ageing driver*. (Doctoral dissertation), Loughborough University, Loughborough, UK.
- Buzsáki, G. (2005). Theta rhythm of navigation: Link between path integration and landmark navigation, episodic and semantic memory. *Hippocampus* 15, 827–840. doi: 10.1002/hipo.20113
- Caduff, D., and Timpf, S. (2008). On the assessment of landmark salience for human navigation. *Cogn. Process* 9, 249–267. doi: 10.1007/s10339-007-0199-2
- Chang, K. T., and Zheng, M. C. (2016). “Study on landmark design of wayfinding map in taipei main station,” in *Advances in Ergonomics in Design*, eds. F. Revelo and M. Soares (Cham: Springer), 571–581. doi: 10.1007/978-3-319-41983-1_52
- Claramunt, C., and Winter, S. (2007). Structural salience of elements of the city. *Environ. Plan. B Plan. Des.* 34:1030. doi: 10.1068/b32099
- Clark, H. H. (1996). *Using Language*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511620539
- Clay, F., Howett, D., FitzGerald, J., Fletcher, P., Chan, D., and Price, A. (2020). Use of immersive virtual reality in the assessment and treatment of alzheimer’s disease: a systematic review. *J. Alzheimer’s Dis.* 75, 23–43. doi: 10.3233/JAD-191218
- Cornwell, B. R., Johnson, L. L., Holroyd, T., Carver, F. W., and Grillon, C. (2008). Human hippocampal and parahippocampal theta during goal-directed spatial navigation predicts performance on a virtual morris water maze. *J. Neurosci.* 28, 5983–5990. doi: 10.1523/JNEUROSCI.5001-07.2008
- Darken, R. P., and Peterson, B. (2001). “Spatial orientation, wayfinding, and representation,” in *Handbook of Virtual Environment Technology*, ed. K. Stanney (Mahwah, NJ: Lawrence Erlbaum Associates), 493–518.
- De Tommaso, M., Ricci, K., Delussi, M., Montemurno, A., Vecchio, E., Brunetti, A., and Bevilacqua, V. (2016). Testing a novel method for improving wayfinding by means of a P3b virtual reality visual paradigm in normal aging. *Springerplus* 5:1297. doi: 10.1186/s40064-016-2978-7
- Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009
- Dhindsa, K., Drobinin, V., King, J., Hall, G. B., Burgess, N., and Becker, S. (2014). Examining the role of the temporo-parietal network in memory, imagery, and viewpoint transformations. *Front. Hum. Neurosci.* 8:709. doi: 10.3389/fnhum.2014.00709
- Duckham, M., Winter, S., and Robinson, M. (2010). Including landmarks in routing instructions. *J. Loc. Based Serv.* 4, 28–52. doi: 10.1080/17489721003785602
- Egenhofer, M., Mark, D. (1995). “Naive geography,” in *Spatial Information Theory A Theoretical Basis for GIS. Volume 988 of Lecture Notes in Computer Science*, eds. A. U. Frank and W. Kuhn (Springer: Berlin), 1–15. doi: 10.1007/3-540-60392-1_1
- Elias, B. (2003). “Extracting landmarks with data mining methods,” in *Spatial Information Theory. Foundations of Geographic Information Science. COSIT 2003. Lecture Notes in Computer Science*, Vol. 2825, eds. W. Kuhn, M. F. Worboys and S. Timpf (Berlin, Heidelberg: Springer). doi: 10.1007/978-3-540-39923-0_25
- Epstein, R. A., and Vass, L. K. (2014). Neural systems for landmark-based wayfinding in humans. *Philos. Trans. R Soc. B* 369:20120533. doi: 10.1098/rstb.2012.0533
- Fewings, R. (2001). Wayfinding and airport terminal design. *J. Navig.* 54, 177–184. doi: 10.1017/S037346301001369
- Fischer, L. F., Mojica Soto-Albors, R., Buck, F., and Harnett, M. T. (2020). Representation of visual landmarks in retrosplenial cortex. *Elife* 9:e51458. doi: 10.7554/eLife.51458
- Gazova, I., Laczó J., Rubinova, E., Mokrisova, I., Hyncicova, E., Andel, R., et al. (2013). Spatial navigation in young versus older adults. *Front. Aging Neurosci.* 5:94. doi: 10.3389/fnagi.2013.00094
- Goodman, J., Brewster, S. A., and Gray, P. (2005). How can we best use landmarks to support older people in navigation? *Behav. Inform. Technol.* 24, 3–20. doi: 10.1080/01449290512331319021
- Grabler, F., Agrawala, M., Sumner, R. W., and Pauly, M. (2008). Automatic generation of tourist maps. *ACM Trans. Graphics* 27, 1–11. doi: 10.1145/1399504.1360699
- Hanslmayr, S., Volberg, G., Wimber, M., Raabe, M., Greenlee, M. W., and Bäuml, K.-H. T. (2011). The relationship between brain oscillations and BOLD signal during memory formation: a combined EEG–fMRI study. *J. Neurosci.* 31, 15674–15680. doi: 10.1523/JNEUROSCI.3140-11.2011
- He, G., Ishikawa, T., and Takemiya, M. (2015). Collaborative navigation in an unfamiliar environment with people having different spatial aptitudes. *Spatial Cogn. Comput.* 15, 285–307. doi: 10.1080/13875868.2015.1072537
- Heft, H., and Nasar, J. L. (2000). Evaluating environmental scenes using dynamic versus static displays. *Environ. Behav.* 32, 301–322. doi: 10.1177/0013916500323001
- Hile, H., Vedantham, R., Cuellar, G., Liu, A., Gelfand, N., Grzeszczuk, R., et al. (2008). “Landmark-based pedestrian navigation from collections of geotagged photos,” in *Proceedings of the 7th International Conference on Mobile and Ubiquitous Multimedia*, (Umeå: Association for Computing Machinery), 145–152. doi: 10.1145/1543137.1543167
- Howard, I. P., and Templeton, W. B. (1966). *Human Spatial Orientation*. London: John Wiley & Sons.
- Howett, D., Castegnaro, A., Krzywicka, K., Hagman, J., Marchment, D., Henson, R., et al. (2019). Differentiation of mild cognitive impairment using an entorhinal cortex-based test of virtual reality navigation. *Brain* 142, 1751–1766. doi: 10.1093/brain/awz116
- Itti, L., and Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Res.* 40, 1489–1506. doi: 10.1016/S0042-6989(99)00163-7
- Itti, L., and Koch, C. (2001). Computational modelling of visual attention. *Nat. Rev. Neurosci.* 2, 194–203. doi: 10.1038/35058500
- Jaiswal, N., Ray, W., and Slobounov, S. (2010). Encoding of visual-spatial information in working memory requires more cerebral efforts than retrieval: evidence from an EEG and virtual reality study. *Brain Res.* 1347, 80–89. doi: 10.1016/j.brainres.2010.05.086
- Jansen-Osmann, P., Schmid, J., and Heil, M. (2007). Wayfinding behavior and spatial knowledge of adults and children in a virtual environment: the role of the environmental structure. *Swiss J. Psychol.* 66, 41–50. doi: 10.1024/1421-0185.66.1.41
- Jeffery, K. (2019). Urban architecture: a cognitive neuroscience perspective. *Des. J.* 22, 853–872. doi: 10.1080/14606925.2019.1662666
- Jiang, C. F., and Li, Y. S. (2009). “Development and verification of a VR platform to evaluate wayfinding abilities,” in *Proceedings of the 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, ed. B. He (Minneapolis: IEEE), 2396–2399.
- Joseph, A. (2006). *The Impact of the Environment on Infections in Healthcare Facilities*. Concord, CA: Center for Health Design.
- Julian, J., and Epstein, R. (2013). The landmark expansion effect: navigational relevance influences memory of object size. *J. Vision*, 13, 49. doi: 10.1167/13.9.49
- Julian, J. B., Keinath, A. T., Frazzetta, G., and Epstein, R. A. (2018). Human entorhinal cortex represents visual space using a boundary-anchored grid. *Nat. Neurosci.* 21, 191–194. doi: 10.1038/s41593-017-0049-1
- Kalantari, S., and Neo, J. R. J. (2020). Virtual environments for design research: lessons learned from use of fully immersive virtual reality in interior design research. *J. Interior Design.* 45, 27–42. doi: 10.1111/joid.12171
- Kalantari, S., and Shepley, M. (2020). Psychological and social impacts of high-rise buildings: a review of the post-occupancy evaluation literature. *Hous. Stud.* 1–30. doi: 10.1080/02673037.2020.1752630
- Kerzel, D., Schönhammer, J., Burra, N., Born, S., and Souto, D. (2011). Saliency changes appearance. *PLoS ONE* 6:e28292. doi: 10.1371/journal.pone.0028292
- Klippel, A., and Winter, S. (2005). “Structural salience of landmarks for route directions,” in *International Conference on Spatial Information Theory*, (Springer: Berlin, Heidelberg), 347–362. doi: 10.1007/11556114_22
- Knöll, M., Neuheuser, K., Cleff, T., and Rudolph-Cleff, A. (2018). A tool to predict perceived urban stress in open public spaces. *Environ. Plan. B* 45, 797–813. doi: 10.1177/0265813516686971
- Koch, C., and Ullman, S. (1985). *Shifts in selective visual attention: towards the underlying neural circuitry*. *Hum. Neurobiol.* 4, 219–227.
- Koike, B. D. V., Farias, K. S., Billwiller, F., Almeida-Filho, D., Libourel, P.-A., Tiran-Cappello, A., et al. (2017). Electrophysiological evidence that the retrosplenial cortex displays a strong and specific activation phased with

- hippocampal theta during paradoxical (REM) sleep. *J. Neurosci.* 37, 8003–8013. doi: 10.1523/JNEUROSCI.0026-17.2017
- Kondo, M., Fluehr, J., McKeon, T., and Branas, C. (2018). Urban green space and its impact on human health. *Int. J. Environ. Res. Public Health* 15:445. doi: 10.3390/ijerph15030445
- Kothe, C. (2014). *Lab Streaming Layer (LSL)*. Available online at: <https://github.com/scn/labstreaminglayer> (accessed October 26, 2020).
- Kuliga, S. F., Thrash, T., Dalton, R. C., and Hölscher, C. (2015). Virtual reality as an empirical research tool: exploring user experience in a real building and a corresponding virtual model. *Comput. Environ. Urban Syst.* 54, 363–375. doi: 10.1016/j.compenvurbsys.2015.09.006
- Lee, S. A., Shusterman, A., and Spelke, E. S. (2006). Reorientation and landmark-guided search by young children: evidence for two systems. *Psychol. Sci.* 17, 577–582. doi: 10.1111/j.1467-9280.2006.01747.x
- Lin, C.-T., Chiu, T.-C., and Gramann, K. (2015). EEG correlates of spatial orientation in the human retrosplenial complex. *Neuroimage* 120, 123–132. doi: 10.1016/j.neuroimage.2015.07.009
- Lovelace, K. L., Hegarty, M., and Montello, D. R. (1999). “Elements of good route directions in familiar and unfamiliar environments,” in *Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science. COSIT 1999. Lecture Notes in Computer Science*, Vol. 1661, eds. C. Freksa, D. M. Mark (Berlin, Heidelberg: Springer), 65–82. doi: 10.1007/3-540-48384-5_5
- Lynch, K. (1960). *The Image of the City*. Cambridge, MA: MIT Press.
- Maass, W., and Schmauks, D. (1998). MOSES: an example for the modelling of spatial services by means of route description system. *Zeitschrift Semiotik* 20, 105–118.
- Makransky, G., Borre-Gude, S., and Mayer, R. E. (2019). Motivational and cognitive benefits of training in immersive virtual reality based on multiple assessments. *J. Comput. Assist. Learn.* 35, 691–707. doi: 10.1111/jcal.12375
- May, A. J., Ross, T., Bayer, S. H., and Tarkiainen, M. J. (2003). Pedestrian navigation aids: Information requirements and design implications. *Pers. Ubiquitous Comput.* 7, 331–338. doi: 10.1007/s00779-003-0248-5
- Michon, P. E., and Denis, M. (2001). “When and why are visual landmarks used in giving directions?,” in *Spatial Information Theory. COSIT 2001. Lecture Notes in Computer Science*, Vol. 2205, ed. D. R. Montello (Berlin, Heidelberg: Springer), 292e305. doi: 10.1007/3-540-45424-1_20
- Millonig, A., and Schechtner, K. (2007). Developing landmark-based pedestrian-navigation systems. *IEEE Trans. Intell. Transport. Syst.* 8, 43–49. doi: 10.1109/TITS.2006.889439
- Nasar, J. L. (1994). Urban design aesthetics: the evaluative qualities of building exteriors. *Environ. Behav.* 26, 377–401. doi: 10.1177/001391659402600305
- Nothegger, C., Winter, S., and Raubal, M. (2004). Selection of salient features for route directions. *Spatial Cogn. Comput.* 4, 113–136. doi: 10.1207/s15427633sc0402_1
- Nys, M., Gyselinck, V., Orriols, E., and Hickmann, M. (2015). Landmark and route knowledge in children's spatial representation of a virtual environment. *Front. Psychol.* 5:1522. doi: 10.3389/fpsyg.2014.01522
- Onton, J., Westerfield, M., Townsend, J., and Makeig, S. (2006). Imaging human EEG dynamics using independent component analysis. *Neurosci. Biobehav. Rev.* 30, 808–822. doi: 10.1016/j.neubiorev.2006.06.007
- Paller, K. A., and Wagner, A. D. (2002). Observing the transformation of experience into memory. *Trends Cogn. Sci.* 6, 93–102. doi: 10.1016/S1364-6613(00)01845-3
- Peters, D., Wu, Y., and Winter, S. (2010). “Testing landmark identification theories in virtual environments,” in *International Conference on Spatial Cognition*, (Springer: Berlin, Heidelberg), 54–69. doi: 10.1007/978-3-642-14749-4_8
- Pfurtscheller, G., and Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin. Neurophysiol.* 110, 1842–1857. doi: 10.1016/S1388-2457(99)00141-8
- Plancher, G., Tirard, A., Gyselinck, V., Nicolas, S., and Piolino, P. (2012). Using virtual reality to characterize episodic memory profiles in amnesic mild cognitive impairment and Alzheimer's disease: influence of active and passive encoding. *Neuropsychologia* 50, 592–602. doi: 10.1016/j.neuropsychologia.2011.12.013
- Pu, Y., Cornwell, B. R., Cheyne, D., and Johnson, B. W. (2017). The functional role of human right hippocampal/parahippocampal theta rhythm in environmental encoding during virtual spatial navigation. *Hum. Brain Mapp.* 38, 1347–1361. doi: 10.1002/hbm.23458
- Ramanoël, S., York, E., Le Petit, M., Lagrené K., Habas, C., and Arleo, A. (2019). Age-related differences in functional and structural connectivity in the spatial navigation brain network. *Front. Neural Circuits* 13:69. doi: 10.3389/fncir.2019.00069
- Raubal, M., and Winter, S. (2002). “Enriching wayfinding instructions with local landmarks,” in *International Conference on Geographic Information Science*, eds. M. J. Egenhofer and D. M. Mark (Berlin: Springer), 243–259. doi: 10.1007/3-540-45799-2_17
- Reagan, I., and Baldwin, C. L. (2006). Facilitating route memory with auditory route guidance systems. *J. Environ. Psychol.* 26, 146–155. doi: 10.1016/j.jenvp.2006.06.002
- Röser, F., Hamburger, K., Krumnack, A., and Knauff, M. (2012). The structural saliency of landmarks: results from an on-line study and a virtual environment experiment. *J. Spat. Sci.* 57, 37–50. doi: 10.1080/14498596.2012.686362
- Sadalla, E. K., Burroughs, W. J., and Staplin, L. J. (1980). Reference points in spatial cognition. *J. Exp. Psychol.* 6:516. doi: 10.1037/0278-7393.6.5.516
- Scheeringa, R., Petersson, K. M., Oostenveld, R., Norris, D. G., Hagoort, P., and Bastiaansen, M. C. M. (2009). Trial-by-trial coupling between EEG and BOLD identifies networks related to alpha and theta EEG power increases during working memory maintenance. *Neuroimage* 44, 1224–1238. doi: 10.1016/j.neuroimage.2008.08.041
- Scholz, S., Schneider, S. L., and Rose, M. (2017). Differential effects of ongoing EEG beta and theta power on memory formation. *PLoS ONE* 12:e0171913. doi: 10.1371/journal.pone.0171913
- Sharma, G., Kaushal, Y., Chandra, S., Singh, V., Mittal, A. P., and Dutt, V. (2017). Influence of landmarks on wayfinding and brain connectivity in immersive virtual reality environment. *Front. Psychol.* 8:1220. doi: 10.3389/fpsyg.2017.01220
- Siegel, A. W., and White, S. H. (1975). The development of spatial representations of large-scale environments. *Adv. Child Dev. Behav.* 10, 9–55. doi: 10.1016/S0065-2407(08)60007-5
- Slone, E., Burles, F., Robinson, K., Levy, R. M., and Iaria, G. (2015). Floor plan connectivity influences wayfinding performance in virtual environments. *Environ. Behav.* 47, 1024–1053. doi: 10.1177/0013916514533189
- Sorrows, M. E., and Hirtle, S. C. (1999). “The nature of landmarks for real and electronic spaces,” in *International Conference on Spatial Information Theory*, eds. C. Freksa and D. M. Mark (Berlin: Springer), 37–50. doi: 10.1007/3-540-48384-5_3
- Stark, A., Riebeck, M., and Kawalek, J. (2007). “How to design an advanced pedestrian navigation system: field trial results,” in *2007 4th IEEE Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*, (Dortmund: IEEE), 690–694. doi: 10.1109/IDAACS.2007.4488511
- Stirk, J. A., and Underwood, G. (2007). Low-level visual saliency does not predict change detection in natural scenes. *J. Vision* 7, 3. doi: 10.1167/7.10.3
- Tang, C. H., Wu, W. T., and Lin, C. Y. (2009). Using virtual reality to determine how emergency signs facilitate way-finding. *Appl. Ergon.* 40, 722–730. doi: 10.1016/j.apergo.2008.06.009
- Tenbrink, T. (2008). *Space, Time, and the Use of Language: An Investigation of Relationships* (Vol. 36). Berlin: Walter de Gruyter. doi: 10.1515/9783110198829
- Tolman, E. (1948). Cognitive maps in rats and men. *Psychol. Rev.* 55, 189–208.
- Underwood, G., Foulsham, T., and Humphrey, K. (2009). Saliency and scan patterns in the inspection of real-world scenes: eye movements during encoding and recognition. *Visual Cogn.* 17, 812–834. doi: 10.1080/13506280902771278
- United Nations (2018). *2018 Revision of World Urbanization Prospects*. Available online at: <https://population.un.org/wup/> (accessed June 15, 2019).
- Vaidya, S. P., and Johnston, D. (2013). Temporal synchrony and gamma-to-theta power conversion in the dendrites of CA1 pyramidal neurons. *Nat. Neurosci.* 16, 1812–1820. doi: 10.1038/nn.3562
- Werner, S., and Schindler, L. E. (2004). The role of spatial reference frames in architecture: misalignment impairs way-finding performance. *Environ. Behav.* 36, 461–482. doi: 10.1177/0013916503254829
- White, D. J., Congedo, M., Ciorciari, J., and Silberstein, R. B. (2011). Brain oscillatory activity during spatial navigation: theta and gamma activity link medial temporal and parietal regions. *J. Cogn. Neurosci.* 24, 686–697. doi: 10.1162/jocn_a_00098

- White, T. P., Jansen, M., Doege, K., Mullinger, K. J., Park, S. B., Liddle, E. B., et al. (2013). Theta power during encoding predicts subsequent-memory performance and default mode network deactivation. *Hum. Brain Mapp.* 34, 2929–2943. doi: 10.1002/hbm.22114
- Winter, S., Tomko, M., Elias, B., and Sester, M. (2008) Landmark hierarchies in context. *Environ Plann B* 35, 381–398. doi: 10.1068/b33106
- Wolbers, T., and Büchel, C. (2005). Dissociable retrosplenial and hippocampal contributions to successful formation of survey representations. *J. Neurosci.* 25, 3333–3340. doi: 10.1523/JNEUROSCI.4705-04.2005
- Zhang, X., Li, Q. Q., Fang, Z. X., Lu, S. W., and Shaw, S. L. (2014). An assessment method for landmark recognition time in real scenes. *J. Environ. Psychol.* 40, 206–217. doi: 10.1016/j.jenvp.2014.06.008
- Zumer, J. M., Scheeringa, R., Schoffelen, J.-M., Norris, D. G., and Jensen, O. (2014). Occipital alpha activity during stimulus processing gates the information flow to object-selective cortex. *PLoS Biol.* 12:e1001965. doi: 10.1371/journal.pbio.1001965
- Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Rounds, Cruz-Garza and Kalantari. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



What to Expect When the Unexpected Becomes Expected: Harmonic Surprise and Preference Over Time in Popular Music

Scott A. Miles^{1,2*}, David S. Rosen^{2,3}, Shaun Barry², David Grunberg² and Norberto Grzywacz^{1,4,5}

¹ Interdisciplinary Program in Neuroscience, Georgetown University, Washington, DC, United States, ² Secret Chord Laboratories, Norfolk, VA, United States, ³ Music and Entertainment Technology Laboratory, Drexel University, Philadelphia, PA, United States, ⁴ Department of Psychology, Loyola University Chicago, Chicago, IL, United States, ⁵ Department of Molecular Pharmacology and Neuroscience, Loyola University Chicago, Chicago, IL, United States

OPEN ACCESS

Edited by:

Jose Luis Contreras-Vidal,
University of Houston, United States

Reviewed by:

Georgios N. Yannakakis,
University of Malta, Malta
Avi Mendelsohn,
University of Haifa, Israel
Jesus Gabriel Cruz-Garza,
Cornell University, United States

*Correspondence:

Scott A. Miles
scott@secretchordlaboratories.com

Specialty section:

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

Received: 30 June 2020

Accepted: 29 March 2021

Published: 30 April 2021

Citation:

Miles SA, Rosen DS, Barry S,
Grunberg D and Grzywacz N (2021)
What to Expect When the Unexpected
Becomes Expected: Harmonic
Surprise and Preference Over Time in
Popular Music.
Front. Hum. Neurosci. 15:578644.
doi: 10.3389/fnhum.2021.578644

Previous work demonstrates that music with more *surprising* chords tends to be perceived as more enjoyable than music with more conventional harmonic structures. In that work, harmonic surprise was computed based upon a static distribution of chords. This would assume that harmonic surprise is constant over time, and the effect of harmonic surprise on music preference is similarly static. In this study we assess that assumption and establish that the relationship between harmonic surprise (as measured according to a specific time period) and music preference is not constant as time goes on. Analyses of harmonic surprise and preference from 1958 to 1991 showed increased harmonic surprise over time, and that this increase was significantly more pronounced in preferred songs. Separate analyses showed similar increases over the years from 2000 to 2019. As such, these findings provide evidence that the human perception of tonality is influenced by exposure. Baseline harmonic expectations that were developed through listening to the music of “yesterday” are violated in the music of “today,” leading to preference. Then, once the music of “today” provides the baseline expectations for the music of “tomorrow,” more pronounced violations—and with them, higher harmonic surprise values—become associated with preference formation. We call this phenomenon the “Inflationary-Surprise Hypothesis.” Support for this hypothesis could impact the understanding of how the perception of tonality, and other statistical regularities, are developed in the human brain.

Keywords: music, surprise, harmony, preference, predictive coding

1. INTRODUCTION

In Miles et al. (2017), the examination of harmonic surprise and music preference tested two seemingly contradictory hypotheses about harmonic surprise and music preference. The *Absolute-Surprise Hypothesis* states that moderate increases in harmonic surprise are perceived as “good”: there is a relationship between music preference and musical popularity that is dependent on how much total harmonic surprise is contained in a piece of music. The *Contrastive-Surprise Hypothesis* states that increases in harmonic surprise are perceived as “bad”: sections of music with elevated

harmonic surprise lead to a mild pain signal in the brains of listeners, and when this is relieved during a subsequent section with lower harmonic surprise, the result is pleasure, leading to music preference.

Evidence supporting both hypotheses led to the formulation of the Hybrid-Surprise Hypothesis. Analyses showed that the per-song average harmonic surprise of the top quartile (Q_1) of Billboard charting songs is significantly higher than that of the bottom-quartile (Q_4) songs, providing support for the *Absolute-Surprise Hypothesis*. Furthermore, the results revealed that there is increased variation in surprise across the sections of Q_1 songs than Q_4 songs, providing support for the *Contrastive-Surprise Hypothesis*. However, each of these analyses included songs across all 33 years of the corpus, without taking release date into account. The findings thus assumed uniform effects of surprise on preference, based on a distribution of chords that never varies. This uniformity is the simplest hypothesis.

It is possible, however, that there are significant changes over the years in either the effects of surprise, the underlying chord distributions, or both. In fact, the musical properties of popular music have been shown to evolve over time (Mauch et al., 2015). The songs examined in Miles et al. (2017) from the McGill Billboard corpus (Burgoyne et al., 2011) span 33 years in their release date (1958–1991). Thus, it is likely that the effects of absolute or contrastive surprise on preference measurably change over this span of years, and any such change must be accounted for to further understand the nature of music preference and harmonic surprise. Through a separate set of analyses, the present study is designed to address this gap in the understanding of the relationship between harmonic surprise and preference in popular music.

In addition to addressing this gap in understanding, we also set out to address a gap in time. The present study introduces the possibility of a dynamic relationship between harmonic surprise and preference as time goes by. This highlights the lack of recently released songs in the McGill Billboard Corpus, which ends in 1991. To see if any change in the effects of harmonic surprise on preference over time extend to the current era of popular music, we added an analysis of the Secret Chord Laboratories (SCL) corpus. This corpus features a nearly exhaustive list of Billboard-charting songs released from January 1, 2000 to December 31, 2019.

2. LITERATURE REVIEW

The process by which the human brain perceives music is being extensively studied by researchers in the field (McDermott et al., 2016; Reybrouck et al., 2018; Daly et al., 2020). One recent development of note is the determination that the content of music may matter less to a listener's perception of a musical work than whether or not the listener enjoys that work (Wilkins et al., 2015). In other words, a listener is likely to have a more similar cognitive reaction to hearing two pieces of music that he or she prefers, even if those two works are very different, than to hearing two similar pieces of music, one of which he likes and one of which he or she dislikes. As such, determining how the human

brain determines whether a particular piece of music is enjoyable or not is of particular interest.

One prevailing theory as to how music evokes pleasurable responses in the human brain is that the music, by adhering to or deviating from what a listener would expect, can stimulate a neural reward (Meyer, 1956; Huron, 2006). This is reinforced by the evidence that music perception is based at least in large part on cultural knowledge. It has been shown that the musical culture that a listener grows up with has an influence on the understanding and perception of music later in life (Curtis and Bharucha, 2008), to the point where an individual's perception of, and enjoyment of, a new musical piece is heavily dependent on music already heard (McDermott et al., 2016). However, it has historically been difficult to evaluate this idea, in part because the concepts involved (such as the "amount of surprise" in the music or the expected "reward") are difficult to quantify. A system that could precisely estimate the amount of "surprise" in a piece of music and how much people might be expected to like it would thus be of utility to the research community.

In order to estimate surprise, one approach is drawn from the field of information theory (Rohrmeier and Koelsch, 2012). This approach has proven useful at describing various aspects of music. However, until recently, information theory-based approaches have not been able to take the additional step of describing how those aspects influence the perception of a given musical work.

One potentially useful aspect of information theory is the concept of "surprise," a measure of how much a given element deviates from what that element would be expected to be (Atick, 1992). This concept has been applied to music in order to determine how much a specific musical element deviates from the norm (Egermann et al., 2013). By taking a musical piece and calculating the amount of surprise in its components (such as its harmonies, its melody, its rhythm, its timbre, etc.), it could thus become possible to quantify the total amount of surprise in the music.

Furthermore, just as work has been done to quantify surprise, progress has also been made in quantifying musical perception. Because the position of a musical work on charts such as the Billboard Hot 100 is a function of how many people listen to it and buy it (Parker, 1992), music which delivers more reward to listeners can be expected to place higher on the charts than music which does not deliver as much reward. As such, features such as chart position can be used as an approximation of popularity and musical reward.

However, while there do exist projects which have sought to analyze musical response in terms of surprise, particularly the surprise of the music's harmonic content, much of this work relies on artificial datasets comprised of individual chords (Koelsch et al., 2001) or old, relatively simple music such as Bach chorales (Steinbeis et al., 2005), as opposed to modern songs. We therefore developed prior work on this topic (Miles et al., 2017) in which we calculated the harmonic surprise of actual popular music and assessed how this surprise related to the music's popularity on the charts. This was done to ensure that our results were relevant to the music people actually listen to in the modern age, or at least

in the time since the Billboard Hot 100 began being published in 1958.

Our prior work on this subject considered two possible hypotheses as to how surprise affects musical perception. One is the *Absolute-Surprise Hypothesis*, which states that musical popularity is determined by the overall amount of surprise in a piece. This is based on the theoretical foundation that dopamine, a pleasurable brain chemical, is often associated with novelty (Suhara et al., 2001), and thus that listening to music which is surprising or novel will likely produce more of this chemical. In addition to our own work, another prominent paper in this field used functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) machines to discover that harmonically unexpected elements of a work do tend to induce more dopamine production than more conventional musical elements (Salimpoor et al., 2011).

The second hypothesis we considered is the *Contrastive-Surprise Hypothesis*. In this approach, the response to a piece of music is not dependent on the total amount of surprise in the piece but on the contrast between high-surprise and low-surprise sections within a given musical work. This hypothesis is in line with previously-advanced theories noting that pleasure can be derived from first building up tension (such as with a surprising element) and then relieving it (as with a non-surprising element; Huron, 2006). Previous electrophysiological work (Koelsch et al., 2001) has found an association with the perception of unexpected chords to the neural correlate of mild irritation known as the early right anterior negativity, associated with prediction error being processed in the brain. The *Contrastive-Surprise Hypothesis* is consistent with a model of reward resulting from the relief of such irritation.

It is worth noting that our prior work (Miles et al., 2017) used a single distribution of chords over all time and thus assumed that the “expected” harmony of music was constant over the years across Western popular music. This assumption, however, may not have been valid; it is entirely possible that the common harmonies which can be reasonably expected to occur in music may change from year to year. As such, we present this current work, which seeks to investigate this possibility and determine if using a more sophisticated model for the expected harmonies allows for a more accurate model.

We also investigate the hypothesis that preferred music tends to increase in surprise over time, whether absolute, contrastive, or both, at a rate higher than any such increase over time in less preferred music. This hypothesis, which we call the *Inflationary-Surprise Hypothesis*, might be due to the ever-increasing requirement for music being released at any given time, considered “high-surprise,” to build upon the already increased surprise within preferred music that already exists. This phenomenon could be largely driven by effects of the listening habits during critical periods in the formation of harmonic expectations by listeners. The Billboard Hot 100 is known to be driven by an adolescent cohort of consumers. At this stage in their lives, teenagers generally want to be associated with the most popular new song or artist, since music preference is important to identity formation. In four studies, North and Hargreaves (1999) reported that music preference of a particular

style functions as an “identity badge,” whereby adolescents form their self-concepts and social judgments. It appears that these personal music definitions and choices for adolescents are likely to elicit emotional or spiritual experiences (Bosacki and O’Neill, 2015). The heightened social and emotional impact of music for adolescents creates a strong nostalgia, rekindling images of past selves, experiences, and friends who shared those musical preferences. Furthermore, the emotional content and subject matter of popular music connect with adolescents, because its sound and lyrics match the extreme emotional experience of their daily lives (Wells and Hakanen, 1991). It has been reported that music can function therapeutically to reduce feelings of stress and loneliness in adolescents (Zillmann and Gan, 1997).

If the effect of absolute and contrastive surprise on music preference indeed increases over time, this may be due to cascading cohorts of primarily adolescent listeners whose baseline expectations have been formed during a critical period of statistical learning at an earlier age. Each of these successive cohorts could be driving an apparent effect whereby overall harmonic surprise of preferred songs, as measured against a constant distribution of chords from the past, increases over time. Evidence of the *Inflationary-Surprise Hypothesis* would also support broader theories about musical expectations being learned through exposure early in life. In his song about advancements as a result of human progress, “Boy in the Bubble,” Simon (1986) sings: “...every generation throws a hero up the pop charts.” It is possible that these “heroes” use increasing harmonic surprise, over time, in their songs.

3. MATERIALS AND METHODS

The songs included in the McGill Billboard corpus of songs from 1958 to 1991 were separated into four consecutive time bins to examine how the effects of harmonic surprise on music preference change over time. The songs of a more recent corpus, a set of 6,051 songs on the Billboard Hot 100 chart released from 2000 to 2019 (the *SCL Corpus*), were also separated into four consecutive time bins (see **Table 1**). The *SCL Corpus* features a considerable representation of the 7,988 total unique songs that charted on the Billboard Hot 100 over that 20-year span of time. The null hypotheses stipulated that for each corpus, each type of harmonic surprise effect on preference across the four time bins would not significantly differ from one another. Support for this null hypothesis would suggest that there is no impact of time on the effects of harmonic surprise on music preference.

For the McGill corpus, we chose to group the earliest released songs of 1958–1975 together as a baseline to compare the effects of the remaining bins. In preliminary analyses when comparing average absolute surprise measures of each Billboard quartile, we did not observe any differences across time in these measures in Q_1 relative to Q_4 through this period. This allowed for the resulting first time bin to serve as a substantial baseline chord distribution from which to compute a uniform measure of surprise for the remaining time bins. We attribute this lack of change in the effect of harmonic surprise on preference during

TABLE 1 | Four time bins and corresponding periods within which songs of each time bin were released for both the “past” and “present” aspects of this study.

Time bin	Period
McGill 1	August 1958–January 1975
McGill 2	February 1975–July 1980
McGill 3	August 1980–January 1986
McGill 4	February 1986–November 1991
SCL 1	January 2000–December 2004
SCL 2	January 2005–December 2009
SCL 3	January 2010–December 2014
SCL 4	January 2015–December 2019

The first time bin establishes baseline harmonic surprise. In the McGill corpus, this bin includes songs released over 16.5 years, and each remaining time bin includes songs released within 5.5 years. In the SCL corpus, each bin includes songs released with 5 years.

this initial time-period to the establishment of a new genre: rock-n-roll.

The process resulted in four time bins for the McGill corpus (see **Table 1**), with 1958–1975 representing the first time bin. We then separated the SCL corpus into four time bins, each spanning 5 years of release dates, as well. We finally compared per-song average surprise and variation in surprise across sections of songs from Q_1 to those from Q_4 .

Next, we investigated trends over time within each of the two corpora to see if absolute surprise effects on preference, contrastive surprise effects on preference, or both, were changing over time.

All of the chords of songs from the McGill Billboard corpus were transcribed by hand (Burgoyne et al., 2011). The resulting labels describe each chord “up to seventh” level. In other words, each chord label specified all notes including and up to the seventh of the chord, if appropriate. The chords for the songs of the SCL corpus were estimated by using a neural network trained to predict the root note of the chord and whether the chord is major or minor (Korzeniowski and Widmer, 2016). The difference in these two approaches to transcription results in the possibility of a slight distinction between the results in the two corpora: surprise as measured in songs of the SCL corpus is more likely to reflect pure tonality, whereas the “color” from elements other than the root and third could influence surprise in songs of the McGill corpus.

For each corpus, the relationship between harmonic surprise and preference was calculated using a method based on the approach outlined in Miles et al. (2017) and Miles (2020). Chord labels from each corpus were normalized to the key of each song. For the songs of the McGill corpus, this means that the transcribed key was identified, and then each transcribed chord was labeled according to its relationship to that key. For the SCL corpus, this means that the probable key was detected using Korzeniowski’s and Widmer’s Convolutional Neural Network key-detection algorithm included in the Madmom software package (Böck et al., 2016; Korzeniowski and Widmer, 2018), and then each identified chord was then labeled according to its relationship to that key.

Next, zeroth-order harmonic surprise was calculated for each chord, based on the prevalence of chords in the years from 1958 to 1975 in the case of the McGill corpus, or 2000–2004 in the case of the SCL corpus. As in Miles et al. (2017), the analysis was limited to zeroth-order harmonic surprise, which does not take into account the ordering of chords. This limitation is due to the increased statistical power necessary to determine any higher-order surprise effects in such a small corpus. Surprise was calculated by first finding N , the total number of unique chords in the corpus, and then for each unique chord C_j finding M_j , the number of times that chord appears. This gave a total number of chords in the corpus (including repetitions) of $\sum_{i=1}^N M_i$. We then calculated the probability of unique chord C_j as in Equation (1).

$$P(C_j) = \frac{M_j}{\sum_{i=1}^N M_i} \quad (1)$$

Given the probability of a given chord, that chord’s surprise was calculated using the standard information theory equation, as shown in Equation (2).

$$S(C_j) = -\log_2(P(C_j)) \quad (2)$$

The total number of unique chords in the McGill corpus in songs from 1958 to 1975 was 348. This included chords with all twelve possible roots, and various modes and extensions for each of those root notes. The range of unique chords in songs from 2000 to 2004 in the SCL corpus, since it was based only on root and third notes, was 24. This included all twelve possible roots, each with either major or minor thirds.

Once the surprise for each chord in a piece of music was obtained, then that piece’s overall absolute and contrastive surprise were found. Absolute surprise was estimated by taking the mean surprise of each chord in a song, and contrastive surprise was estimated by finding the standard deviation (SD) of mean surprise for each section of a song, with sections being calculated algorithmically according to Nieto’s and Bello’s Music Structure Analysis Framework (Nieto and Bello, 2016).

A uniform chord distribution of “all songs” was used in the previous analyses of Miles et al. (2017). In this analysis, we calculated surprise based on the uniform chord distribution statistics either of “1958–1975 (combined),” in the case of the McGill corpus, or of “2000–2004 (combined),” in the case of the SCL corpus.

To determine the classification of songs as “top quartile” or “bottom quartile”— Q_1 or Q_4 —the process was slightly different for each of the two corpora. For the *McGill Billboard Corpus*, the 545 total songs were ordered by peak Billboard chart position, with number of weeks on the chart breaking any ties in chart position. The resulting 136 top-ranking songs were then classified as Q_1 songs, and the 136 bottom-ranking songs were classified as Q_4 songs. For the *SCL Corpus*, the songs were first broken into groups by year of release, and then further broken into groups within each year of release into genres. This was done to reduce any variability in preference for any particular genre within the charts. Genres were taken from metadata tagged by Apple Music. The resulting “year*genre” groups were then ordered by peak

Billboard chart position and number of weeks on the chart. The top 25% of songs in each group were then classified as Q_1 and the bottom 25% of songs in each group were classified as Q_4 .

4. RESULTS

4.1. The Effect of Absolute Surprise Over Time— Q_1 and Q_4 Per-Song Harmonic Surprise Over Time

We looked at per-song average surprise in the four newly defined time bins of each corpus, across top and bottom quartiles. We then computed linear regression lines best fitting the Q_1 and Q_4 data over time. The results of the analyses are presented in **Figures 1, 2** below.

In the analysis of absolute surprise over time in the McGill corpus, with per-song average surprise calculated using the chord distribution of bins 1958–1975 (and excluding newly introduced chords), there were significant upward trends in Q_1 (Jonckheere-Terpstra, $p < 0.001$) and in Q_4 (Jonckheere-Terpstra, $p < 0.05$). This suggests an increase in surprise over time for songs of both quartiles. Tests showed that while there was no significant difference in per-song average surprise between Q_1 and Q_4 in time-bin “1958–1975,” per-song average surprise was significantly higher among the three remaining time bins (t -test, Q_1 mean = 5.46; Q_4 mean = 4.78, $p < 0.01$). In the data of **Figure 1**, we observed that the average rate of change for songs in Q_1 , +0.096 bits/year, was three times the trend in Q_4 , +0.032 bits/year. In order to test the significance of the slopes calculated for trends over time in Q_1 and Q_4 , we also tested the null hypothesis that slopes are equivalent using the regression slopes test provided by Zaiontz (2013). Equation (3) below gives the statistic t that was calculated using the slopes obtained from linear regression of surprise data for Q_1 and Q_4 . The test was normalized by using the standard error of the slope for Q_1 and Q_4 , s_{Q_1} , and s_{Q_4} , respectively. In this equation, n_1 and n_2 represent the number of songs in Q_1 and Q_4 , respectively, and T represents the student's t distribution.

$$t = \frac{\text{slope}(Q_1) - \text{slope}(Q_4)}{\sqrt{s_{Q_1}^2 + s_{Q_4}^2}} \sim T(n_1 + n_2 - 4) \quad (3)$$

For the data in **Figure 1**, we obtain a p -value of 1.8×10^{-12} . This suggests that the increase in absolute surprise for Q_1 songs was more pronounced than that of Q_4 songs in the McGill corpus.

In the analysis of absolute surprise over time of the SCL corpus, with per-song average surprise calculated using the chord distribution of 2000–2004, there were upward trends in Q_1 and Q_4 . The average rate of change for songs in Q_1 , +0.0146 bits/year, is also greater than three times the rate in Q_4 , +0.0046 bits/year. This suggests an increase in surprise over time for songs of both quartiles. With the same statistical test comparing the slopes of the two trend lines for the SCL corpus data, we obtained a p -value of 3×10^{-11} , suggesting that the increase in Q_1 songs was more pronounced than that of Q_4 songs in the SCL corpus. Note that the scale of calculated surprise in songs of the SCL corpus is distinct from that of the McGill corpus. This is the result of the different methodology in calculating surprise: in the McGill

corpus, chords were transcribed by humans up to the seventh tone, while in the SCL corpus, chords were algorithmically determined only at the root and third tones. This resulted in far fewer unique chords in the SCL corpus than in the McGill corpus (24 and 348, respectively), thereby lowering the overall surprise values for each chord in the SCL corpus.

4.2. Effect of Contrastive Surprise Over Time: Variation in Surprise Among Sections Within Songs Over Time

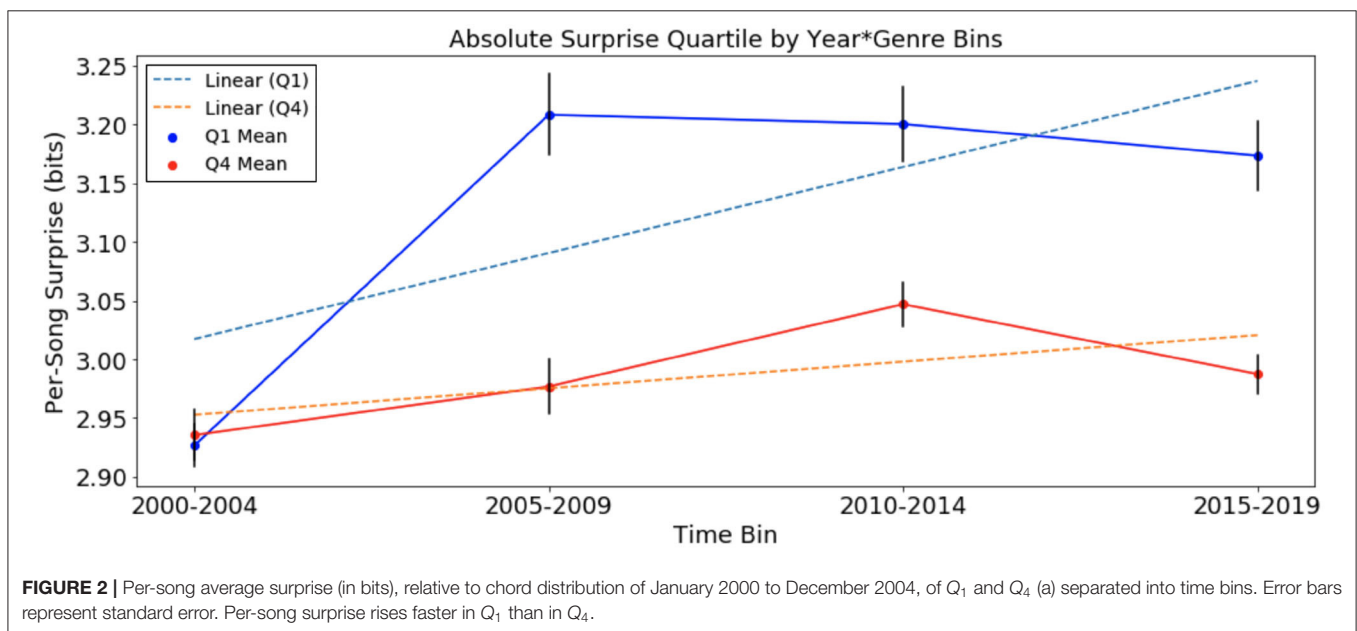
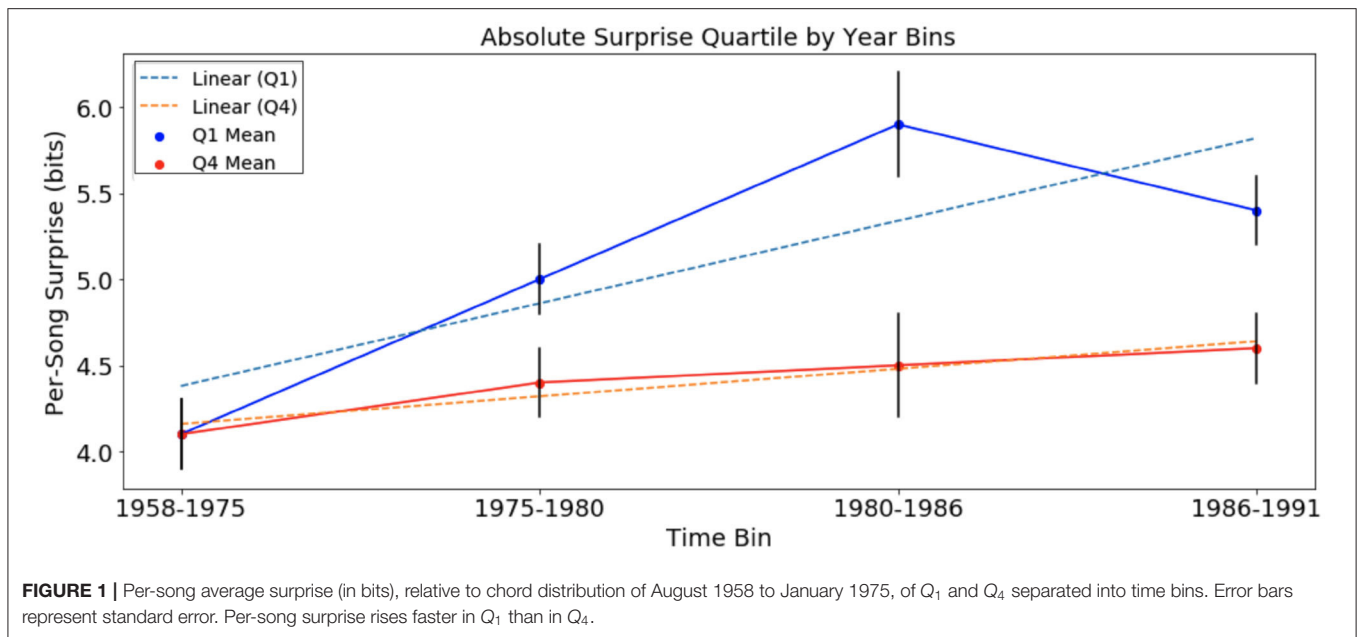
Next, we looked at variation of surprise across sections within Q_1 and Q_4 songs in the same time bins used in the previous section with surprise computed using the chord distribution of the original 1958–1975 distribution and 2000–2004 distribution for McGill and SCL corpora, respectively. The results are presented below in **Figures 3, 4**.

In this analysis, there was an increase in variation across sections of songs in both Q_1 and Q_4 (Jonckheere-Terpstra, Q_1 : $p < 0.001$; Q_4 : $p < 0.05$). The rate of change for songs in Q_1 at +0.044 bits/year is more than twice the slope of the trend in Q_4 at +0.0166 bits/year. Tests showed that while there was no significant difference in variation across sections between Q_1 and Q_4 in time bin 1958–1975, variation was significantly higher among time bins 2–4 (t -test, Q_1 mean = 1.45; Q_4 mean = 1.21, $p < 0.05$). Additionally, $p < 1 \times 10^{-100}$ for the slope test suggests that the increase in Q_1 songs was more pronounced than that of Q_4 songs in the McGill corpus.

In the analysis of contrastive surprise over time in **Figure 4**, with variation of surprise across sections using the chord distribution of 2000–2004 in the SCL corpus, there were upward trends in Q_1 and Q_4 . The rate of change in Q_1 of +0.0032 bits/year is approximately 1.5 times the rate of change in Q_4 of +0.002 bits/year. This suggests an increase in surprise over time for songs of both quartiles. As we did in the previous analysis, we test the null hypothesis that slopes are equivalent and obtain a p -value of 0.008. This suggests that the increase in contrastive surprise for Q_1 songs was significantly more pronounced than that of Q_4 songs in the SCL corpus. Note that the scale of these results from songs of the SCL corpus is again smaller, due to the much smaller number of possible unique chords.

4.3. New Chords Introduced in Q_1 and Q_4 Songs

In the McGill Billboard corpus, “up to seventh” labels of chords included myriad permutations of each chord based on the possible 12 roots and 12 thirds. This was not an issue with the SCL corpus, which only included root and third notes in its chord labels. By using a chord distribution, in the McGill corpus analyses, that reflects only songs of 1958–1975, we failed to account for any surprise effects due to new chords that might be introduced into the distribution over time. Chords being introduced for the first time in the corpus are likely to be highly harmonically unexpected and could significantly contribute to any effects of harmonic surprise on preference. To examine the relative contribution to surprise of new chords, we examined the prevalence of chords in each quartile that appeared in years 1959 to 1991 that had not previously appeared in any year



(Figure 5). This analysis showed that 28.7% of newly introduced chords appeared in Q_1 songs, and 19.8% of newly introduced chords appeared in Q_4 songs. New chords were found to appear significantly more frequently in Q_1 than in Q_4 ($\chi^2 = 133.5$, $df = 1$, $p < 0.001$).

5. DISCUSSION

5.1. Findings and Their Support for the Proposed Hypotheses

In the statistical corpus analyses reported in Miles et al. (2017), we found evidence that the brain uses both absolute- and

contrastive-surprise strategies in determining preference for popular songs. In the present study, we tested whether absolute and contrastive harmonic surprise effects on preference varied over time across the 33 years of the McGill corpus and across the 20 years of the SCL corpus.

Tests of trend showed that in both corpora, Q_1 and Q_4 average per-song surprise increased across consecutive time bins when surprise was calculated using the chord distribution of the first time bin, and that Q_1 average per-song surprise increased at a significantly greater rate. Also in both corpora, tests of trend showed that variation in surprise across sections of Q_1 and Q_4 songs increased over time, and that variation in surprise across sections of Q_1 songs increased at a significantly greater rate.

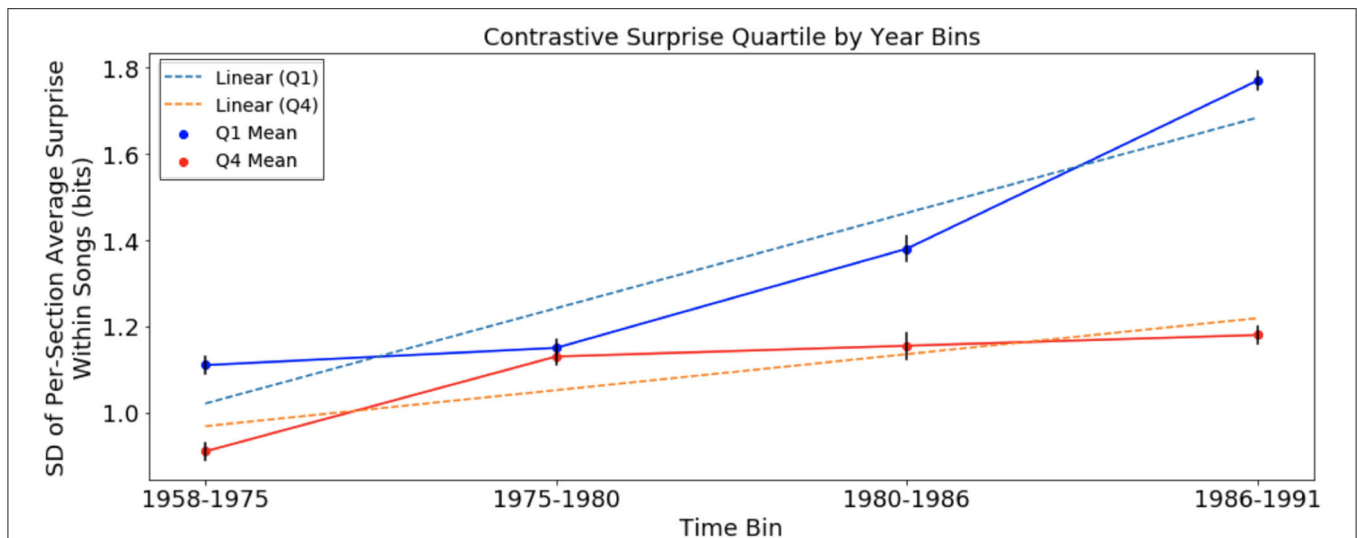


FIGURE 3 | Variation of harmonic surprise across sections of songs, with harmonic surprise calculated using chord distribution of August 1958 to January 1975, of Q_1 and Q_4 , separated into time bins. Error bars represent standard error. Variation of surprise across sections rises faster in Q_1 than in Q_4 .

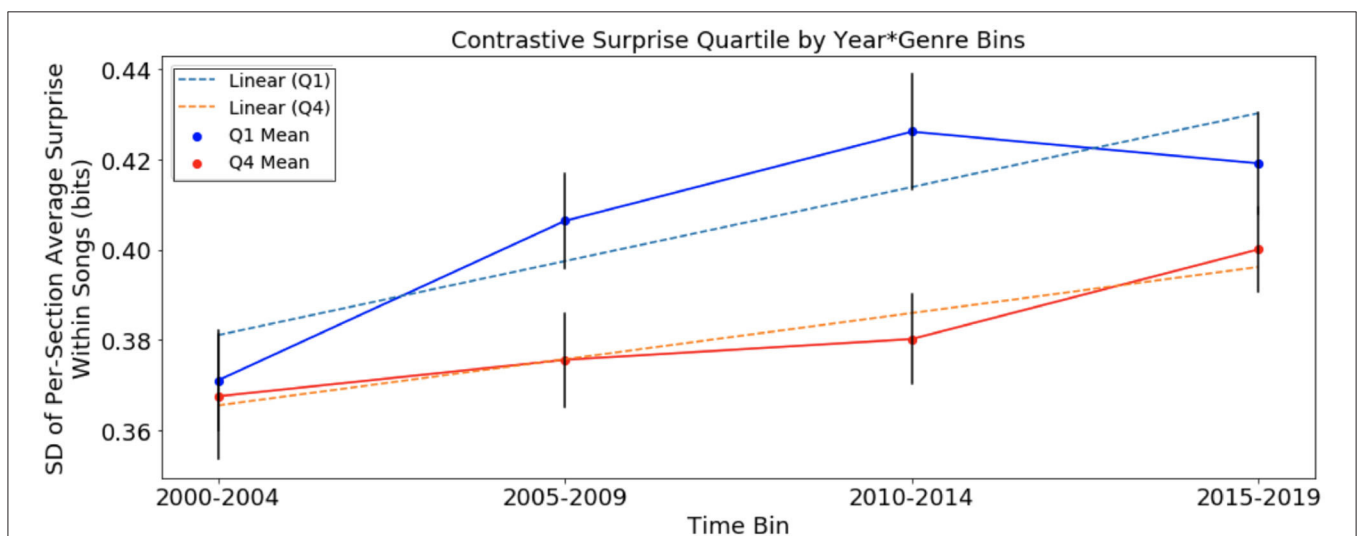


FIGURE 4 | Variation of harmonic surprise across sections of songs, with harmonic surprise calculated using chord distribution of January 2000 to December 2004, of Q_1 and Q_4 , separated into time bins. Error bars represent standard error. Variation of surprise across sections rises faster in Q_1 than in Q_4 .

These results rule out the null hypothesis that absolute and contrastive harmonic surprise effects on music preference do not vary over time.

Rather, these results support the Inflationary-Surprise Hypothesis: it appears that the effects of harmonic surprise on music preference have to be more pronounced over time to get the same effect. The force driving harmonic surprise upward in the Hot 100 chart, and more forcefully upward in Q_1 songs specifically, could be due to the statistical learning of harmonic regularities. Expectations for these regularities could be learned early in life, during a critical window for developing models of tonality. The window of consumers driving performance of

songs on the chart features a cycle of successive cohorts of 13- to 19-year old listeners. These listeners are likely to develop their baseline expectations from the preferred music released in the past, and because of the relationship between moderate increases in harmonic surprise and preference, the cycle results in preferred music that increases in harmonic surprise relative to a fixed composition of chords.

Successful musicians from the present use chords that are surprising not just for the current moment but also relative to the past. These successful musicians also keep introducing new chords. Consequently, the probability distribution of chords is ever changing, such that musicians must create new surprise

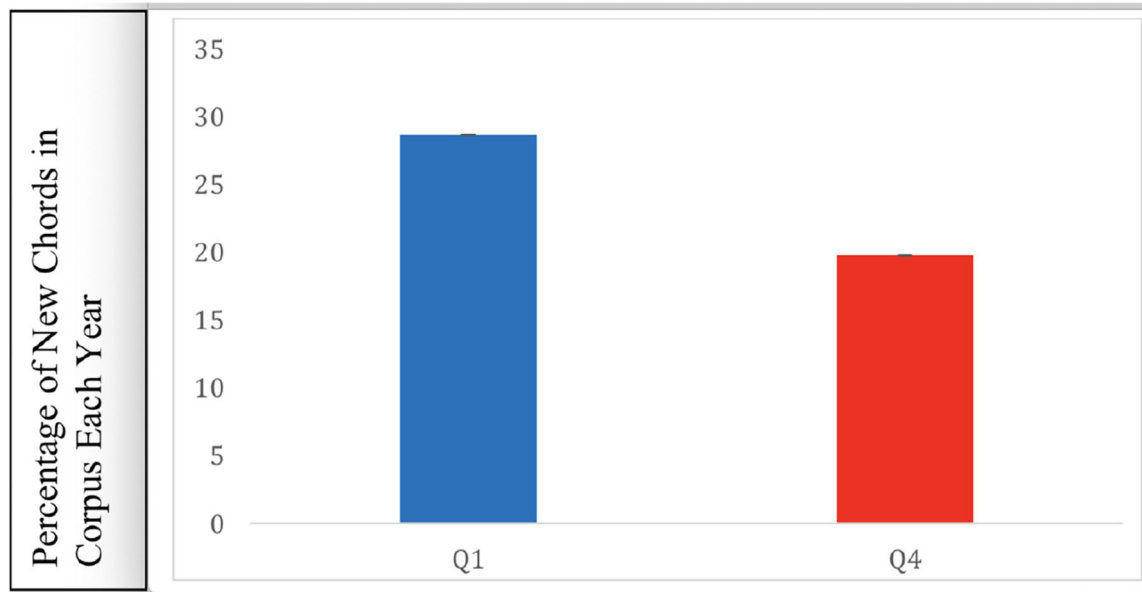


FIGURE 5 | Percentage of chords that appear over the years from 1959 to 1991, that had not appeared in songs of previous years, in Q₁ and Q₄. Error bars represent standard error. There were significantly more novel chords featured in Q₁ songs.

to accomplish the same level of preference, leading to an inflationary effect.

Another interesting result is that in the McGill corpus, per-song average surprise (**Figure 1**), the rise of Q₁ surprise seems to level off around six bits. There is a corresponding plateau effect of per-song average surprise in the SCL corpus around 3.2 bits. This could be evidence of a ceiling effect for absolute surprise. The data here are not sufficient to test this hypothesis. It is possible that this is related to an “inverted-U” effect of the relationship between complexity and pleasure (Berlyne, 1973). It is also possible that the inflationary effect of the relationship between harmonic surprise and preference can only go so far, and the trend is not infinitely sustainable. Future research might be useful in exploring whether there is an ideal range of harmonic surprise in popular music, such that too much or too little harmonic surprise is inversely related to pleasure.

When calculated relative to the chord distribution statistics from songs released from August 1958 to February 1975 in the McGill corpus, and when calculated relative to the chord distribution statistics from songs released from 2000 to 2004 in the SCL corpus: per-song average surprise in Q₁ and variation in surprise across sections of Q₁ songs were shown to increase more over time than the corresponding measures of Q₄ songs. Additionally, chords introduced for the first time in a given year were more likely to appear in a Q₁ song than in a Q₄ song in the McGill corpus. Taken together, these findings are consistent with the Surprise-Inflation Hypothesis.

Evidence for measurable increases of harmonic surprise over time and for apparent increases over time in the magnitude of measured surprise advantages in preferred music

is probably linked to how schematic information about musical systems is acquired. Capacity for perceiving fundamental pitch features in music, such as octave equivalence, are thought to be evolutionarily conserved, extending to other species (Greenwood, 1997). Higher-level aspects of music processing such as tonality, however, have been shown to be statistically learned through exposure (Tillmann et al., 2000; Loui and Wessel, 2008). In addition, structural expectations have been demonstrated to be learned through regularities within auditory sequences (Saffran et al., 1999; Tillmann and Poulin-Charronnat, 2010). While there is empirical evidence of statistical learning (even within the time frame of a behavioral experiment) of schematic regularities within artificial music systems, there is less evidence of shifts in higher-level expectations based on tonality in ecologically valid music. A notable exception, however, is presented in Rohrmeier and Widdess (2012), where exposure to a novel tonal system of regularities impacted subsequent expectations by participants. Investigations into tonality in Western music (e.g., Krumhansl and Keil, 1982; Tillmann et al., 2000), approach its system of tonality as a relatively fixed hierarchy. The finding of these differing surprise measures over time, however, is evidence of non-static harmonic expectations within Western popular music, as well as shifts in preference for various harmonic elements within it. Such shifts in preference are consistent with several components of the framework of how aesthetic values are learned over time presented in Aleem et al. (2020). These components include the shaping of reward value according to probabilistic information from exposure to stimuli and a “peak-shift” effect marked by the exaggeration of desirable features.

6. CONCLUSION

In an examination of the relationship between harmonic surprise and preference in popular music over the years, we found that surprise relative to a fixed distribution of chords seems to increase over time, and that this increase is significantly more pronounced in preferred songs. Such dynamic harmonic expectations highlight the interactions between individual listeners and musicians with the culture around them. The Surprise-Inflation Hypothesis raised by the results presented here suggests that the brain's craving for surprise causes continuous changes in harmonic distributions in popular music. A musician exposed to changes advanced by other musicians must innovate to be successful. It could be that musicians, learning from the success of high-surprise songs from one year, end up producing more high-surprise songs the next year. This could be an explicit strategy to improve on the part of musicians, rather than an implicit change in expectation on the part of the listeners. However, listeners' preferences change as a result, forcing musicians to incorporate further changes. Hence, the inherent craving for surprise in each of us may push our entire culture in an endless evolution of musical preferences.

DATA AVAILABILITY STATEMENT

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

REFERENCES

- Aleem, H., Correa-Herran, I., and Grzywacz, N. (2020). A theoretical framework for how we learn aesthetic values. *Front. Hum. Neurosci.* 14:345. doi: 10.3389/fnhum.2020.00345
- Atick, J. (1992). Could information theory provide an ecological theory of sensory processing? *Network* 22, 4–44. doi: 10.3109/0954898X.2011.638888
- Berlyne, D. E. (1973). "The vicissitudes of aplopathematic and thelematoscopic pneumatology (or the hydrography of hedonism)," in *Pleasure, Reward, Preference: Their Nature, Determinants, and Role in Behavior*, eds D. E. Berlyne and K. B. Madsen (Elsevier), 1–33.
- Böck, S., Korzeniowski, F., Schlüter, J., Krebs, F., and Widmer, G. (2016). "Madmom: a new Python audio and music signal processing library," in *Proceedings of the 24th ACM International Conference on Multimedia (Amsterdam)*, 1174–1178. doi: 10.1145/2964284.2973795
- Bosacki, S. L., and O'Neill, S. A. (2015). Early adolescents' emotional perceptions and engagement with popular music activities in everyday life. *Int. J. Adolesc. Youth* 20, 228–244. doi: 10.1080/02673843.2013.785438
- Burgoyne, J. A., Wild, J., and Fujinaga, I. (2011). "An expert ground truth set for audio chord recognition and music analysis," in *ISMIR, Vol. 11* (Montreal, CA), 633–638.
- Curtis, M., and Bharucha, J. (2008). Memory and musical expectation for tones in cultural context. *Music Percept.* 26, 365–375. doi: 10.1525/mp.2009.26.4.365
- Daly, I., Nicolaou, N., Williams, D., Hwang, F., Kirke, A., Miranda, E., et al. (2020). Neural and physiological data from participants listening to affective music. *Sci. Data* 7, 1–7. doi: 10.1038/s41597-020-0507-6
- Egermann, H., Pearce, M., Wiggins, G., and McAdam, S. (2013). Probabilistic models of expectation violation predict psychophysiological emotional responses to live concert music. *Cogn. Affect. Behav. Neurosci.* 13, 533–553. doi: 10.3758/s13415-013-0161-y
- Greenwood, D. (1997). The mel scale's disqualifying bias and a consistency of pitch-difference equisections in 1956 with equal cochlear distances and equal frequency ratios. *Hear. Res.* 103, 199–224. doi: 10.1016/S0378-5955(96)00175-X
- Huron, D. (2006). *Sweet Anticipation: Music and the Psychology of Expectation*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/6575.001.0001
- Koelsch, S., Gunter, T., Schröger, E., Tervaniemi, M., Sammler, D., and Friederici, A. (2001). Differentiating ERAN and MMN: an ERP study. *NeuroReport* 12, 1385–1389. doi: 10.1097/00001756-200105250-00019
- Korzeniowski, F., and Widmer, G. (2016). "A fully convolutional deep auditory model for musical chord recognition," in *2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP)*, 1–6. doi: 10.1109/MLSP.2016.7738895
- Korzeniowski, F., and Widmer, G. (2018). "Genre-agnostic key classification with convolutional neural networks," in *Proceedings of the 19th International Society for Music Information Retrieval Conference (ISMIR)*, 264–270.
- Krumhansl, C., and Keil, F. (1982). Acquisition of the hierarchy of tonal functions in music. *Mem. Cogn.* 10, 243–251. doi: 10.3758/BF03197636
- Loui, P., and Wessel, D. (2008). Learning and liking an artificial musical system: effects of set size and repeated exposure. *Music. Sci.* 12:207. doi: 10.1177/102986490801200202
- Mauch, M., MacCallum, R., Levy, M., and Leroi, A. (2015). The evolution of popular music: USA 1960–2010. *R. Soc. Open Sci.* 2. doi: 10.1098/rsos.150081
- McDermott, J., Schultz, A., Undurraga, E., and Godoy, R. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature* 535, 547–550. doi: 10.1038/nature18635
- Meyer, L. (1956). *Emotion and Meaning in Music*. Chicago, IL: University of Chicago Press.
- Miles, S. (2020). *Predicting the Popularity of a Song Based on Harmonic Surprise*. Available online at: <https://patents.google.com/patent/US20200265083A1> (accessed June 25, 2020).

AUTHOR CONTRIBUTIONS

SM was involved in the design of the project, developing, and using programming scripts for data analysis, interpreting results, writing, and revising the manuscript. SB was involved in the data analysis, implementation of neural networks, and writing the manuscript. DG was involved in data collection, interpreting the results, and writing the manuscript. DR was involved in the design of the project, interpreting results, writing, and revising the manuscript. NG was involved in the design of the project, devising the formulas, interpreting results, writing, and revising the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The research was also supported by funds provided by the President and Board of Trustees of Loyola University Chicago, and by Research Funds from Georgetown University.

ACKNOWLEDGMENTS

NG would like to thank Helen Ryan and Gina Lopez for their administrative support, and Jo Ann Rooney for creating an exceptional work environment.

- Miles, S., Rosen, D., and Grzywacz, N. (2017). A statistical analysis of the relationship between harmonic surprise and preference in popular music. *Front. Hum. Neurosci.* 11:263. doi: 10.3389/fnhum.2017.00263
- Nieto, O., and Bello, J. (2016). "Systematic exploration of computational music structure research," in *Proceedings of the 17th International Society for Music Information Retrieval Conference (ISMIR)* (New York City, NY), 547–553.
- North, A. C., and Hargreaves, D. J. (1999). Music and adolescent identity. *Mus. Educ. Res.* 1, 75–92. doi: 10.1080/1461380990010107
- Parker, M. (1992). Reading the charts - making sense with the hit parade. *Pop. Mus.* 10, 205–217. doi: 10.1017/S0261143000004517
- Reybrouck, M., Vuust, P., and Brattico, E. (2018). Brain connectivity networks and the aesthetic experience of music. *Brain Sci.* 8:107. doi: 10.3390/brainsci8060107
- Rohrmeier, M., and Koelsch, S. (2012). Predictive information processing in music cognition. A critical review. *Int. J. Psychophysiol.* 83, 164–175. doi: 10.1016/j.ijpsycho.2011.12.010
- Rohrmeier, M., and Widdess, R. (2012). "Incidental learning of modal features of north Indian music," in *Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music* (Thessaloniki), 23–28.
- Saffran, J., Johnson, E., Aslin, R., and Newport, E. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition* 70, 27–52. doi: 10.1016/S0010-0277(98)00075-4
- Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A., and Zatorre, R. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nat. Neurosci.* 14, 257–264. doi: 10.1038/nn.2726
- Simon, P. (1986). *Boy in the Bubble, on Album "Graceland"*. New York, NY: Warner Brothers.
- Steinbeis, N., Koelsch, S., and Sloboda, J. (2005). Emotional processing of harmonic expectancy violations. *Ann. N. Y. Acad. Sci.* 1060, 457–461. doi: 10.1196/annals.1360.055
- Suhara, T., Yasuno, F., Sudo, Y., Yamamoto, M., Inoue, M., Okubo, Y., et al. (2001). Dopamine D2 receptors in the insular cortex and the personality trait of novelty seeking. *NeuroImage* 13, 891–895. doi: 10.1006/nimg.2001.0761
- Tillmann, B., Bbarucha, J., and Bigand, E. (2000). Implicit learning of tonality: a self-organizing approach. *Psychol. Rev.* 107, 885–913. doi: 10.1037/0033-295X.107.4.885
- Tillmann, B., and Poulin-Charronnat, B. (2010). Auditory expectations for newly acquired structures. *Q. J. Exp. Psychol.* 63, 1646–1664. doi: 10.1080/17470210903511228
- Wells, A., and Hakanen, E. A. (1991). The emotional use of popular music by adolescents. *J. Q.* 68, 445–454. doi: 10.1177/107769909106800315
- Wilkins, R., Hodges, D., Laurienti, P., Steen, M., and Burdette, J. (2015). Network science and the effects of music preference on functional brain connectivity: from Beethoven to Eminem. *Sci. Rep.* 4:6130. doi: 10.1038/srep06667
- Zaiontz, C. (2013). *Real Statistics Resource Pack Software (Release 6.8)*. Copyright (2013–2020). Available online at: www.real-statistics.com (accessed June 4, 2020).
- Zillmann, D., and Gan, S. (1997). *Musical Taste in Adolescence*. Oxford University Press.

Conflict of Interest: The authors declare that this study received funding from Secret Chord Laboratories. The authors are employees of SCL and proprietary algorithms belonging to SCL were used to analyze the data.

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



Stochasticity, Nonlinear Value Functions, and Update Rules in Learning Aesthetic Biases

Norberto M. Grzywacz^{1,2*}

¹Department of Psychology, Loyola University Chicago, Chicago, IL, United States, ²Department of Molecular Pharmacology and Neuroscience, Loyola University Chicago, Chicago, IL, United States

OPEN ACCESS

Edited by:

Surjo R. Soekadar,
Charité—Universitätsmedizin Berlin,
Germany

Reviewed by:

Tyler Davis,
Texas Tech University, United States
Mariya Cherkasova,
West Virginia University,
United States

*Correspondence:

Norberto M. Grzywacz
norberto@luc.edu

Specialty section:

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

Received: 08 December 2020

Accepted: 31 March 2021

Published: 10 May 2021

Citation:

Grzywacz NM (2021) Stochasticity,
Nonlinear Value Functions, and
Update Rules in Learning Aesthetic
Biases.
Front. Hum. Neurosci. 15:639081.
doi: 10.3389/fnhum.2021.639081

A theoretical framework for the reinforcement learning of aesthetic biases was recently proposed based on brain circuitries revealed by neuroimaging. A model grounded on that framework accounted for interesting features of human aesthetic biases. These features included individuality, cultural predispositions, stochastic dynamics of learning and aesthetic biases, and the peak-shift effect. However, despite the success in explaining these features, a potential weakness was the linearity of the value function used to predict reward. This linearity meant that the learning process employed a value function that assumed a linear relationship between reward and sensory stimuli. Linearity is common in reinforcement learning in neuroscience. However, linearity can be problematic because neural mechanisms and the dependence of reward on sensory stimuli were typically nonlinear. Here, we analyze the learning performance with models including optimal nonlinear value functions. We also compare updating the free parameters of the value functions with the delta rule, which neuroscience models use frequently, vs. updating with a new Phi rule that considers the structure of the nonlinearities. Our computer simulations showed that optimal nonlinear value functions resulted in improvements of learning errors when the reward models were nonlinear. Similarly, the new Phi rule led to improvements in these errors. These improvements were accompanied by the straightening of the trajectories of the vector of free parameters in its phase space. This straightening meant that the process became more efficient in learning the prediction of reward. Surprisingly, however, this improved efficiency had a complex relationship with the rate of learning. Finally, the stochasticity arising from the probabilistic sampling of sensory stimuli, rewards, and motivations helped the learning process narrow the range of free parameters to nearly optimal outcomes. Therefore, we suggest that value functions and update rules optimized for social and ecological constraints are ideal for learning aesthetic biases.

Keywords: reinforcement learning, aesthetic value, value function, delta rule, regret minimization, stochastic dynamics

INTRODUCTION

Values and in particular aesthetic ones are a significant part of our lives because they contribute to our process of decision making (Skov, 2010). Because humans are highly social animals, the set of values of each person must be in tune with their cultures and surroundings. Therefore, learning is an essential component of how our values come to be. In the case of aesthetic values, they begin to be learned early on in life, such that by preschool age, cultural idiosyncrasies are observed in children (Senzaki et al., 2014). In addition, these values continue to progress over our lifespans (Park and Huang, 2010).

How does the brain learn aesthetic values? An important meta-analysis of neuroimaging considered commonalities of aesthetic biases across multiple sensory modalities (Brown et al., 2011). The results of this and many other imaging studies indicated general mechanisms for appraisal involving a well-studied (Schultz, 1998, 2016) reward-based learning circuit (Lacey et al., 2011; Vartanian and Skov, 2014; Wang et al., 2015). However, these studies suggest that many independent factors impact this process of reward-based learning, with Brown et al. (2011) in particular discussing a novel role for motivation.

Because the development of aesthetic biases involves a rewards-based learning circuitry, a mechanism akin to reinforcement learning (O'Doherty et al., 2003; Sutton and Barto, 2018) likely mediates the process. Several theoretical frameworks for aesthetic values have elements of reward-circuitry and reinforcement learning. Some of these theories are computational (Martindale, 1984; Schmidhuber, 2010; Van de Cruys and Wagemans, 2011; Aleem et al., 2019, 2020) and some are not (Biederman and Vessel, 2006; Skov, 2010; Vessel and Rubin, 2010; Chatterjee and Vartanian, 2014). Of the computational theories, the only one considering motivation is that of Aleem et al. This is also the only theory studying the temporal evolution of learning. Simulations of a model based on the Aleem et al. theoretical framework and mathematical analysis lead to three main findings. First, different people may develop distinct weighing of aesthetic variables because of individual variability in motivation (Nelson and Morrison, 2005; Brown and Dissanayake, 2009; Silvia et al., 2009). Demonstration of the development of individuality is especially important in a theory in which learning leads to a degree of coordination of aesthetic values across society. Second, individuals from different cultures and environments may develop different aesthetic values because of unique sensory inputs and social rewards. Third, because learning is stochastic stemming from probabilistic sensory inputs, motivations, and rewards, aesthetic values vary in time.

A potential problem for reinforcement-learning models for the brain is the linearity of many of the most important mechanisms. For example, the model used by Aleem et al. (2020) assumes a linear value function (Sutton and Barto, 2018), that is, a linear relationship between sensory inputs and values. Furthermore, this model makes a linearity assumption for the update rule of the value function. Thus, although the reward has a nonlinear dependence on sensory inputs, brain actions would approximate this dependence linearly. Biologically, these linear mechanisms are not reflective of typical reward-related

neural signaling (Schultz, 2015). Moreover, recent studies have signaled the need for a new conception of aesthetics that incorporates distributed processing and nonlinear recurrent networks (Leder and Nadal, 2014; Nadal and Chatterjee, 2019). Assuming such linear mechanisms is common even in Machine Learning to lighten computations and mathematical analysis (Chung et al., 2018). In addition, linear methods have also been well-explored theoretically (Tsitsiklis and Van Roy, 1997; Maei, 2011; Mahmood and Sutton, 2015; Iigaya et al., 2020) and empirically (Dann et al., 2014; White and White, 2016) in the Machine Learning literature. Finally, arguments have been made that linear rules perform comparably to deep neural networks when predicting subjective aesthetic values (Iigaya et al., 2020). However, modeling nonlinear processes with linear approximations should produce errors, or equivalently, regret in Machine Learning terminology (Kaelbling et al., 1996; Sutton and Barto, 2018; formally, regret is the difference between an agent's performance with that of an agent that acts optimally). Hence, increasing effort has begun in Machine Learning to develop methods for nonlinear value functions (Tesauro, 2005; Xu et al., 2007; Kober et al., 2013; Gu et al., 2016; Osband et al., 2016; Chung et al., 2018).

In this article, we present mathematical and computational analyses of linear and nonlinear reinforcement-learning models for the acquisition of aesthetic values. We analyze 16 models. They stem from the combination of four types of value function (one linear and three nonlinear) and four types of value-function update rule (two making a linearity assumption for the updates and two assuming nonlinearities). All these models incorporate motivation (Brown et al., 2011). The comparisons between the models use different metrics, the most important of which is “regret.” We measure regret as the difference between reward and the prediction of reward. We choose this metric because humans often experience emotional responses to regret as a decision error (Gilbert et al., 2004; Filiz-Ozbay and Ozbay, 2007; Somasundaram and Diecidue, 2016). Another metric is time of convergence, which is important because a good learning mechanism should acquire its values as quickly as possible.

THEORETICAL CONSIDERATIONS

We have split the description of the theoretical considerations into two subsections, general and mathematical. The “General Description of the Theoretical Considerations” section has a description of the ideas without any equations. Our goal in that section is to help the reader understand the elements of the theoretical considerations at an intuitive level. That section may allow some readers to skip the equations (“Mathematical Description of the “Theoretical Considerations” section) and the “Materials and Methods” section, and go directly to the “Results” section. The subsections of “General Description of the Theoretical Considerations” and “Mathematical Description of the Theoretical Considerations” sections have parallel titling. The parallel subjects of these subsections may help the reader when connecting the intuitive and mathematical levels of understanding.

General Description of the Theoretical Considerations

Motivation-Gated Reinforcement Learning of Aesthetic Values

The starting point for the analyses in this article is the theoretical framework of Aleem et al. (2020). The core of the framework is reinforcement learning. As it is normal for reinforcement learning, the system first receives inputs from the external world, that is, the sensory inputs. Moreover, the system receives internal inputs on the motivation to act. The system then uses these external and internal inputs to estimate what will be the expected reward during the period in which these signals are arriving. This estimate is commonly referred to as value. When rewards arrive, they are compared with the values (i.e., the estimated rewards). If there is a mismatch (i.e., non-zero regret), the system learns by updating the parameters of the internal model (the value function). This update allows the system to achieve its goal of producing better reward predictions in the future.

While reinforcement learning is at the heart of the theoretical framework, it has four notable extensions. First, the estimate of reward itself is equivalent to aesthetic value. Second, the reinforcement-learning circuitry includes the concept of motivation within, which, by our definition, refers to the internal drive of an individual to act given an input. More specifically, motivation is a component of the likelihood of a person to act, which in turn is akin to policy in Machine Learning (Sutton and Barto, 2018). Third, both motivation and sensory inputs to the theoretical framework are probabilistic. Fourth, the inputs to our theoretical framework depend not only on individuals but also across societies.

Linear and Nonlinear Value Functions

In this article, we investigate the performance of aesthetic learning with four types of value function. First, we probe the linear value function, which yields an estimate of reward that is proportional to the sensory inputs. The constants of proportionality, which Aleem et al. (2020) call aesthetic weights, are the free parameters that the process of learning should estimate. Second, we follow the linear step with a saturation function characteristic of many neurobiological processes (Hudspeth et al., 2013; Schultz, 2015). Such saturation function added to the output of the linear function models a value-function nonlinearity resulting from diminishing marginal utility (Kreps, 1990). We call this mechanism the Output-saturation model because we apply the saturating process at the output of the linear stage. Third, we apply the same saturation mechanism to each component of the linear model. Appropriately, we call this mechanism the Component-saturation model. Fourth, we use the value function developed by Aleem et al. (2020) in their theoretical framework for aesthetic learning.

Update Rules for Value Functions

In the Aleem et al. article, the updates of the value function are performed with the delta rule (Sutton and Barto, 2018). This rule implements a gradient descent on the magnitude of

regrets (errors) of the predictions of reward. The delta rule stipulates that the change of the free parameters of the value function should be proportionate to the difference between observed and predicted rewards, typically denoted δ . Thus, the larger this difference is, the faster this change becomes. In all the simulations and mathematical analyses in this article, this component of the delta rule applies. Furthermore, the delta rule prescribes in what direction the vector of free parameters of the value function should change (Here, we often use “free parameters” when referring to the vector of free parameters of the value function). This change should be in the direction opposite to the gradient of the value function with respect to this vector. If the value function is linear, then this gradient is equal to the vector of sensory stimuli (Sutton and Barto, 2018).

However, the standard delta rule has some disadvantages, suggesting an important modification. To understand these disadvantages, let us start with some of the advantages of this rule. The first worth mentioning is that it attempts to minimize regret. This minimization holds for both standard reinforcement learning (Sutton and Barto, 2018) and the version here with motivation gating (Aleem et al., 2020). In addition, for the linear value function, the delta rule tends to optimize the trajectory of the free parameters (Aleem et al., 2020). However, as we will illustrate in the “Hypotheses Tested in This Article” sections, this advantage does not apply in general to nonlinear value functions. Fortunately, a related rule that has this advantage does exist. This new rule points the trajectory of the free parameters directly to the closest point in the isoline corresponding to the reward received (the target isoline). Because this rule takes the vector through the shortest route, we say that the rule implements the Shortest-path strategy. We sometimes also call this the Phi rule because the vertical line in Φ bisects its ellipse with the shortest path.

Hypotheses Tested in This Article

In this article, we probe the performance of learning under various value functions (“Linear and Nonlinear Value Functions” section) and their various update rules (“Update Rules for Value Functions” section). At the simplest level, the expectations for these probes are straightforward. For example, an update rule appropriate for a linear value function should do poorly with a nonlinear one. However, we wish to develop expectations that are more granular for the various value functions and update rules. **Figure 1** helps us formulate hypotheses based on these rules and functions.

From **Figure 1**, if one disregards the stochastic nature of the learning process, we can draw the following seven hypotheses about the interactions between values functions and their update rules:

- I. Assume that the value function is linear and the update delta rule follows the gradient at the position of the vector of free parameters. The final regret should be zero and the update convergence should be fast. After convergence, the recovery from fluctuation errors should also be fast.

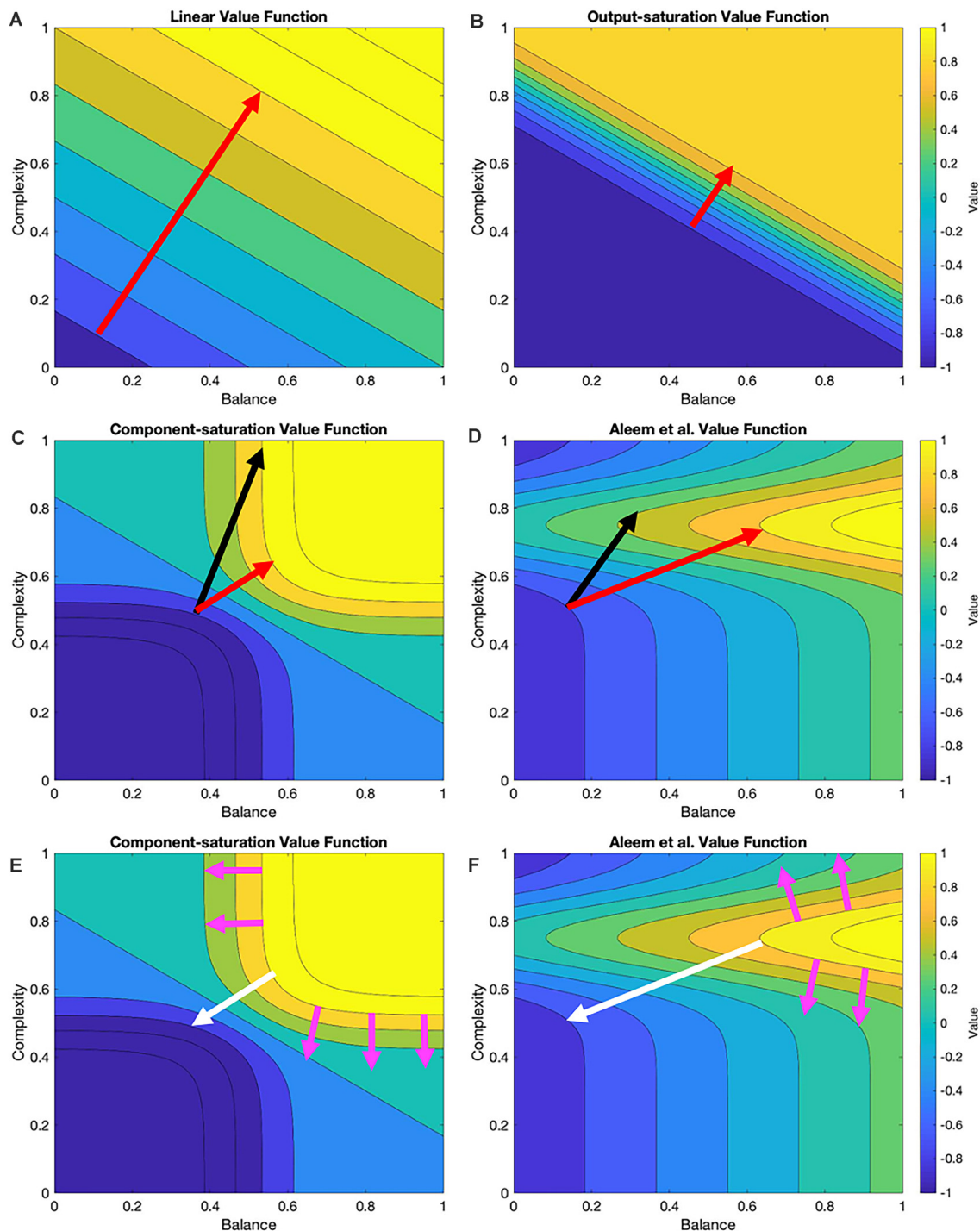


FIGURE 1 | Contour plots of four examples of value functions and relationships to possible update rules. These value functions are two-dimensional in this research ("Stochastic Sampling" section) but multi-dimensional in general. In the examples given here, the two dimensions are measures of balance and complexity in visual images. These measures range from 0 to 1 for the value functions illustrated in this research. The parameters of the value functions are those of **Table 2**. The red arrows indicate the optimal trajectory from the current position of the vector of free parameters of the value function to the closest point of the isoline corresponding to the sampled reward (Here, we call this curve the target isoline, but in general, it is an isosurface.) The black arrows indicate the trajectory based on gradient computation. **(A,B)** For the Linear and Output-saturation value functions, the gradient and optimal trajectories coincide. **(C,D)** For the Component-saturation and Aleem et al. models, the gradient trajectory is not optimal. **(E,F)** However, if one computed the gradients from the target isoline instead of the current position (magenta and white arrows), the gradient at the optimal point on the target isoline would be parallel to the optimal trajectory ("Results" section; white arrows). We call this computation the Shortest-path or Phi rule ("Update Rules for Value Functions" section).

- II. A similar hypothesis applies to the Output-saturation value function because of its straight and parallel contour lines.
- III. We hypothesize that the regret magnitude should be larger with the Component-saturation and Aleem et al. value functions than with the Linear and Output-saturation ones. The convergence and the recovery from fluctuation errors should also be slower for the former two. These problems will be especially acute for the Aleem et al. value function.
- IV. Similarly, we hypothesize a straight trajectory for the Linear and Output-saturation value functions (except for small stochastic fluctuations). But the trajectory should be curved for the other two functions.
- V. Instead, the Phi Rule should yield no regret and fast convergence and recovery from fluctuation errors for all value functions.
- VI. Regardless of the rule, although value will reach a unique fixed point, the free parameters will not. The reason for the lack of uniqueness is that many parameter combinations yield the same value (isolines in **Figure 1**).
- VII. Let a parameter of the model of reward have higher sensitivity coefficient than another parameter. Thus, if we increase the former parameter, we get more reward than if we raise the latter by a similar amount (Vidal et al., 1966; Saltelli et al., 2008). The corresponding free parameters of the value function should exhibit the same hierarchy of contributions to the estimation of reward.

However, if one does not disregard the stochastic nature of the learning process, these hypotheses could be wrong. With the stochastic sampling, the contour plots in **Figure 1** would change across samples, possibly making the convergence more complex. The computer simulations in the “Results” section test this possibility.

Mathematical Description of the Theoretical Considerations

Motivation-Gated Reinforcement Learning of Aesthetic Values

Much of the work described in this section appears in Aleem et al. (2020). We will only sketch the relevant work in that work here, leaving details to that article but pointing out the new ideas in this article.

Let the sensory inputs be a N dimensional vector, $\vec{u}(t)$, with the various components u_i corresponding to variables that the brain uses to represent the external world. Without loss of generality, Aleem et al. assumed that $0 \leq u_i \leq 1$. Moreover, and more importantly, Aleem et al. assumed that the value function was linear. Instead, we assume a general value function

$$v(t) = m(t)\mu(\vec{u}(t):\vec{w}(t)), \quad (1)$$

where $0 \leq m \leq 1$ is the motivation function, μ is a general nonlinear differentiable function representing the fully motivated value, and $\vec{w}(t)$ is the vector of free parameters of the value function (In this research, we use the colon to indicate parameters and thus, $\mu(\vec{u}(t):\vec{w}(t))$ means that the function μ has \vec{u} as variables and \vec{w} as parameters. The reason \vec{w} varies with time is that learning operates by parametric optimization.) Thus,

if we interpret m as the probability of acting around time, then the expected received reward is

$$r(t) = m(t)r^*(t), \quad (2)$$

where r^* is the reward that a fully motivated person would get.

Aleem et al. (2020) used a delta-rule update of the value function by first computing

$$\delta(t) = r(t) - v(t). \quad (3)$$

They then used the gradient update rule assuming a linear value function. We instead must use the value function in Equation (1), which yields

$$\frac{d\vec{w}(t)}{dt} = \varepsilon \delta(t) \nabla_{\vec{w}} \mu(\vec{u}(t):\vec{w}(t)), \quad (4)$$

where $\varepsilon > 0$ is a constant.

To complete the theoretical framework, we need to specify the statistical properties of $\vec{\mu}$, m , and r^* . Following Aleem et al., we define the probability density functions

$$P(\vec{I}_u|\vec{B}), P((\vec{u}(t), r^*(t))|\vec{I}_u), \quad (5)$$

where \vec{B} indicates the vector of parameters characteristic of the social and environmental background under consideration and \vec{I}_u is the vector of parameters of an individual in this society. We also define the probability density function of m as

$$P(\vec{I}_m|\vec{B}), P(m(t)|\vec{u}(t), \vec{I}_m), \quad (6)$$

where we insert \vec{B} to indicate that individual motivation may depend on environmental and social backgrounds.

Linear and Nonlinear Value Functions

For reinforcement learning to work well, the value function should be able to capture the structure of the incoming rewards. From Equations (1–3), (5) and (6), the expected least-square error (dropping both the dependence on t and the parameters for the sake of conciseness) is

$$E = \int \int \int \vec{u}, r^*, m P(\vec{u}, r^*) P(m|\vec{u}) (m(\mu(\vec{u}) - r^*))^2. \quad (7)$$

This error is a function of the value function $\mu(\vec{u})$ (Riesz and Szökefalvi-Nagy, 1990). As shown in Appendix: Optimal Value Function, the minimal of this function occurs when

$$\mu_{opt}(\vec{u}:\vec{w}, \vec{k}) = \langle r^* \rangle (\vec{u}:\vec{I}_u = [\vec{w}, \vec{k}]), \quad (8)$$

where $\langle r^* \rangle (\vec{u}:\vec{I}_u = [\vec{w}, \vec{k}])$ indicates the mean of r^* given the sampled sensory inputs, and the free (\vec{w}) and constant (\vec{k}) parameters of the value function.

We are now ready to specify the optimal value functions obtained after setting the mean rewards in our models.

Linear Value Function

$$\begin{aligned} \langle r_{lin}^* \rangle (\vec{u}:\vec{I}_u = \vec{w}^{(lin)}) &= \vec{w}^{(lin)} \cdot \vec{u}, \\ \mu_{lin}(\vec{u}:\vec{w}) &= \vec{w} \cdot \vec{u}, \end{aligned} \quad (9)$$

where $\vec{w}^{(lin)}$ are constant parameters of the model of reward.

Output-Saturation Value Function

$$\langle r_{out}^* \rangle (\vec{u}; \vec{I}_u = [\vec{w}^{(out)}, \vec{k} = [\alpha_1, \beta_1]]) = \frac{e^{\alpha_1(\vec{w}^{(out)} \cdot \vec{u} - \beta_1)} - 1}{e^{\alpha_1(\vec{w}^{(out)} \cdot \vec{u} - \beta_1)} + 1},$$

$$\mu_{out}(\vec{u}; \vec{w}, \alpha_1, \beta_1) = \frac{e^{\alpha_1(\vec{w} \cdot \vec{u} - \beta_1)} - 1}{e^{\alpha_1(\vec{w} \cdot \vec{u} - \beta_1)} + 1}, \quad (10)$$

where $\vec{w}^{(out)}$, $\alpha_1 > 0$, and β_1 are constant parameters of the model of reward. The right-hand side of Equation (10) is the hyperbolic tangent, a sigmoidal function centered on β_1 and with speed of rise controlled by α_1 .

Component-Saturation Value Function

$$\langle r_{com}^* \rangle (\vec{u}; \vec{I}_u = [\vec{w}^{(com)}, \vec{k} = [\alpha_{2,1}, \beta_{2,1}, \alpha_{2,2}, \beta_{2,2}]])$$

$$= \sum_{i=1}^N \frac{e^{\alpha_{2,i}(w_i^{(com)} \cdot u_i - \beta_{2,i})} - 1}{e^{\alpha_{2,i}(w_i^{(com)} \cdot u_i - \beta_{2,i})} + 1},$$

$$\mu_{com}(\vec{u}; \vec{w}, \alpha_{2,1}, \beta_{2,1}, \alpha_{2,2}, \beta_{2,2})$$

$$= \sum_{i=1}^N \frac{e^{\alpha_{2,i}(w_i \cdot u_i - \beta_{2,i})} - 1}{e^{\alpha_{2,i}(w_i \cdot u_i - \beta_{2,i})} + 1}, \quad (11)$$

where again, $\vec{w}^{(com)}$, $\alpha_{21} > 0$, β_{21} , $\alpha_{22} > 0$, and β_{22} are constant parameters of the model of reward. In turn, $\vec{w}^{(com)}$, w_i , and u_i are the i th components of the vectors $\vec{w}^{(com)}$, \vec{w} , and \vec{u} respectively.

Aleem et al. Value Function

$$(\vec{u}; \vec{I}_u = [\vec{w}^{(ale)}, \vec{k} = [\alpha_3, \beta_3]])$$

$$= -w_1^{(ale)} + 2w_1^{(ale)}u_1 - w_2^{(ale)}\theta(\alpha_3, \beta_3)$$

$$+ w_2^{(ale)}e^{-\frac{(u_2 - \alpha_3)^2}{2\beta_3^2}},$$

$$\mu_{ale}(\vec{u}; \vec{w}, \alpha_3, \beta_3)$$

$$= -w_1 + 2w_1u_1 - w_2\theta(\alpha_3, \beta_3) + w_2e^{-\frac{(u_2 - \alpha_3)^2}{2\beta_3^2}}. \quad (12)$$

where again, $\vec{w}^{(ale)}$, $\alpha_3 > 0$, and β_3 are constant parameters of the model of reward. In turn, $w_i^{(ale)}$ is the i th components of the vector $\vec{w}^{(ale)}$. Finally, the function θ is (α_3, β_3) is

$$\theta(\alpha_3, \beta_3) = \int_0^1 e^{-\frac{(u_2 - \alpha_3)^2}{2\beta_3^2}} du_2.$$

The derivation of Equation (12) follows from the equations in Aleem et al. (2020).

Update Rules for Value Functions

We use two update rules for the free parameters, with the first being the gradient-based delta rule in Equation (4). To implement this rule, we must first sample $\vec{u}(t)$, $r^*(t)$, and $m(t)$ from Equations (5) and (6) (details in “Materials and Methods” section). From, these samples, we can compute the value functions as in the second part of Equations (9)–(12) and thus, $\delta(t)$. Finally, we must compute the gradient, $\nabla_w \mu(\vec{u}(t); \vec{w}(t))$, for these value functions.

The second update rule that we use in this article is what we call the Phi (or Shortest-path) rule (Figures 1E,F). To define this rule, we begin by considering

$$\{\vec{w}_r(t) | \mu(\vec{u}(t); \vec{w}_r(t)) = r^*(t)\}, \quad (13)$$

which is the set of all free parameters of the value function that yield the sampled reward. Thus, $\vec{w}_r(t)$ are the points of the target isolines in Figure 1. Now, define the optimal point in the target isoline, that is, the point closest to \vec{w} :

$$\vec{w}_{opt}(t) = \operatorname{argmin}_{\vec{w}_r(t)} \|\vec{w}_r(t) - \vec{w}(t)\|. \quad (14)$$

This point may not be unique, but the lack of uniqueness is rare (and one can break it with tiny random perturbations), and thus, we neglect it here. We now define the vector $\vec{\Phi}(t)$ as

$$\vec{\Phi}(\vec{w}(t); \vec{u}(t), r^*(t)) = \frac{\vec{w}_{opt}(t) - \vec{w}(t)}{\|\vec{w}_{opt}(t) - \vec{w}(t)\|}, \quad (15)$$

that is, the unit vector pointing from $\vec{w}(t)$ to $\vec{w}_{opt}(t)$. With $\vec{\Phi}(t)$ in hand, we propose a new rule instead that in Equation (4):

$$\frac{d\vec{w}(t)}{dt} = \varepsilon \delta(t) \vec{\Phi}(\vec{w}(t); \vec{u}(t), r^*(t)). \quad (16)$$

Hypotheses Tested in This Article

As mentioned in the “Update Rules for Value Functions” section, the gradient-based delta rule attempts to minimize regret. This minimization holds for both standard reinforcement learning (Sutton and Barto, 2018) and the version here with motivation gating (Aleem et al., 2020). In the latter study, the demonstration of the minimization of regret was for the linear value function (Equation 9). In Appendix: Minimization of Regret Under Optimal Value Functions and the Delta Rule, we extend the demonstration for nonlinear value functions in the presence of motivation. Specifically, we show that the rule in Equation (4) tends to perform a stochastic gradient descent on the error

$$E = \left\langle m(t)(r^*(t) - \mu(\vec{u}(t); \vec{w}(t)))^2 \right\rangle_t, \quad (17)$$

where $\langle \rangle_t$ stands for time average. Consequently, the rule in Equation (4) performs a gradient descent on the error of value weighed statistically by the motivation.

Another implication of the delta rule is that it tends to maximize the rate of convergence for the linear value function (Aleem et al., 2020). The delta rule also maximizes the rate of recovery from fluctuation errors after convergence. These maximizations are contingent on the gradient being perpendicular to the isolines. However as seen in Figures 1C,D, the gradient is not generally perpendicular to the isolines for nonlinear value functions.

These conclusions on the gradient-based delta rule underlie Hypotheses I–IV.

In contrast, the Shortest-path Phi rule overcomes the deficiencies of the gradient-based delta rule. The Phi rule does so by going directly to the optimal point, \vec{w}_{opt} , on the target isoline (Equations 13 and 16). Appendix:

Perpendicularity Condition Under the Phi Rule adds to this conclusion, demonstrating an important property of \vec{w}_{opt} :

$$\vec{w}_{opt} - \vec{w} \propto \nabla_w \mu(\vec{u}; \vec{w}_{opt}). \quad (18)$$

Consequently, the perpendicular of the target isoline through \vec{w}_{opt} is parallel to the vector connecting \vec{w} to \vec{w}_{opt} . This result extends the conclusion for the delta rule that it tends to maximize the rate of convergence of \vec{w} for the linear value function. The result also extends the conclusion that the delta rule tends to maximize the rate of recovery from fluctuation errors after convergence. These conclusions are now valid for nonlinear value functions if one uses the Phi rule.

MATERIALS AND METHODS

We tested the hypotheses of “Hypotheses Tested in This Article” sections through mathematical analyses and computer simulations. The “Simulated Conditions” section lists all the conditions (mixtures of value functions and update rules) simulated in this article. Then the “Methods for Computer Simulations” and “Stochastic Sampling” sections describe the technical details of the simulations. These sections are followed by a summary of the simulation procedures (“Summary of the Simulation Procedures” section) and the parameters of the simulations (“Standard Simulation Parameters” section). Finally, the “Statistics to Test the Hypotheses” section describes the statistics used to test the hypotheses. The detailed mathematical analyses are left to the appendices, but the results are explained at appropriate places in this article.

Simulated Conditions

This article compares the performance of various value functions and their update rules in the learning of aesthetic biases. Hence, we performed simulations combining conditions of value functions and update rules. The simulated conditions appear in **Table 1**.

The logic of these conditions is as follows: The 16 conditions are divided in sets of four, with the title indicated in the first column of this table. Every set includes all four types of reward model. In the first set, the value function is linear and the update rules assumes a gradient descent based on the linear value function. This set makes these assumptions despite the reward model not always being linear (Conditions 2–4). Because of the doubly linear assumptions, we call this set the Purely-linear conditions. In contrast, the second set assumes a value function matched to the reward models. However, the update rule continues to be linear and thus, we call this set the Mixed-linear conditions. Next is the set called the Full-gradient conditions. This is the only set respecting fully the reward models in both the value functions and the gradient-descent update rules. Finally, the Shortest-path conditions also have values functions respectful of rewards but use the Phi rule instead of the delta rule.

The main model in Aleem et al. (2020) corresponds to Condition 4.

Methods for Computer Simulations

We must simulate Equations (1–4) to implement the delta rule and Equations (1–3), and (16) for the Phi rule. Combining these two sets of equations, we get respectively

$$\begin{aligned} \frac{d\vec{w}(t)}{dt} &= \epsilon_\delta m(t) (r^*(t) - \mu(\vec{u}(t); \vec{w}(t))) \nabla_w \mu(\vec{u}(t); \vec{w}(t)), \\ \frac{d\vec{w}(t)}{dt} &= \epsilon_\Phi m(t) (r^*(t) - \mu(\vec{u}(t); \vec{w}(t))) \\ &\quad \vec{\Phi}(\vec{w}(t); \vec{u}(t), r^*(t)). \end{aligned} \quad (19)$$

We use possibly different ϵ_δ and ϵ_Φ to allow for a fair comparison between the convergence rates of the two processes, as explained in the “Standard Simulation Parameters” section. Equations (19) are stochastic differential equations (Aleem et al., 2020).

We simplify our simulations through a mean-field approximation of Equation (6):

$$\begin{aligned} \frac{d\vec{w}(t)}{dt} &= \epsilon_\delta \bar{m}(\vec{u}(t); \vec{I}_m) \\ &\quad (r^*(t) - \mu(\vec{u}(t); \vec{w}(t))) \nabla_w \mu(\vec{u}(t); \vec{w}(t)), \\ \frac{d\vec{w}(t)}{dt} &= \epsilon_\Phi \bar{m}(\vec{u}(t); \vec{I}_m) \\ &\quad (r^*(t) - \mu(\vec{u}(t); \vec{w}(t))) \vec{\Phi}(\vec{w}(t); \vec{u}(t), r^*(t)), \end{aligned} \quad (20)$$

where $\bar{m}(\vec{u}(t); \vec{I}_m)$ is the mean motivation as a function of the sensory inputs $\vec{u}(t)$ and parametric on \vec{I}_m (Aleem et al., 2020).

To approximate a solution to Equations (20), we must discretize time and sample, \vec{u} , m , and r^* for every t . We do this discretization as follows:

$$\begin{aligned} \vec{w}(t_{k+1}) &= \vec{w}(t_k) + \epsilon'_\delta \bar{m}(\vec{u}(t_{k+1}); \vec{I}_m) (r^*(t_{k+1}) \\ &\quad - \mu(\vec{u}(t_{k+1}); \vec{w}(t_k))) \nabla_w \mu(\vec{u}(t_{k+1}); \vec{w}(t_k)), \\ \vec{w}(t_{k+1}) &= \vec{w}(t_k) + \epsilon'_\Phi \bar{m}(\vec{u}(t_{k+1}); \vec{I}_m) (r^*(t_{k+1}) \\ &\quad - \mu(\vec{u}(t_{k+1}); \vec{w}(t_k))) \vec{\Phi}(\vec{w}(t_k); \vec{u}(t_{k+1}), r^*(t_{k+1})), \end{aligned} \quad (21)$$

where $\epsilon'_\delta = \epsilon_\delta (t_{k+1} - t_k)$ and $\epsilon'_\Phi = \epsilon_\Phi (t_{k+1} - t_k)$, with $t_{k+1} - t_k$ being constant (for $k = 0, 1, 2, \dots$).

In this article, we compute $\nabla_w \mu$ analytically. These gradients are relatively easy to compute, so we omit them here from the sake of space. As for the computation of $\vec{\Phi}$, we use the method of Marching Squares Algorithm to obtain the value isolines (Maple, 2003), and then apply Equations (14) and (15). We apply this algorithm to a 101×101 pixels approximation of the value function.

Stochastic Sampling

To simulate Equations (21), one must sample \vec{u} and r^* stochastically from the probability distributions in Equation (5), and compute $\bar{m}(\vec{u}(t); \vec{I}_m)$ for use in Equations (20). We follow Aleem et al. (2020) and take five steps to simplify the sampling to make the simulations fast. See Aleem et al. (2020) for more details and justifications.

- We did not simulate social “noise” by implementing explicitly $P(\vec{I}_u | \vec{B})$ and $P(\vec{I}_m | \vec{B})$, instead setting individual parameters by hand.

TABLE 1 | Conditions simulated.

Set	Condition	Reward	Value function	Update rule
Purely linear	1	$\langle r_{lin}^* \rangle$; Eq. (9)	μ_{lin} ; Eq. (9)	$\nabla_w \mu_{lin}$; Eqs. (4) and (9)
	2	$\langle r_{out}^* \rangle$; Eq. (10)	μ_{lin} ; Eq. (9)	$\nabla_w \mu_{lin}$; Eqs. (4) and (9)
	3	$\langle r_{com}^* \rangle$; Eq. (11)	μ_{lin} ; Eq. (9)	$\nabla_w \mu_{lin}$; Eqs. (4) and (9)
	4	$\langle r_{ale}^* \rangle$; Eq. (12)	μ_{lin} ; Eq. (9)	$\nabla_w \mu_{lin}$; Eqs. (4) and (9)
Mixed linear	5	$\langle r_{lin}^* \rangle$; Eq. (9)	μ_{lin} ; Eq. (9)	$\nabla_w \mu_{lin}$; Eqs. (4) and (9)
	6	$\langle r_{out}^* \rangle$; Eq. (10)	μ_{out} ; Eq. (10)	$\nabla_w \mu_{lin}$; Eqs. (4) and (9)
	7	$\langle r_{com}^* \rangle$; Eq. (11)	μ_{com} ; Eq. (11)	$\nabla_w \mu_{lin}$; Eqs. (4) and (9)
	8	$\langle r_{ale}^* \rangle$; Eq. (12)	μ_{ale} ; Eq. (12)	$\nabla_w \mu_{lin}$; Eqs. (4) and (9)
Full gradient	9	$\langle r_{lin}^* \rangle$; Eq. (9)	μ_{lin} ; Eq. (9)	$\nabla_w \mu_{lin}$; Eqs. (4) and (9)
	10	$\langle r_{out}^* \rangle$; Eq. (10)	μ_{out} ; Eq. (10)	$\nabla_w \mu_{out}$; Eqs. (4) and (10)
	11	$\langle r_{com}^* \rangle$; Eq. (11)	μ_{com} ; Eq. (11)	$\nabla_w \mu_{com}$; Eqs. (4) and (11)
	12	$\langle r_{ale}^* \rangle$; Eq. (12)	μ_{ale} ; Eq. (12)	$\nabla_w \mu_{ale}$; Eqs. (4) and (12)
Shortest path	13	$\langle r_{lin}^* \rangle$; Eq. (9)	μ_{lin} ; Eq. (9)	$\vec{\Phi}$; Eq. (16)
	14	$\langle r_{out}^* \rangle$; Eq. (10)	μ_{out} ; Eq. (10)	$\vec{\Phi}$; Eq. (16)
	15	$\langle r_{com}^* \rangle$; Eq. (11)	μ_{com} ; Eq. (11)	$\vec{\Phi}$; Eq. (16)
	16	$\langle r_{ale}^* \rangle$; Eq. (12)	μ_{ale} ; Eq. (12)	$\vec{\Phi}$; Eq. (16)

- B. We split the individual parameters \vec{I}_u into sensory related (\vec{I}_s) and reward related (\vec{I}_r):

$$\vec{I}_u = [\vec{I}_s, \vec{I}_r]. \quad (22)$$

- C. We made \vec{u} two-dimensional. One component was visual balance (u_b) and the other was visual complexity (u_c), making

$$\vec{u} = [u_b, u_c],$$

where $0 \leq u_b, u_c \leq 1$, as per the definitions elsewhere (Aleem et al., 2017). Thus, while our model is amenable to a range of sensory inputs, we simplified it to the visual domain for illustrative purposes. Accordingly, $N = 2$ in Equation (11), and $u_1 = u_b$ and $u_2 = u_c$ in Equation (12) and **Figure 1**.

- D. To be compatible with the two-dimensional \vec{u} and so that all value functions have the same number of free parameters, we have set the number of free parameters in each model of reward to two. The models in Equations (10–12) have other parameters, namely, $\alpha_1, \beta_1, \alpha_2, \beta_2, \alpha_3$, and β_3 . However, we treat them as constants, with values specified in the “Standard Simulation Parameters” section.

- E. We split the second term of Equation (5) as follows:

$$P((\vec{u}, r^*) | \vec{I}_u) = P(\vec{u} | \vec{I}_s) P(r^* | \vec{u}, \vec{I}_r). \quad (23)$$

With these simplifications in hand, we followed Aleem et al. for the sampling of \vec{u} through the first term of the right-hand side of Equation (23). We also followed them for the subsequent computation of $\vec{m}(\vec{u}(t); \vec{I}_m)$. For the sake of space, we refer the reader to their article (see their Equations 12, 13, 18, and 19).

Finally, we must specify how to sample r^* through the second term of the right-hand side of Equation (23). We model $P(r^* | \vec{u}, \vec{I}_r)$ as a Gaussian distribution with one of the means as in Equations (9–12):

$$P(r^* | \vec{u}; \vec{I}_r) = \left[\vec{w}^{(x)}, \vec{k} = [\vec{k}^{(x)}, \sigma_x] \right] \\ = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{(r^* - \langle r_x^* \rangle(\vec{u}; \vec{w}^{(x)}, \vec{k}^{(x)}))^2}{2\sigma_x^2}}, \quad (24)$$

where $x \in \{lin, out, com, ale\}$, and $\vec{w}^{(x)}, \vec{k}^{(x)}$, and $\sigma_x > 0$ are constant parameters.

Summary of the Simulation Procedures

The simulations proceed with the following algorithm:

- Suppose that at time t_k , the vector of free parameters is $\vec{w}(t_k)$.
- Sample $\vec{u}(t_{k+1}) = [u_b(t_{k+1}), u_c(t_{k+1})]$ from Equation (12) of Aleem et al. (2020).
- Sample $\langle r_x^* \rangle$ from Equation (24), with the definitions of $\langle r_x^* \rangle$ in Equations (9–12).
- Compute $\vec{m}(\vec{u}(t_{k+1}); \vec{I}_m)$ from Equation (18) of Aleem et al. (2020).
- Compute $\vec{w}(t_{k+1})$ from Equation (21).
- Start the process again at Step a but at time t_{k+1} .

See Aleem et al. (2020) for more details on this algorithm.

All simulations were performed with code specially written in MATLAB.

Standard Simulation Parameters

In this article, we report on simulations with different parameter sets to explore the various models. We have designated one of these sets as our standard set because the corresponding results capture the data in the literature reasonably well (Aleem et al., 2020). **Table 2** shows the parameters of the standard simulations. Parameters for other simulations are indicated as appropriate in the Results (“Results” section).

A parameter in this table merits special discussion, namely, $\epsilon_\Phi = 0.007454$. We chose this value to make the comparison of the convergence rates of the gradient delta rule and the Phi rule fair. Changes of \vec{w} in both rules are proportional to δ times a vector indicating the direction of change. In the delta rule, the vector is $\nabla_w \mu$ whereas in the Phi rule, the vector is $\vec{\Phi}$, with the latter being a unit vector, while the former possibly having variable magnitudes. To make the convergence rate fair, we wanted to make the magnitudes of $\epsilon_\delta \times \nabla_w \mu$ comparable to the magnitudes of $\epsilon_\Phi \times \vec{\Phi}$. We did so by obtaining the root mean square of the

TABLE 2 | Standard set of parameters.

Parameter(s)	Equation(s)	Values
$\bar{w}(t_0)$	(21)	[0, 0]
ϵ_δ	(21)	0.01
ϵ_Φ	(21)	0.007454
$t_{k+1} - t_k$	(21)	1
$[\bar{w}^{(lin)}, \sigma_{lin}]$	(9) and (24)	[0.6, 0.9, 0.1414]
$[\bar{w}^{(out)}, \alpha_1, \beta_1, \sigma_{out}]$	(10) and (24)	[1.2, 1.8, 10, 1.5, 0.1414]
$[\bar{w}^{(com)}, \alpha_{21}, \beta_{21}, \alpha_{22}, \beta_{22}, \sigma_{com}]$	(11) and (24)	[1.2, 1.8, 10, 0.6, 10, 0.9, 0.1414]
$[\bar{w}^{(out)}, \alpha_3, \beta_3, \sigma_{out}]$	(12) and (24)	[0.6, 1, 0.75, 0.1, 0.1414]

magnitude of $\epsilon_\delta \times \nabla_w \mu_{lin}$, which is 0.7454, and thus because $\epsilon_\delta = 0.01$, we got $\epsilon_\Phi = 0.007454$.

Statistics to Test the Hypotheses

All analyses comparing these statistics across the stimulated conditions (**Table 1**) used one-way ANOVA followed by *post-hoc* two-sided *t*-tests. For each of the Conditions 1–16, we ran 10 repetitions with 1,000,000 iterations each.

The statistics used to test our hypotheses (“Hypotheses Tested in This Article” sections) are summarized in **Table 3**.

To start the estimation of these statistics, we began by obtaining the fully motivated value curve obtained for the most common stimulus, namely, $\vec{u} = [u_b, u_c] = [0.5, 0.5]$ (Aleem et al., 2020). This curve was $v^*(t) = \mu([0.5, 0.5] : \vec{w}(t))$ [Equation (1)].

From this curve, we first estimated τ_c as the number of iterations needed for $v^*(t)$ to reach 90% of the median of $v^*(t)$ during the last 100,000 iterations. Similarly, we used these 100,000 iterations of $v^*(t)$ to estimate τ_r . This statistic was important because it determined how many iterations we had to consider to avoid correlated measurements of the variable under consideration. We estimated this statistic through the autocorrelation coefficient (Park, 2018), by measuring when it decayed to 0.1 and setting that time to τ_r . We also tested whether τ_c and τ_r were correlated across all the conditions in **Table 1**. For this purpose, we used the robust Kendall’s τ correlation coefficient (Bonett and Wright, 2000).

With τ_r in hand, we could proceed to measure the values of δ_f and \vec{w}_f . To measure these statistics, we obtained the medians of δ and \vec{w} respectively over the last $2 \times \tau_r$ iterations of each simulation. By considering $2 \times \tau_r$ iterations, we could make sure to have two sets of temporally independent measurements.

Finally, to measure ρ_- we first obtained the phase diagram of the free parameters, that is, $w_2(t)$ vs. $w_1(t)$. As we will see in the “Results” section, we can model the initial portion of this plot in our simulations as the straight line $w_2(t) = kw_1(t)$, where $k > 0$ is a constant, and $w_1(t), w_2(t) > 0$ for $t > 0$. We estimated this line by robust linear regression, using M-estimation with Tukey’s biweight function (Rousseeuw and Leroy, 2003) from all the iterations such that $t \leq \tau_c$. The plot then sometimes deviated from this line, meandering from it a certain distance. To measure the deviation from straightness, we used three points: $\vec{w}(t_0)$ (**Table 2**), $\vec{w}_f = [w_{f,1}, w_{f,2}]$ (**Table 3**), and the point $\vec{w}(t_n)$ in the line $w_2(t) = kw_1(t)$ that was nearest to \vec{w}_f . From these points, we defined the deviation from straightness as the signed ratio of the distance from \vec{w}_f to $\vec{w}(t_n)$ to the distance from

$\vec{w}(t_n)$ to $\vec{w}(t_0)$. The sign was positive if \vec{w}_f was above the line and negative otherwise. This definition using a signed ratio was valid because the denominator was always positive with our simulations. Straightforward calculus and algebra showed

$$\rho_- = \frac{w_{f,2} - kw_{f,1}}{kw_{f,2} + w_{f,1}}. \quad (25)$$

Consequently, $-\infty \leq \rho_- \leq \infty$, with $\rho_- = 0$ if and only if \vec{w}_f was on the initial straight line. Highly positive ρ_- meant that final aesthetic preferences had a strong bias towards complexity, whereas highly negative ρ_- meant a strong balance bias.

To test Hypothesis VI, we ran a one-way ANOVA on each of the components of \vec{w}_f over the 10 repetitions of each condition.

RESULTS

Limitations of the Purely-Linear Conditions

If the brain acquires aesthetic biases through reinforcement learning, neural circuitries implementing suitable value functions and update rules are necessary for good performance. We propose that good value functions and update rules depend on the statistics of sensory inputs, motivations, and rewards. Here, we focus on the latter. We do so because learning to predict rewards is the goal of the learning process. We thus built several models of reward, one linear and three nonlinear, and tested the learning performance of four value functions and three types of update rules (**Table 1**).

The simplest and thus, the most used combination of value function and update rule for reinforcement learning in the brain is purely linear (Conditions 1–4 in **Table 1**). Is learning performance with this combination good even when facing nonlinear reward models? **Figure 2** shows the results of simulations with this combination of value function and update rule. The figure includes the temporal progression of free parameters, their phase diagrams, and errors in the prediction of reward. The simulations are performed for the four types of reward model studied in this article.

In all the simulations, the free parameters rose rapidly initially (**Figures 2A–D**). This rise occurred because these free parameters correlated positively with reward (Aleem et al., 2020). However, for some conditions, the fast rise ended and one of the free parameters started to fall as the other continued to climb (**Figures 2A,C**). This apparent competition of free parameters eventually stopped and the simulations reached steady state. We will address the reason for this apparent competition in

TABLE 3 | Statistics used to test the hypotheses.

Symbol	Title	Hypotheses
τ_c	Time of convergence	I–III and V
τ_r	Time of recovery from fluctuation errors	I–III and V
δ_f	Final regret	I–III and V
\bar{w}_f	Final free parameters of the value function	VI
ρ_{-}	Deviation from straightness	IV

the “Failing Hypotheses: How Stochasticity Helps and Shapes Learning” section. The apparent competition was especially evident in the phase diagrams (**Figure 2E**). With apparent competition, the phase diagram first seemed to rise linearly and then meandered away from the straight line (see the Linear and the Aleem et al.’s reward models in **Figure 2E**).

The apparent competition between the free parameters was not reflected in the temporal dependence of values. They rose and reached a steady state without any inflection points (**Figure 7** of Aleem et al. —the results here were similar; data not shown). The lack-of-inflection point result is not surprising, because as shown in the “Update Rules for Value Functions” section, although free parameters do not statistically reach a unique fixed point, values do (Aleem et al., 2020). Furthermore, the delta rule used in these simulations tends to minimize value regret in a gradient-decent manner (“Update Rules for Value Functions” section; Appendix: Minimization of Regret Under Optimal Value Functions and the Delta Rule). Hence, values monotonically approach optimal results, even if the free parameters display strange behaviors.

Going back to the temporal plots, we almost always observed the free parameter of complexity being larger than that of balance in these simulations (**Figures 2A,B,D**). This advantage of complexity was not surprising. We set up the simulations such that the fixed parameters of complexity made it contribute more to reward than those of balance (**Table 2**). However, when the reward model used the Component-saturation nonlinearity, the opposite happened and balance won (**Figures 2C,E**). The plots of regret provided further evidence of the inadequacy of the Purely-linear conditions (**Figure 2F**). Only when the reward model was linear did the final regret stay near zero. For all nonlinear reward models, the final regret was significantly negative (overestimation of reward).

To quantify the performance of the Purely-linear conditions, we measured the five statistics indicated in **Table 3**. The first statistic, regret (δ_f), indicated the overall error of the estimation of reward after the learning process had converged. Next, the time of convergence (τ_c), estimated how long the learning process took to converge. A related statistic was τ_r , which captured how long the learning process took to recover from a fluctuation error. In turn, the deviation from straightness (ρ_{-}), captured how directly the learning trajectory went to the final goal. Finally, we measured with \bar{w}_f where the free parameters converged at the end of the simulation. These results are summarized in **Figure 3**.

As expected, the magnitudes of the final regrets were large when using the Purely-linear strategy with nonlinear reward models (**Figure 3A**). These regrets were negative (overestimation of reward). However, the regrets were not significantly different

from zero for the linear reward model ($\delta_f = -0.0001 \pm 0.0001$; mean \pm standard error). Although the regrets were statistically different from each other (one-way ANOVA and *post-hoc* two-sided *t*-test), the times of convergence were roughly similar ($\tau_c \approx 1,400$ iterations—**Figure 3B**). Likewise, the times of recovery of fluctuation errors were roughly comparable ($\tau_r \approx 1,200$ iterations—**Figure 3B**). The times of recovery exhibited a strong positive correlation with the times of convergence across all the conditions of **Table 1** (**Figure 3B**; Kendall’s $\tau = 0.93$, $p < 4 \times 10^{-10}$). As for deviations from straightness, all but the Output-saturation reward yielded results significantly different from zero (**Figure 3C**). These deviations were positive (advantage to complexity) or negative (advantage to balance). Interestingly, the Purely-linear simulations deviated from zero even for the linear reward model ($\rho_{-} = 0.084 \pm 0.004$; $t = 20.0$, 9 d. f., $p < 1 \times 10^{-8}$).

In conclusion, the simulations with the Purely-linear conditions rule out Hypothesis I (“Hypotheses Tested in This Article” sections). This hypothesis fails because of the non-zero final regrets observed despite using a linear value function. We also rule out Hypothesis IV, since the linear value function yielded curved trajectories for all but the Output-saturation reward model. Finally, the inversion of complexity and balance preferences in **Figure 2C** rules out Hypothesis VII. On **Table 2**, the parameters of complexity are larger than those of balance, making the sensitivity coefficients for the former larger than for the latter. Therefore, Hypothesis VII would predict complexity preferences to be always larger than those for balance.

Simulations With the Mixed-Linear Conditions

Using a linear value function tends to lead to a poor learning performance when the reward model is nonlinear (**Figures 2, 3**), but does the outcome improve when one uses the appropriate nonlinear value function? Would we observe an improvement even if the update rule continues to be linear? To answer these questions, we performed the simulations for the Mixed-linear conditions (**Table 1**). **Figure 4** shows the results of these simulations. These results are important because they address the Hypotheses II and III in the “Hypotheses Tested in This Article” sections.

A comparison of **Figure 4** with **Figure 2** revealed that the Purely and Mixed-linear conditions yielded qualitatively, but not quantitatively, similar learning performances. The ordering of the free-parameter curves (**Figures 4A–D**) were largely similar for the two sets of conditions. So were the shapes

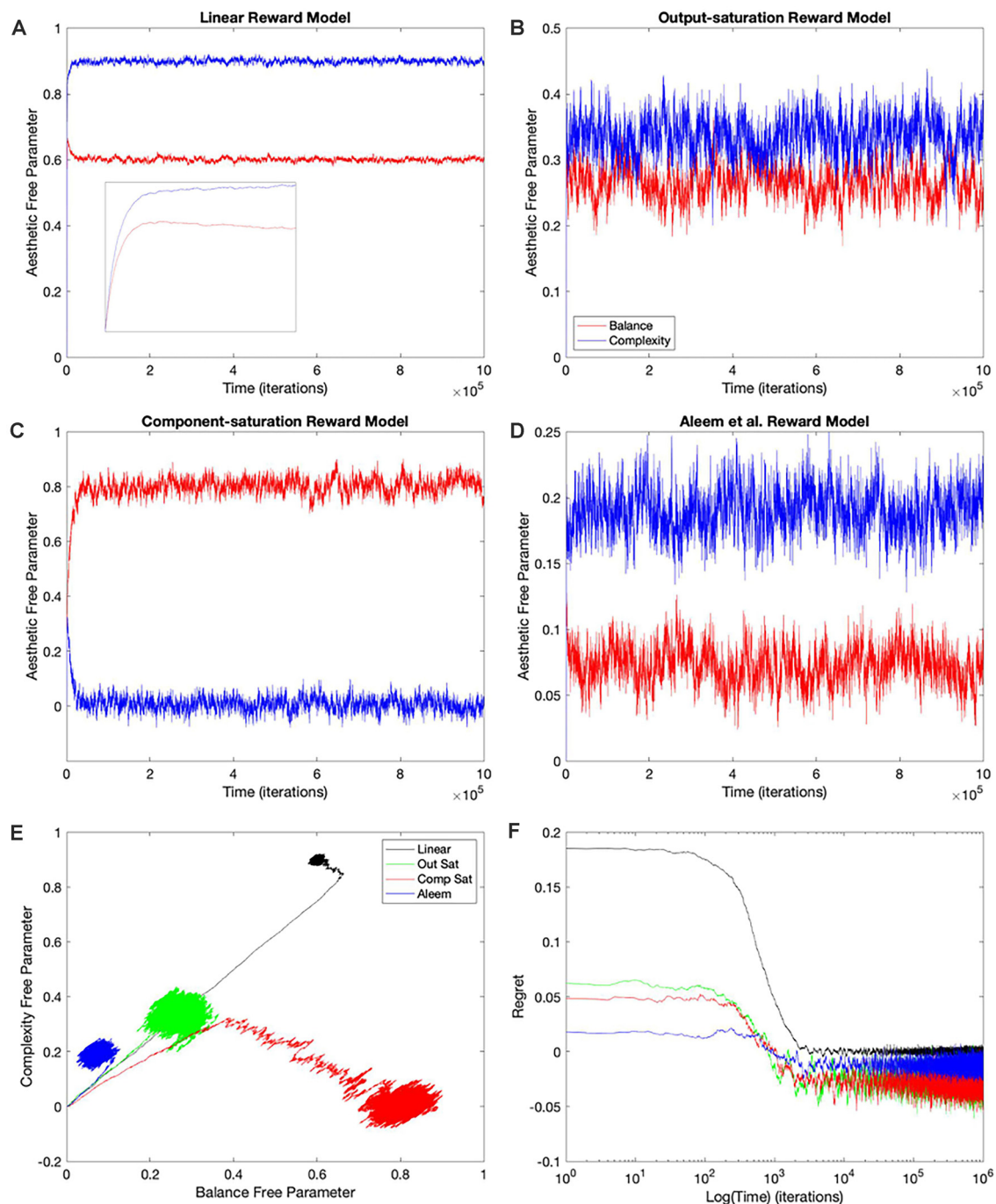


FIGURE 2 | Dynamics of the free parameters of the value function for the purely-linear conditions (Table 1). (A–D) Linear, Output-saturation, Component-saturation, and Aleem et al.’s reward models respectively. Red and blue lines correspond respectively to free parameters related to balance and complexity in the sensory inputs. The inset in (A) provides details of the early dynamics (first 10,000 iterations). (E) Phase diagrams. (F) Time dependence of regrets, smoothed with a 500-iterations moving average. This figure indicates that linear value functions and update rules yield poor learning performance when the reward models are nonlinear. For example, regrets are significantly negative (overestimation of reward) for all nonlinear reward models (Panel F).

of the phase diagrams (Figure 4E) and the regret behaviors (Figure 4F). This similarity included the surprising “error” in ordering for the behavior of the Component-saturation curves (Figure 4C). However, the final free parameters were smaller

for the Saturation reward models and larger for Aleem et al. in the Mixed-linear conditions. In addition, the magnitudes of final regrets were smaller. Figure 3A quantifies the improvement of the final regret for the Aleem reward model. In contrast,

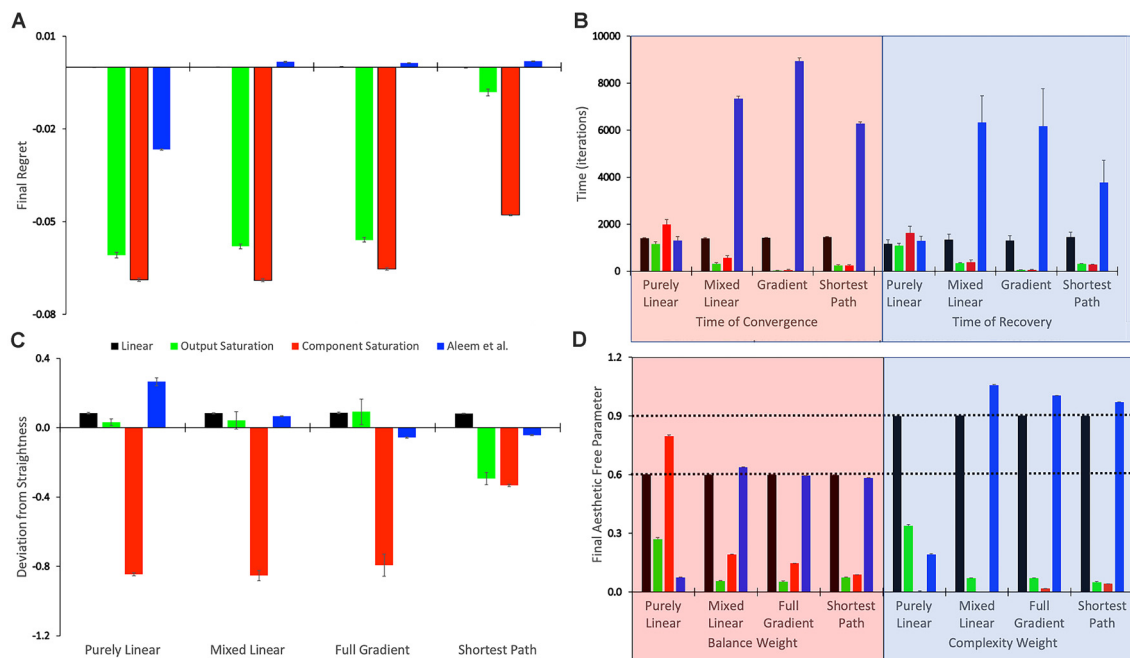


FIGURE 3 | Statistics of the tested conditions (Table 1). (A) Final regret. (B) Times of convergence and of recovery from fluctuation errors. (C) Deviation from straightness. (D) Final free parameters of the value functions. Each of these statistics was measured 10 times for each of the reward models, with the means and standard errors displayed. Color bins indicate the different reward models (colors matched to Figures 2E,F). We group the 16 conditions in sets of four according to the experimental conditions (Table 1). These sets were the purely linear, mixed linear, full gradient, and shortest path. The sets appear twice in Panel (B), for time of convergence (transparent red) and time of recovery from fluctuation errors (transparent blue). Similarly, the sets appear twice in Panel (D), for balance (transparent red) and complexity (transparent blue) free parameters. The dotted horizontal lines indicate the parameters of the linear reward models.

both Saturation reward models did not show statistically significant changes in terms of regret. Surprisingly, however, the time of convergence became faster for the Saturation reward models ($\tau_c \approx 430$ iterations) and slower for Aleem et al. reward $\tau_c = 7,300 \pm 100$ iterations (Figure 3B). The times of recovery from fluctuation errors exhibited similar results (Figure 3B). Finally, the magnitude of deviations from straightness fell for the Aleem et al.'s reward model (Figure 3C).

We conclude that Hypothesis II is also not valid. It fails because the Mixed-linear conditions include the Output-saturation value function, which yields no improvement in the final regret. Moreover, we can reject Hypothesis III because the magnitude of final regret for the Aleem et al. value function is smaller than for the Linear one. However, the slowness of both convergence and recovery from fluctuation errors with the Aleem et al. value function is predicted by the second part of Hypothesis III. Similarly, the straightness of the trajectory with the Output-saturation value function supports the second part of Hypothesis IV. The curvatures with the Component-saturation and Aleem et al. value functions also do so.

Simulations With the Full-Gradient Conditions

Why does the Mixed-linear conditions not improve the performance with the Output and Component-saturation reward

models despite using the appropriate value functions? Is the failure due to the use of an inappropriate (linear) update rule? A simple way to answer these questions is to implement the gradient update fully in the simulations. This is exactly what the Full-gradient conditions of Table 1 aim to achieve. The results of the simulations with these conditions appear in Figure 5.

The learning performances in Figure 5 were like those in Figure 4. The only apparent changes in Figure 5 were more noise in the Saturation conditions, and closer final free parameters of balance and complexity for Component Saturation (Figure 5C). However, inspection of the statistics in Figure 3 revealed small but significant improvements with the Full-gradient conditions. For example, the final regret improved slightly for the Aleem et al. function from $\delta_f = 0.0017 \pm 0.0001$ to $\delta_f = 0.0013 \pm 0.0001$ ($t = 2.26$, 18 d. f., $p < 0.04$). The statistics also revealed faster times to convergence for the Saturation value functions ($\tau_c \approx 40$ iterations; $t = 4.64$, 18 d. f., $p < 3 \times 10^{-4}$ for Output Saturation). The times of recovery from fluctuation errors again exhibited similar results. In terms of deviation from straightness, the notable result was the change of sign for the Aleem et al. value function. The deviation from straightness changed from $\rho_{-} = 0.066 \pm 0.002$ to $\rho_{-} = -0.057 \pm 0.003$.

Consequently, employing the appropriate update rules in a gradient-based delta-rule model helps the learning performance, but the effects are minor.

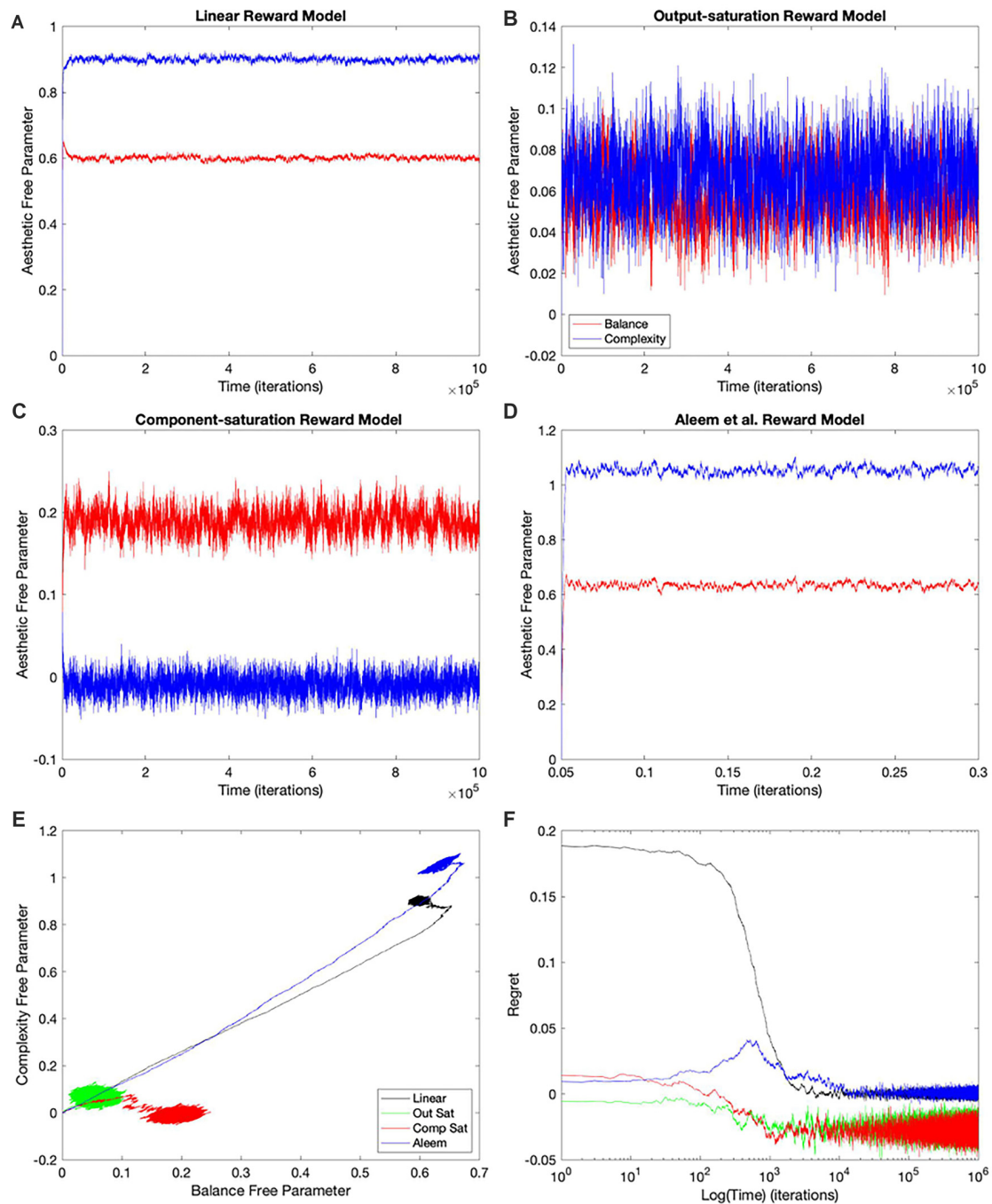


FIGURE 4 | Dynamics of the free parameters of the value function for the Mixed-linear conditions (**Table 1**). The conventions for this figure are the same as those for **Figure 2**. The results here are qualitatively like those in **Figure 2**, but some significant quantitative differences are readily apparent. The panels and conventions for this figure are the same as those for **Figure 2**.

Improved Performance With the Shortest-Path (Phi Rule) Conditions

Even with the Full-gradient conditions, the learning performance is still wanting (**Figure 3**), especially for nonlinear reward models. **Figure 1** provides a possible explanation for the deficiency of performance based on gradient-based delta rules.

The gradient is taken at the position of the vector of free parameters. Therefore, the direction of the gradient is generally blind to the curvatures of the isolines of the value function (**Figures 1C,D**). We have then proposed a new update rule that bypasses this deficiency of the gradient-based delta rule. If the value function is known, a calculation

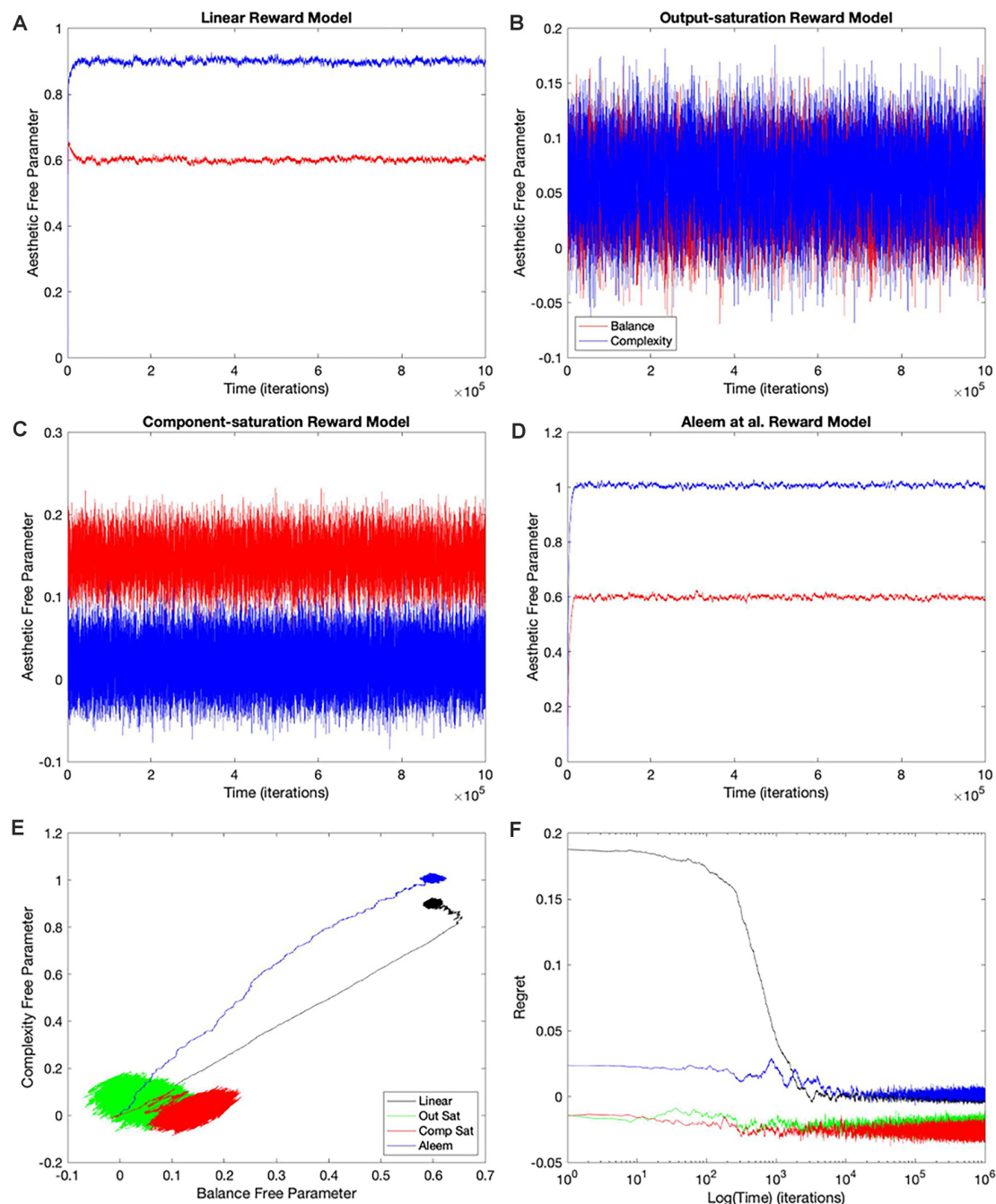


FIGURE 5 | Dynamics of the free parameters of the value function for the full-gradient conditions (Table 1). The conventions for this figure are the same as those for Figure 2. The results here are qualitatively like those in Figure 4, with only minor quantitative differences being easily observable. The panels and conventions for this figure are the same as those for Figure 2.

can be performed of the direction minimizing the path from the vector of free parameters to the target isoline (Figures 1E,F). We have called this update rule the Shortest-path or Phi rule (“Update Rules for Value Functions” section). The results of the simulations with this new rule appear in Figure 6.

Figure 6 shows that the Shortest-path (Phi) update rule produces superior performance when compared to the Full-gradient delta rule (Figure 5). The best evidence for the improved performance is that the magnitudes of final regrets are smaller with the Phi rule than with the delta rule (red curves in Figures 5C, 6C). This is confirmed in

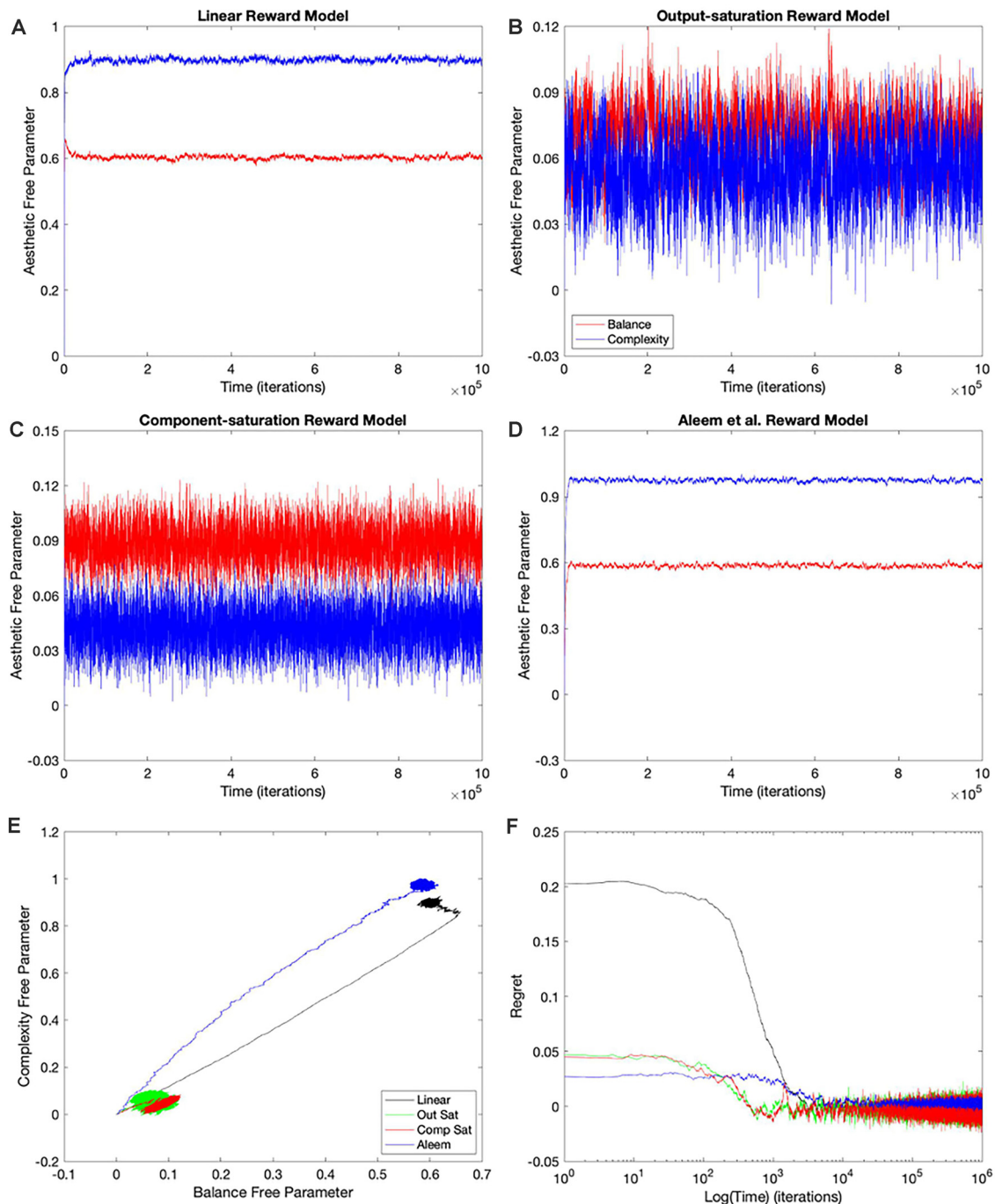


FIGURE 6 | Dynamics of the free parameters of the value function for the shortest-path (Phi Rule) conditions (Table 1). The conventions for this figure are the same as those for Figure 2. The free-parameter and phase-diagram results here are qualitatively like those in Figure 5, with minor apparent quantitative differences (Panels A–E). However, although the changes in the free-parameter curves are subtle, the improvement of the regrets are dramatic (Panel F).

Figure 3A, especially for the Saturation conditions. The magnitude of the deviation from straightness also fell for the Component-saturation condition (Figure 3C; $t = 7.23$, 18 *d. f.*, $p < 2 \times 10^{-6}$). Furthermore, this deviation fell for the Aleem et al. value functions ($t = 3.76$, 18

d. f., $p < 0.002$). Finally, the time of convergence fell for Aleem et al. value function from $\tau_c = 8,900 \pm 100$ to $\tau_c = 6,290 \pm 70$ iterations (Figure 3). The time of recovery from fluctuation errors also exhibited similar results (Figure 3B).

In conclusion, the Shortest-path (Phi) rule leads to superior learning performance as compared to the delta rule. However, the performance is not perfect. Imperfections include the small but non-zero regrets, small but significant deviations from straightness, and the relatively slow convergence and recovery for the Aleem et al. value function. Hence, the results reject Hypothesis V that states that the Phi Rule should yield no regret, and fast convergence and recovery from fluctuation errors.

Non-uniqueness of the Learned Free Parameters

Hypothesis VI predicts that regardless of the update rule for the value function, the value reaches a unique fixed point (albeit only statistically), but the free parameters do not. The reason for the lack of uniqueness is that many parameter combinations yield the same value (isolines in **Figure 1**). To test this non-uniqueness hypothesis, we have inspected the statistics of the final free parameters of the simulations. The statistics appear in **Figures 3D, 7**, which shows box plots for each of the 10 individual simulations in some of the conditions in **Table 1**.

The statistics in **Figure 3D** initially suggested that at least for some conditions, the free parameters converged statistically to a unique fixed point. For example, the linear value function, which we repeated over the four sets of conditions, yielded final estimated parameters indistinguishable from those of the reward function (dotted horizontal lines in **Figure 3D**). The estimated value-function parameters for the delta rule ($N = 30$) were $w_{f,1} = 0.5999 \pm 0.0007$ and $w_{f,2} = 0.9005 \pm 0.0008$. In turn, the estimated value-function parameters for the Phi rule ($N = 10$) were $w_{f,1} = 0.598 \pm 0.001$ and $w_{f,2} = 0.900 \pm 0.001$. These estimated value-function parameters were statistically the same as the reward parameters, which were $\vec{w}^{(lin)} = [0.6, 0.9]$ (**Table 2**).

However, closer inspection of the data reveals that the free parameters do not converge statistically to a unique fixed point. **Figure 7** illustrates this conclusion with four examples of conditions in **Table 1**. (However, the conclusion applies to all conditions—data not shown). In these examples, we focus on the final balance free parameter and break down the results into the 10 simulations that give rise to each bin of **Figure 3D**. The first example to comment here is the one described in the last paragraph). As the **Figure 7A** shows, although the final balance free parameter hovers close to 0.6 ($\approx 2.5\%$ variation), the outcomes of the different simulations are not statistically homogeneous (one-way ANOVA, $F = 5,960$, 9 numerator d.f., 26,080 denominator d.f. $p < 10^{-15}$). This inhomogeneity is not due to autocorrelations of the value signal (“Statistics to Test the Hypotheses” section). In addition, the inhomogeneity is applicable if one uses the Phi instead of the delta rule (**Figure 7B**, $\approx 1.5\%$ variation, $p < 10^{-15}$). Finally, the inhomogeneity remains if the value function is nonlinear. **Figures 7C,D** illustrate this latter conclusion for the Component-saturation value function, using the delta and Phi rules respectively. The respective variations are approximately 25% and 15%. And the one-way ANOVA tests yield $p < 10^{-15}$ for both cases.

In closing, we cannot strictly speaking reject Hypothesis VI, because the free parameters do not converge statistically to a unique fixed point. However, the breakdown of uniqueness is less than expected from **Figure 1**. For example, the variation of final balance free parameters is small, being less than 2.5% for the linear value function. The small variation and non-uniqueness of convergence, leads us to define the concept of region (instead of point) of convergence.

Failing Hypotheses: How Stochasticity Helps and Shapes Learning

The sections “Limitations of the Purely-linear Conditions” to “Non-uniqueness of the Learned Free Parameters” sections ruled out the hypotheses raised in the “Hypotheses Tested in This Article” sections, except possibly for Hypothesis VI, whose test nevertheless yielded a surprising result. Why did those hypotheses fail? In the “Hypotheses Tested in This Article” sections, we mentioned that we formulated the hypotheses by disregarding the stochastic nature of the learning process. In this section, we show that the stochasticity of the process has more effect on the learning outcome than expected.

To understand why stochasticity led to the rejection of all but one of the hypotheses raised by **Figure 1**, we dove deeper into the surviving hypothesis. Although the final free parameters did not predictably exhibit uniqueness according to Hypothesis VI, their variation was much less than expected (**Figure 7**). Why was the variation so small? To answer this question, consider initially the linear value function (**Figure 1A**). The expectation of large variation of final free parameters was due to every point on the target isoline giving the same prediction of reward. However, because we drew the sensory inputs and rewards randomly across iterations, the slopes and intercepts of the isolines changed. Consequently, the target isoline changed across iterations. But the intersections of the target isolines crossed in a small region around the fixed parameters of the reward model (**Figure 8A**). Therefore, the variations of the final free parameters were smaller than we would expect by only considering the non-stochastic process (**Figure 1A**). The same low-variation result applied to the nonlinear value functions (data not shown). The stochasticity of the learning process thus helped improve the acquired final free parameters.

Similarly, the stochasticity helped explain the failure of the other hypotheses. Hypothesis I failed because of the non-zero final regrets observed despite using a linear value function when the reward model was nonlinear (**Figures 2F, 3A**). Consider for example the nonlinear Output-saturation model in **Figure 1B**. In this model, the contour plot also consisted of parallel straight isolines. When the learning converged around the right solution, the stochastic process sometimes took the free parameters beyond the target isoline and sometimes before it. As shown in **Figure 1B**, the gradient was larger before than beyond that isoline. The larger gradient caused the recovery to be faster for the former kind of error. Thus, the value overestimated reward on average, that is, the free parameters spent more time recovering beyond the target isoline than before it. The consequence was that when the regret is positive, it stayed so for fewer iterations than when it was negative (**Figure 8B**). The regret was thus negative on

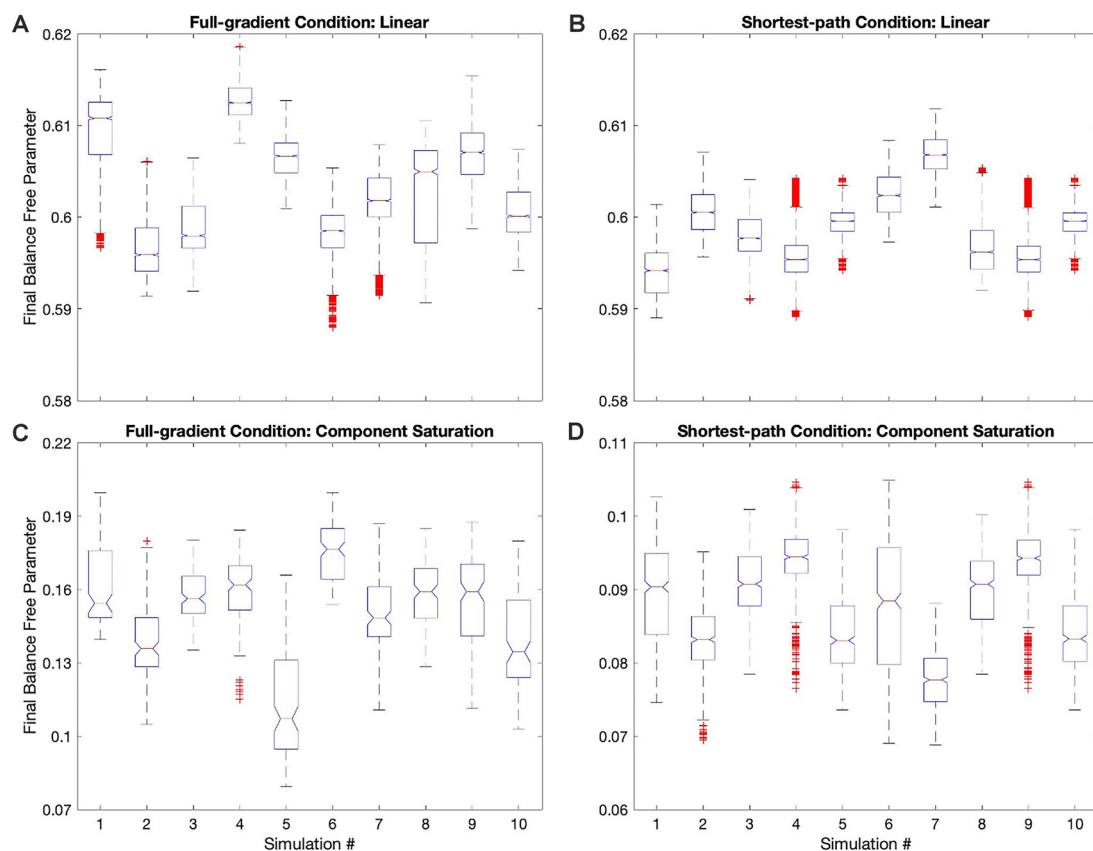


FIGURE 7 | Box plots of the final balance free parameter in each of the simulations of four conditions in **Table 1**. **(A)** Full-gradient delta rule with linear value function. **(B)** Shortest-path (Phi) rule with linear value function. **(C)** Full-gradient delta rule with Component-saturation value function. **(D)** Shortest-path (Phi) rule with Component-saturation value function. Each box plot contains the 10 simulations of the indicated condition. On each box, the central mark is the median, and the edges of the box are the 25th and 75th percentiles. The whiskers are extended to the most extreme data points that are not considered outliers, with those being plotted individually using red “+” symbols. Box plots include notches for the comparison of the median values. Two medians are significantly different at the 5% significance level if their intervals, represented by notches, do not overlap. In all these four examples, the median final balance free parameters varied significantly across the simulations.

average (**Figures 2F, 3A**). Similar regret reasons helped explain why Hypotheses II, III, and V failed (details not discussed here for the sake of brevity).

Stochasticity also explained why we could reject Hypothesis IV. We ruled it out because the linear value function yielded curved trajectories for all but the Output-saturation reward model (**Figure 2E**). An initial hypothesis for what caused these curved trajectories was the motivation function Equation (1). Aleem et al. (2020) showed that making this function a constant eliminated the curved trajectory in their model. However, their model corresponded only to Condition four in **Table 1**, so we could not be sure that their result would apply to all the conditions in **Figure 2E**. When we probed this possibility by setting the motivation to a constant, we generally did not eliminate the curvatures of the trajectories in that figure. The only exception was for the Aleem et al.’s reward model (**Figure 8C**).

Further investigation revealed that the reason for the curvatures was due to something more fundamental and again, related to the stochasticity of the learning process. The argument

explaining the reason was mathematical. Taking the mean-field approximation of Equation (19) (Chaikin and Lubensky, 2007) and neglecting the probabilistic variations of m (because it does not matter for the curvatures) we get

$$\begin{aligned}\frac{d\vec{w}(t)}{dt} &= \epsilon_{\delta} m \left(\left(r^*(t) \nabla_w \mu(\vec{u}(t); \vec{w}(t)) \right)_{r^*, \vec{u}} \right. \\ &\quad \left. - \left(\mu(\vec{u}(t); \vec{w}(t)) \nabla_w \mu(\vec{u}(t); \vec{w}(t)) \right)_{\vec{u}} \right), \\ \frac{d\vec{w}(t)}{dt} &= \epsilon_{\Phi} m \left(\left(r^*(t) \vec{\Phi}(\vec{w}(t); \vec{u}(t), r^*(t)) \right)_{r^*, \vec{u}} \right. \\ &\quad \left. - \left(\mu(\vec{u}(t); \vec{w}(t)) \vec{\Phi}(\vec{w}(t); \vec{u}(t), r^*(t)) \right)_{\vec{u}} \right).\end{aligned}\quad (26)$$

The term inside the parenthesis in this equation is the subtraction of two averages over r^* and \vec{u} . These averages are like those in Equation (17), using the sampling of r^* and \vec{u} at every t . Consider the situation in which the value $\mu(\vec{u}(t); \vec{w}(t))$ is a poor predictor of the reward $r^*(t)$. If the value underestimates the reward grossly, then the first average dominates the dynamics. If instead the value overestimates the reward grossly, then the

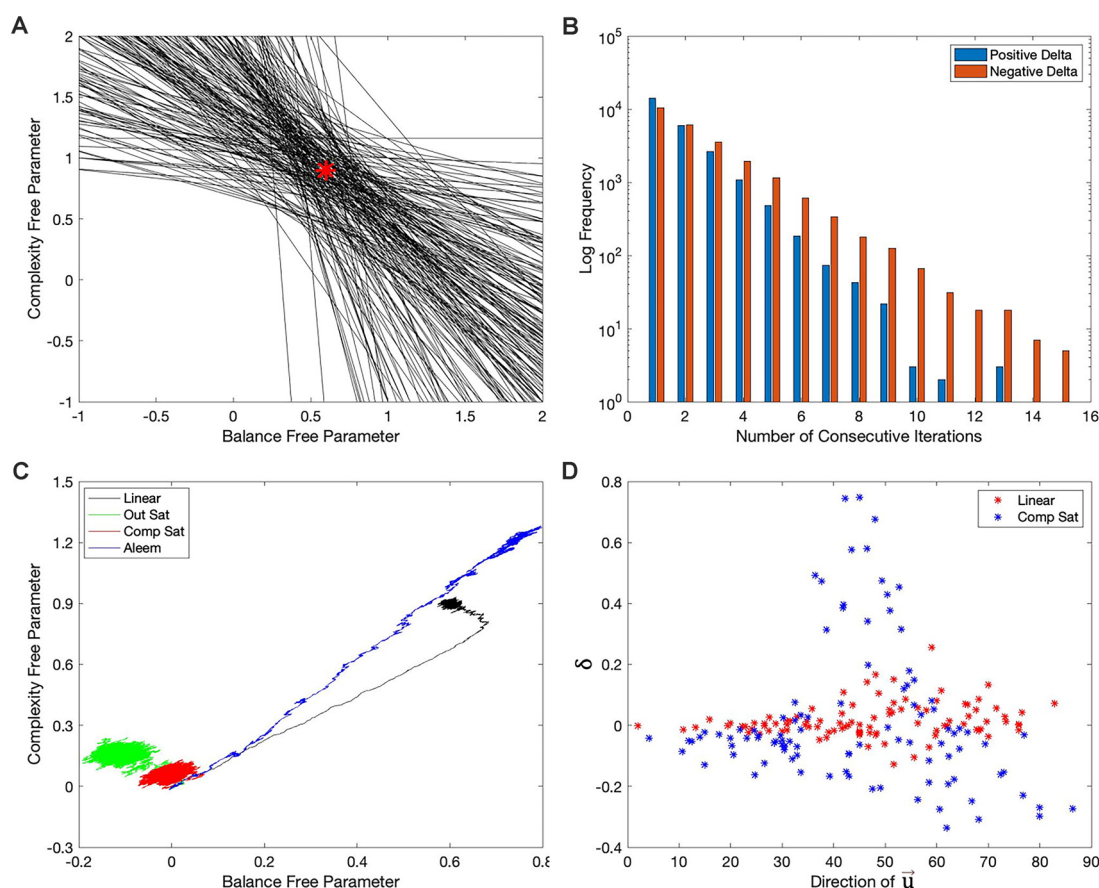


FIGURE 8 | Explanations for the failures of the hypotheses in the “Hypotheses Tested in This Article” sections. **(A)** First 200 target isolines in the simulations of **Figure 7A**. The red star indicates the parameters of the model of reward (**Table 2**). The red star lies in the middle of the small region defined by the intersection of the target isolines. **(B)** Distribution of the number of consecutive iterations spent before (blue) and beyond (red) the target isolines in the last 100,000 iterations of the simulations of **Figure 2B**. The free parameters take longer to recover when they are beyond the target isoline than after it. **(C)** Phase diagram similar to **Figure 2E** but with the motivation function set to 1. The phase diagrams continue to exhibit curvatures, except possibly for that with the Aleem et al.’s reward model. **(D)** Comparison of 100 consecutive iterations (iteration 1,401 to iteration 1,500) with the linear (**Figure 2A**) and Component-saturation (**Figure 2C**) rewards models. The results with the Linear model (red dots) exhibit little correlation between δ and the direction of \vec{u} . But a strong, complex correlation is evident for the Component-saturation model (blue).

second average dominates. Either way, the dominance gives rise to the initial, straight trajectory of the simulations (**Figures 2E, 4E, 5E, 6E**). When the simulations approach the region of convergence, both averages begin to contribute simultaneously to the slower, more random trajectory. Now, the first but not the second average depend on the statistics of r^* . Hence, the initial and final trajectories are generally in different directions, giving rise to the curvatures.

Finally, stochasticity was also at the core of why Hypothesis VII failed. The inversion of complexity and balance free parameters in **Figures 2C, 4C, 5C, 6C** ruled out Hypothesis VII. For this inversion to occur, the right-hand side of Equation (4) had to push the balance free parameters upward faster than the complexity ones. The function δ in Equation (4) was identical for the balance and complexity components of the vector \vec{w} . Similarly, $\nabla_w \mu(\vec{u}(t); \vec{w}(t))$ did not depend on reward and thus, could not differentiate the importance of balance and

complexity. Consequently, because $\nabla_w \mu(\vec{u}(t); \vec{w}(t))$ depended only on \vec{u} , the explanation for why the balance free parameter grew more than the complexity one had to rely on the correlation between δ and \vec{u} . Did certain directions of \vec{u} coincide with larger δ ? **Figure 8D** demonstrated the correlation between δ and \vec{u} with a sector of 100 points in the simulation giving rise to **Figures 2A,C** (This sector was from Iteration 1,401 to Iteration 1,500, but other sectors and other Computer Saturation simulations yielded similar results). The δ in the simulations with the Linear reward model was not strongly correlated with the direction of \vec{u} . However, the Component-saturation model yielded larger positive δ than the Linear model at low angles of \vec{u} (closer to the balance axis). Moreover, for the most part, the Component-saturation model yielded negative δ , specially at the larger angles, that is, closer to complexity. Therefore, **Figure 8D** confirmed the correlation between δ and \vec{u} . This correlation was such that the Component-saturation model yielded statistically

larger balance free parameters than complexity ones. Details of why the Component-saturation model exhibited the correlation seen in **Figure 8D** had to do with the specific shape of the nonlinearity and the statistics of \tilde{u} . We left these details out of this paper for the sake of brevity.

DISCUSSION

An increasingly large number of neuroimaging studies have allowed us to begin understanding the basic neural circuitries underlying the computation of aesthetic biases in the brain (Brown et al., 2011). These circuitries are suggestive of computational mechanisms for the learning of these biases as a set of decision values. Their acquisition would take the form of reinforcement learning gated by internal mechanisms of motivation. Accordingly, a recent theoretical framework for the learning of aesthetic biases followed these computational mechanisms (Aleem et al., 2019, 2020). A model based on that framework could account for interesting features of human aesthetic biases. These features included individuality (Nelson and Morrison, 2005; Brown and Dissanayake, 2009; Silvia et al., 2009), cultural predispositions (Masuda et al., 2008; Park and Huang, 2010; Senzaki et al., 2014), stochastic dynamics of learning and aesthetic biases (Grzywacz and de Juan, 2003; Pouget et al., 2013; Aleem et al., 2020), and the peak-shift effect (Ramachandran and Hirstein, 1999; Costa and Corazza, 2006; Aleem et al., 2020). However, despite the success in explaining these features, a potential major weakness of the model in Aleem et al. (2020) was the linearity of the value function used to predict reward. Such an assumption of linearity is often made in reinforcement-learning models of brain function (Kaelbling et al., 1996; Sutton and Barto, 2018). In this research, we probe what would mean to relax this assumption. In this section, we discuss the effect of relaxing linearity on regret (“Minimization of Regret” section), learning rate (“Efficiency of Learning” and “Phi Versus Delta Rules” sections), and qualitative errors (“Does the Brain Use Ecological Value Functions?” section).

Minimization of Regret

The learning performance exhibited significant regret (error) when using a linear value function to try to predict rewards arising from a nonlinear model. Others have proposed nonlinear value functions (Chung et al., 2018), methods to deal these functions (Xu et al., 2007; Gu et al., 2016; Osband et al., 2016), or their approximators (Tesauro, 2005; Kober et al., 2013; Mahadevan et al., 2013). Here, we attempted to develop optimal nonlinear value functions. The result was exciting because it told us that the optimal nonlinear value function related directly to the statistics of reward in a predictable manner [Equation (8)]. However, incorporating the optimal nonlinear value function helped with some nonlinear reward models but not others. We had expected better performance with these value functions when using the delta rule. Our expectation was due to the mathematical demonstration of the minimization of regret, even with nonlinear value functions. How did we explain this unmet expectation? The expectation of optimization came from a process of gradient descent implemented by the delta rule (Sutton and Barto, 2018).

That the regret did not go to zero could have meant that a local minimum different from the global one trapped the gradient descent (Beck, 2017). Such traps might occur for some nonlinearities but not others.

However, the specific type of stochasticity in our models made it unlikely that their learning processes normally stopped at local minima. The stochastic mechanism arising from the probabilistic sampling of sensory stimuli, rewards, and motivations caused the target isolines to vary. The variation likely helped the free parameters to approach their optimal values (**Figures 7A,C**). This is not surprising because stochasticity often helps optimization processes (Metropolis et al., 1953; Kirkpatrick et al., 1983; Spall, 2003). But for our models, the interaction between stochasticity and the nonlinearities could also cause important errors. Even if the simulation succeeded in reaching exactly a target isoline, the next instant would produce a new one. At this new instant, the vector of free parameters could be before or beyond the new target isoline. The rate of recovery in these two conditions were different because of the model nonlinearity. Consequently, on average, the solution was not optimal, because of the interaction between stochasticity and the nonlinearities of the models. Errors of various forms of stochastic optimization have been described in other studies (Ingber, 1993; Shen et al., 2020). For example, errors could arise if the sampling were not truly stochastic. This could happen to some degree if predictions based on prior learning or motivational factors affected the sampling (Janis and Mann, 1977; Frey, 1986; Schulz-Hardt et al., 2000).

On the other hand, if the learning process occasionally stopped at local minima because of nonlinearities of value functions, it might explain a surprising result from the history of art. An analysis of the statistics of art across the Renaissance and Baroque revealed phase transitions in some measures (Correa-Herran et al., 2020). Another such abrupt transition was observed in a study of the changes in fractal dimension and Shannon entropy in Western paintings (Mather, 2018). The discussion by Correa-Herran et al. (2020) pointed out the essential components of such phase transitions. These components had to be nonlinear interactions between the basic components of a system, which was under the influence of changing external conditions. Correa-Herran et al. (2020) proposed that the basic elements were the values associated with different aesthetic variables. Hence, our proposed use of nonlinear value functions is compatible with the ideas of Correa-Herran et al. (2020). Following their proposal, our nonlinear value functions would generate nonlinear mutual influence between artists learning from each other (Aleem et al., 2019; Correa-Herran et al., 2020). In turn, according to Correa-Herran et al. (2020), the changing external conditions were due to the social pressure to innovate (Barnett, 1953). Such a pressure could come from the desire to increase realism during the Renaissance (Janson et al., 1997). More pressure came from the competition among artists to gain the favor of patrons (Chambers, 1970).

Efficiency of Learning

An implication of the delta rule is that it tends to maximize the rate of learning convergence for the linear value function

(Aleem et al., 2020). Under these conditions, the rate of recovery from fluctuation errors after convergence is also maximal. Therefore, these conditions should implement a highly efficient learning process, albeit with some caveats (Zomaya, 2006; Sutton and Barto, 2018). In contrast, for nonlinear value functions, the delta rule is not expected to lead to efficient learning in general (“Hypotheses Tested in This Article” sections). We thus expected the nonlinear value functions to lead to relatively slow convergence and recovery with the Full-gradient conditions. This expectation did not materialize for the Saturation conditions. We also expected the Shortest-path Phi rule to overcome these deficiencies of the gradient-based delta rule. The Phi rule does so by going directly to the optimal point on the target isoline. But again, this expectation for the Phi rule failed for the Saturation conditions.

How can we explain the failures of the expectations for efficiency of the learning rates of convergence and recovery? Here, we will focus on the time of convergence because its strong correlation with the time of recovery makes the answers similar. As discussed after Equation (26), the time of convergence is dominated by two factors: First, we must consider how far the free parameters must travel to reach the slower, more stochastic portion of the learning trajectory. This phase of the trajectory is reached when the two averages in Equation (26) become similar. Second, we must consider the speed of movement of the free parameters during the early, “straight” portion of the trajectory. This speed depends on the largest average of Equation (26). Hence, three factors may influence this speed. They are the gradient of the value function, the distance from the nearest point on the target isoline, and the correlation between reward and the direction of the vector of motion. Because these factors vary across value functions, the factors modulate the different efficiencies of the learning rates.

These two factors explain the various apparent efficiencies of the time of convergence. For example, the short phase-diagram trajectories of the Saturation conditions explain their fast convergence. In contrast, the long trajectories for the Linear and Aleem et al. conditions help explain their slow convergence. However, for these conditions, the speed of movement of the vector of free parameters during the early, “straight” portion of the trajectory also matters. The phase-diagram trajectories for the Linear and Aleem et al. conditions are almost as long. But the latter converges much more slowly than the former. This slow convergence for the Phi rule provides further evidence against Hypothesis V (“Hypotheses Tested in this Article” and “Failing Hypotheses: How Stochasticity Helps and Shapes Learning” sections).

Phi vs. Delta Rules

The importance of the delta rule arguably derives from its simplicity of implementation, low computational cost, and differential-equation form (Widrow and Hoff, 1960; Stone, 1986). However, we argue here that the delta rule may do poorly when applied to some nonlinear value functions. In those situations, the gradient used in the rule has a non-optimal

direction (**Figures 1C,D**). Alternatives could include gradient-free algorithms, but they do not tend to have the simple and differential forms (Kirkpatrick et al., 1983; Kennedy and Eberhart, 2001; Conn et al., 2009; Mockus, 2012). We thus proposed an alternate differential-equation-based rule that overcomes this deficiency. The new rule (called Phi or Shortest Path) does not estimate the direction of descent based on the gradient at the location of the vector of free parameters. Instead, the new rule uses holistic knowledge of the nonlinear value function to set the direction toward the optimal point on the target isoline. This holistic rule leads to better regret performance. Furthermore, because of its differential form, the Phi rule allows for a simple implementation as the delta rule.

However, the Phi rule has an important disadvantage when compared to the delta rule. The holistic implementation of the Phi rule is bound to make it computationally expensive and consequently, slow. In our implementations, the simulations with the Phi rule conditions were about five times slower than those with the delta rule. But we did not attempt to optimize our implementation of the Phi rule. The main step in such an optimization would be to find efficient algorithms to obtain the isolines of the value function. We used a standard implementation of the Marching Squares algorithm (Maple, 2003), but faster versions exist (Ho et al., 2005; Garrido et al., 2006). We also applied the algorithm to a 101×101 pixels approximation of the value function and perhaps a coarser approximation would be enough. In addition, we could have used other algorithms that are faster for isoline calculations (Yanchang and Junde, 2001). Finally, the holistic isoline computation is parallelizable (Selikhov, 1997; Belikov and Semenov, 2000; Huang, 2001; Dong et al., 2011), making it imminently efficient for brain-network computations.

Does the Brain Implement Nonlinear Value Functions?

The brain has been often argued to linearize what would otherwise be nonlinear input dependencies (Yu and Lewis, 1989; Bernander et al., 1994; Ermentrout, 1998). Such linearization would allow the brain to map conceptual or perceptual dimensions using linear functions. For example, Naselaris et al. (2011) performed successful neuroimaging on many conceptual and perceptual dimensions often assuming such linearization. These authors’ results on linearization have been confirmed by other studies (Qiao et al., 2019). Moreover, linear value functions account for some forms of reinforcement learning in the basal ganglia (Schultz et al., 1997; Hollerman and Schultz, 1998; Schultz, 2015; Sutton and Barto, 2018). Hence, such value functions may sometimes provide a simpler, suitable model for neurobiological or psychological value updating than the models considered in this article.

However, two lines of argument suggest that this linearization argument is only an approximation that is not always valid. The first line is that some of the results above have been disputed. For example, the compressive spatial summation in human cortex (Kay et al., 2013) has challenged the unchecked applicability of the Naselaris et al.’s (2011) conclusions. This challenge is

compatible with the neural representation of stimulus features becoming increasingly nonlinear as one moves along the sensory pathway (Holdgraf et al., 2017). Further limitation of assuming linearization is the nonlinear processing at high-in-the-hierarchy levels of the brain (Andrzejak et al., 2001; Faure and Korn, 2001; Freeman and Vitiello, 2006; Afraimovich et al., 2011). Finally, although some linear models for reinforcement learning in the basal ganglia are good enough, this process is decidedly nonlinear (Frank and Claus, 2006; Hsu et al., 2009; Niv et al., 2012).

Even more important is the argument that many empirically determined value functions in the brain are often nonlinear. An example for this argument in the visual domain comes from psychophysical studies of preference for complexity. The visual-complexity value function in humans is highly nonlinear, lying on an inverted “U” curve, with people liking moderate amounts of complexity (Berlyne, 1971; Aitken, 1974; Nicki and Moss, 1975; Saklofske, 1975; Imamoglu, 2000). Another nonlinear value function for the human visual system is indicated by the saturation relationship between preference and the number of the axes of symmetry in an image (Wu and Chen, 2020). Examples of nonlinear value functions in non-visual sensory modalities also exist. In the auditory system, for instance, the preference for a piece of music is a saturating function of the familiarity with the piece (Szpunar et al., 2004). Relatedly, preference for music has a U-shape dependence on harmonic surprise (Miles, 2018). And even when one leaves the pure sensory domain into social value, value functions are nonlinear. For example, the tendency of humans to adjust values to social conformity by reinforcement learning has a nonlinear dependence on mean social value (Klucharev et al., 2009).

The implications of the brain employing nonlinear value functions in many situations is important. As stated above, using linear value functions would often be good enough. The learning process would always converge because even if the brain erroneously assumes a linear value function, the process minimizes a positive functional (Aleem et al., 2020). And the convergence can occasionally be faster for erroneous linear value functions than for correct nonlinear ones. However, the price that the brain would be pay is large systematic regrets with erroneous linear value functions. Some degree of regret is unavoidable in the learning of aesthetic value because of the stochasticity of the process. But our results show that the brain can minimize regret in a statistical sense by choosing the appropriate value function. Therefore, by choosing to implement nonlinear value functions in many situations, the brain seems to be prioritizing the minimization of regret over the ease of computation.

Does the Brain Use Ecological Value Functions?

Because the Phi rule requires holistic knowledge of the value function, one must ask how would the brain know what the value function is. An answer to this question is that the brain has a bank of socially and ecologically important value functions. Another answer is that the brain uses a single, multidimensional value function, capturing social and ecological values. The

brain may develop such value functions through evolutionary pressure. This proposal echoes ecological and evolutionary ideas for sensory function (Field, 1987; Atick and Redlich, 1992; Grzywacz and de Juan, 2003). Alternatively, the brain could build ecological value functions through developmental and learning mechanisms. Again, this would be akin to the developmental models for optimal receptive fields in the sensory systems of the brain (MacKay and Miller, 1990; Miller, 1994; Burgi and Grzywacz, 1998). And this would be akin to learning new brain representations for familiar objects in adult life (Tarr, 1995; Weinberger, 1995; Booth and Rolls, 1998). Thus, if variables like balance, complexity, and symmetry have evolutionary importance, then the brain would develop dedicated circuitry, facilitating their computation and assignment of value. Such a dedicated circuitry would make sense because the optimal value function depends directly on the external statistics of reward [Equation (8)]. This link between the ease of dedicated computation and aesthetic value is the premise of the Processing Fluency theory (Reber et al., 2004; Aleem et al., 2017; Correa-Herran et al., 2020). The work here and elsewhere suggests that humans learn individually to weigh the various parameters of the ecological value functions (Aleem et al., 2019, 2020). This conclusion suggests that studying the statistics of reward may be as important as investigating the statistics of natural stimuli (Field, 1987; Ruderman and Bialek, 1994; Balboa et al., 2001; Balboa and Grzywacz, 2003).

However, the hypothetical use of ecological value functions implies a couple of limitations in the computation of aesthetic biases. One limitation would be the inability to learn new values outside the set provided by ecological pressures. The alternative would be to use general value functions that could capture both the ecological ones and some that may not be ecological. Examples of such general value functions were introduced elsewhere (Konidaris and Osentoski, 2008; Sutton et al., 2011; Schaul et al., 2015). Another limitation of using just ecological value functions is the error that they would make when a sensory stimulus does not fit their expectations. Using the wrong value function increases the magnitude of regret in the learning process. However, even when the value functions are right and optimal, quantitative and qualitative errors do occur. Errors like these and others are observed after reinforcement learning in the brain (O'Reilly and McClelland, 1994; Clouse, 1997; Niv, 2009; Gold et al., 2012; Dabney et al., 2020). Therefore, these kinds of errors may be unavoidable. The best that one can hope is to make important errors as small as possible. The important errors are not those of free parameters but of value, that is, of the estimation of reward. Value functions and update rules optimized for social and ecological constraints may thus be ideal for the learning of aesthetic biases.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

NG developed the theoretical framework and its equations, carried out the mathematical analyses, computer simulations, statistical analysis of the results, and wrote the manuscript.

FUNDING

The research was supported by funds provided by the President and Board of Trustees of Loyola University Chicago.

REFERENCES

- Afraimovich, V., Young, T., Muezzinoglu, M. K., and Rabinovich, M. I. (2011). Nonlinear dynamics of emotion-cognition interaction: when emotion does not destroy cognition? *Bull. Math. Biol.* 73, 266–284. doi: 10.1007/s11538-010-9572-x
- Aitken, P. (1974). Judgments of pleasingness and interestingness as functions of visual complexity. *J. Exp. Psychol.* 103, 240–244. doi: 10.1037/h0036787
- Aleem, H., Correa-Herran, I., and Grzywacz, N. M. (2017). Inferring master painters' esthetic biases from the statistics of portraits. *Front. Hum. Neurosci.* 11:94. doi: 10.3389/2017.00094
- Aleem, H., Correa-Herran, I., and Grzywacz, N. M. (2020). A theoretical framework for how we learn aesthetic values. *Front. Hum. Neurosci.* 14:345. doi: 10.3389/2020.00345
- Aleem, H., Pombo, M., Correa-Herran, I., and Grzywacz, N. M. (2019). "Is beauty in the eye of the beholder or an objective truth? A neuroscientific answer," in *Mobile Brain-Body Imaging and the Neuroscience of Art, Innovation and Creativity*, eds J. Contreras-Vidal, D. Robleto, J. G. Cruz-Garza, J. M. Azorin and C. S. Nam (Cham: Springer International Publishing), 101–110.
- Andrzejak, R. G., Lehnertz, K., Mormann, F., Rieke, C., David, P., and Elger, C. E. (2001). Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: dependence on recording region and brain state. *Phys. Rev. E Stat. Nonlin. Soft. Matter Phys.* 64:061907. doi: 10.1103/PhysRevE.64.061907
- Atick, J. J., and Redlich, A. N. (1992). What does the retina know about natural scenes? *Neural Comput.* 4, 196–210. doi: 10.1162/neco.1992.4.2.196
- Balboa, R. M., and Grzywacz, N. M. (2003). Power spectra and distribution of contrasts of natural images from different habitats. *Vis. Res.* 43, 2527–2537. doi: 10.1016/s0042-6989(03)00471-1
- Balboa, R. M., Tyler, C. W., and Grzywacz, N. M. (2001). Occlusions contribute to scaling in natural images. *Vis. Res.* 41, 955–964. doi: 10.1016/s0042-6989(00)00302-3
- Barnett, H. G. (1953). *Innovation: The Basis of Cultural Change*. New York, NY: McGraw-Hill.
- Beck, A. (2017). *First-Order Methods in Optimization*. Philadelphia, PA: Society for Industrial and Applied Mathematics.
- Belikov, V. V., and Semenov, A. Y. (2000). Non-sibsonian interpolation on arbitrary system of points in euclidean space and adaptive isolines generation. *Appl. Num. Math.* 32, 371–387. doi: 10.1016/s0168-9274(99)00058-6
- Berlyne, D. E. (1971). *Aesthetics and Psychobiology*. Cambridge, MA: Harvard University Press.
- Bernander, O., Koch, C., and Douglas, R. J. (1994). Amplification and linearization of distal synaptic input to cortical pyramidal cells. *J. Neurophysiol.* 72, 2743–2753. doi: 10.1152/jn.1994.72.6.2743
- Bertsekas, D. P. (1982). *Constrained Optimization and Lagrange Multiplier Methods*. New York, NY: Academic Press.
- Biederman, I., and Vessel, E. A. (2006). Perceptual pleasure and the brain: a novel theory explains why the brain craves information and seeks it through the senses. *Am. Sci.* 94, 247–253. doi: 10.1511/2006.59.247

ACKNOWLEDGMENTS

We thank Hassan Aleem, Ivan Correa-Herran, Maria Pombo, and Jiaan Mansuri for stimulating discussions with us on theoretical and experimental principles of aesthetic biases. Hassan Aleem also provided critical comments on a draft of this article. We also thank Daniel Grzywacz for suggesting the name Phi Rule. And we thank C. Fernando for inspiring this research by suggesting that aesthetic biases have a highly nonlinear dependence on the sensory inputs. Finally, we thank Jo Ann Rooney for creating an exceptional work environment and Gina Lopez for providing administrative support.

- Bonett, D. G., and Wright, T. A. (2000). Sample size requirements for estimating pearson, kendall and spearman correlations. *Psychometrika* 65, 23–28. doi: 10.1007/bf02294183
- Booth, M. C., and Rolls, E. T. (1998). View-invariant representations of familiar objects by neurons in the inferior temporal visual cortex. *Cereb. Cortex* 8, 510–523. doi: 10.1093/cercor/8.6.510
- Brown, S., and Dissanayake, E. (2009). "The arts are more than aesthetics: neuroaesthetics as narrow aesthetics," in *Foundations and Frontiers in Aesthetics. Neuroaesthetics*, eds M. Skov and O. Vartanian (Amityville, NY: Baywood Publishing Co.), 43–57.
- Brown, S., Gao, X., Tisdelle, L., Eickhoff, S. B., and Liotti, M. (2011). Naturalizing aesthetics: brain areas for aesthetic appraisal across sensory modalities. *NeuroImage* 58, 250–258. doi: 10.1016/j.neuroimage.2011.06.012
- Burgi, P. Y., and Grzywacz, N. M. (1998). A biophysical model for the developmental time course of retinal orientation selectivity. *Vis. Res.* 38, 2787–2800. doi: 10.1016/s0042-6989(97)00323-4
- Chaikin, P. M., Lubensky, T. C. (2007). *Principles of Condensed Matter Physics (4th print edition)*. Cambridge, CA: Cambridge University Press.
- Chambers, D. (1970). *Patrons and Artists in the Italian Renaissance*. London, UK: McMillan.
- Chatterjee, A., and Vartanian, O. (2014). Neuroaesthetics. *Trends Cogn. Sci.* 18, 370–375. doi: 10.1016/j.tics.2014.03.003
- Chung, W., Nath, S., Joseph, A., and White, M. (2018). "Two-timescale networks for nonlinear value function approximation," in *Paper Presented at the International Conference on Learning Representations*, New Orleans, LA, 1–32.
- Clouse, J. A. (1997). *The Role of Training in Reinforcement Learning (Vol. 121)*. Amsterdam, Netherlands: North Holland.
- Conn, A. R., Scheinberg, K., and Vicente, L. N. (2009). *Introduction to Derivative-Free Optimization*. Philadelphia, PA: SIAM.
- Correa-Herran, I., Aleem, H., and Grzywacz, N. M. (2020). Evolution of neuroaesthetic variables in portraits paintings throughout the renaissance. *Entropy* 22:146. doi: 10.3390/e22020146
- Costa, M., and Corazza, L. (2006). Aesthetic phenomena as supernormal stimuli: the case of eye, lip, and lower-face size and roundness in artistic portraits. *Perception* 35, 229–246. doi: 10.1068/p3449
- Dabney, W., Kurth-Nelson, Z., Uchida, N., Starkweather, C. K., Hassabis, D., Munos, R., et al. (2020). A distributional code for value in dopamine-based reinforcement learning. *Nature* 577, 671–675. doi: 10.1038/s41586-019-1924-6
- Dann, C., Neumann, G., and Peters, J. (2014). Policy evaluation with temporal differences: a survey and comparison. *J. Mach. Learn. Res.* 15, 809–883.
- Dong, L., Lu, D., and Li, M. (2011). "Parallel algorithm of visualization of reservoir numerical simulation based on pebi grids," in *Paper Presented at the 2011 Fourth International Symposium on Parallel Architectures, Algorithms and Programming*, Tianjin, China, 302–305.
- Ermentrout, B. (1998). Linearization of f-i curves by adaptation. *Neural Comput.* 10, 1721–1729. doi: 10.1162/089976698300017106
- Faure, P., and Korn, H. (2001). Is there chaos in the brain? I. Concepts of nonlinear dynamics and methods of investigation. *C. R. Acad. Sci. III* 324, 773–793. doi: 10.1016/s0764-4469(01)01377-4

- Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A* 4, 2379–2394. doi: 10.1364/josaa.4.002379
- Filiz-Ozbay, E., and Ozbay, E. Y. (2007). Auctions with anticipated regret: theory and experiment. *Am. Econ. Rev.* 97, 1407–1418. doi: 10.1257/aer.97.4.1407
- Frank, M. J., and Claus, E. D. (2006). Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychol. Rev.* 113, 300–326. doi: 10.1037/0033-295X.113.2.300
- Freeman, W. J., and Vitiello, G. (2006). Nonlinear brain dynamics as macroscopic manifestation of underlying many-body field dynamics. *Phys. Life Rev.* 3, 93–118. doi: 10.1016/j.plrev.2006.02.001
- Frey, D. (1986). “Recent research on selective exposure to information,” in *Advances in Experimental Social Psychology*, ed L. Berkowitz (New York, NY: Academic Press), 41–80.
- Garrido, S., Moreno, L., Abderrahim, M., and Martin, F. (2006). “Path planning for mobile robot navigation using voronoi diagram and fast marching”, in *Paper Presented at the Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference* (Beijing, China: IEEE). doi: 10.1109/IROS.2006.282649
- Gilbert, D. T., Morewedge, C. K., Risen, J. L., and Wilson, T. D. (2004). Looking forward to looking backward: the misprediction of regret. *Psychol. Sci.* 15, 346–350. doi: 10.1111/j.0956-7976.2004.00681.x
- Gold, J. M., Waltz, J. A., Matveeva, T. M., Kasanova, Z., Strauss, G. P., Herbener, E. S., et al. (2012). Negative symptoms and the failure to represent the expected reward value of actions: behavioral and computational modeling evidence. *Arch. Gen. Psychiatry* 69, 129–138. doi: 10.1001/archgenpsychiatry.2011.1269
- Grzywacz, N. M., and de Juan, J. (2003). Sensory adaptation as kalman filtering: theory and illustration with contrast adaptation. *Network* 14, 465–482. doi: 10.1088/0954-898x_14_3_305
- Gu, S., Lillicrap, T., Sutskever, I., and Levine, S. (2016). “Continuous deep q-learning with model-based acceleration,” in *Paper Presented at the International Conference on Machine Learning*, New York City, NY, USA, 2829–2838.
- Ho, C. C., Wu, F. C., Chen, B. Y., Chuang, Y. Y., and Ouhyoung, M. (2005). *Cubical Marching Squares: Adaptive Feature Preserving Surface Extraction From Volume Data* (Vol. 24). Oxford, UK and Boston, USA: Blackwell Publishing, Inc.
- Holdgraf, C. R., Rieger, J. W., Micheli, C., Martin, S., Knight, R. T., and Theunissen, F. E. (2017). Encoding and decoding models in cognitive electrophysiology. *Front. Syst. Neurosci.* 11:61. doi: 10.3389/fnsys.2017.00061
- Hollerman, J. R., and Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci.* 1, 304–309. doi: 10.1038/1124
- Hsu, M., Krajbich, I., Zhao, C., and Camerer, C. F. (2009). Neural response to reward anticipation under risk is nonlinear in probabilities. *J. Neurosci.* 29, 2231–2237. doi: 10.1523/JNEUROSCI.5296-08.2009
- Huang, G.-M. (2001). Isoline 3-d display and its parallel algorithm. *Science of Surveying and Mapping*, 26, 20–22.
- Hudspeth, A. J., Jessell, T. M., Kandel, E. R., Schwartz, J. H., and Siegelbaum, S. A. Eds. (2013). *Principles of Neural Science, 5th Edn.* McGraw-Hill, Health Professions Division.
- Iigaya, K., Yi, S., Wahle, I. A., Tanwisuth, K., and O’Doherty, J. P. (2020). Aesthetic preference for art emerges from a weighted integration over hierarchically structured visual features in the brain. *bioRxiv* [Preprint]. doi: 10.1101/2020.02.09.940353
- Imamoglu, Ç. (2000). Complexity, liking and familiarity: architecture and nonarchitecture turkish students’assessments of traditional and modern house facades. *J. Environ. Psychol.* 20, 5–16. doi: 10.1006/jevp.1999.0155
- Inger, L. (1993). Simulated annealing: practice versus theory. *Math. Comput. Model.* 18, 29–57. doi: 10.1016/0895-7177(93)90204-c
- Janis, I. L., and Mann, L. (1977). *Decision Making: A Psychological Analysis of Conflict, Choice, and Commitment*. New York, NY: Free Press.
- Janson, H. W., Janson, A. F., and Marmor, M. (1997). *History of Art*. London: Thames and Hudson.
- Kaelbling, L. P., Littman, M. L., and Moore, A. W. (1996). Reinforcement learning: a survey. *J. Artif. Intell. Res.* 4, 237–285. doi: 10.1613/jair.301
- Kay, K. N., Winawer, J., Mezer, A., and Wandell, B. A. (2013). Compressive spatial summation in human visual cortex. *J. Neurophysiol.* 110, 481–494. doi: 10.1152/jn.00105.2013
- Kennedy, J., and Eberhart, R. C. (2001). *Swarm Intelligence*. San Francisco, CA: Morgan Kaufmann
- Kirkpatrick, S., Gelatt, C. D. Jr., and Vecchi, M. P. (1983). Optimization by simulated annealing. *Science* 220, 671–680. doi: 10.1126/science.220.4598.671
- Klucharev, V., Hytönen, K., Rijpkema, M., Smids, A., and Fernandez, G. (2009). Reinforcement learning signal predicts social conformity. *Neuron* 61, 140–151. doi: 10.1016/j.neuron.2008.11.027
- Kober, J., Bagnell, J. A., and Peters, J. (2013). Reinforcement learning in robotics: a survey. *Int. J. Robot. Res.* 32, 1238–1274. doi: 10.1177/0278364913495721
- Konidaris, G., and Osentoski, S. (2008). “Value function approximation in reinforcement learning using the fourier basis,” in *Computer Science Department Faculty Publication Series, 101* (MA: University of Massachusetts Amherst), 1–11.
- Kreps, D. M. (1990). *A Course in Microeconomic Theory*. Princeton, NJ: Princeton University Press.
- Lacey, S., Hagtvædt, H., Patrick, V. M., Anderson, A., Stilla, R., Deshpande, G., et al. (2011). Art for reward’s sake: visual art recruits the ventral striatum. *NeuroImage* 55, 420–433. doi: 10.1016/j.neuroimage.2010.11.027
- Leder, H., and Nadal, M. (2014). Ten years of a model of aesthetic appreciation and aesthetic judgments: the aesthetic episode-developments and challenges in empirical aesthetics. *Br. J. Psychol.* 105, 443–464. doi: 10.1111/bjop.12084
- MacKay, D. J. C., and Miller, K. D. (1990). Analysis of linsker’s application of hebbian rules to linear networks. *Netw. Comp. Neural Syst.* 1, 257–297. doi: 10.1088/0954-898x_1_3_001
- Maei, H. R. (2011). *Gradient Temporal-Difference Learning Algorithms*. Ph.D. Thesis. Edmonton, AB: University of Alberta.
- Mahadevan, S., Giguere, S., and Jacek, N. (2013). “Basis adaptation for sparse nonlinear reinforcement learning,” in *Paper Presented at the Association for the Advancement of Artificial Intelligence*, Bellevue, Washington, USA, 654–660.
- Mahmood, A. R., and Sutton, R. S. (2015). “Off-policy learning based on weighted importance sampling with linear computational complexity,” in *Paper Presented at the 31st Conference on Uncertainty in Artificial Intelligence*, Amsterdam, Netherlands, 552–561.
- Maple, C. (2003). “Geometric design and space planning using the marching squares and marching cube algorithms,” in *Proceedings of the International Conference on Geometric Modeling and Graphics*, London, United Kingdom, 90–95.
- Martindale, C. (1984). The pleasures of thought: a theory of cognitive hedonics. *J. Mind Behav.* 5, 49–80.
- Masuda, T., Gonzalez, R., Kwan, L., and Nisbett, R. E. (2008). Culture and aesthetic preference: comparing the attention to context of east asians and americans. *Pers. Soc. Psychol. Bull.* 34, 1260–1275. doi: 10.1177/0146167208320555
- Mather, G. (2018). Visual image statistics in the history of western art. *Art Percept.* 6, 97–115. doi: 10.1163/22134913-20181092
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *J. Chem. Phys.* 21, 1087–1092. doi: 10.1063/1.1699114
- Miles, S. A. (2018). *The Relationship Between the Perception of Unexpected Harmonic Events and Preference in Music*. Ph.D. Dissertation. Washington, DC: Georgetown University.
- Miller, K. D. (1994). A model for the development of simple cell receptive fields and the ordered arrangement of orientation columns through activity-dependent competition between on- and off-center inputs. *J. Neurosci.* 14, 409–441. doi: 10.1523/JNEUROSCI.14-01-00409.1994
- Mockus, J. (2012). *Bayesian Approach to Global Optimization: Theory and Applications*. Dordrecht, Netherlands: Kluwer Academic.
- Nadal, M., and Chatterjee, A. (2019). Neuroaesthetics and art’s diversity and universality. *Wiley Interdiscip. Rev. Cogn. Sci.* 10:e1487. doi: 10.1002/wcs.1487

- Naselaris, T., Kay, K. N., Nishimoto, S., and Gallant, J. L. (2011). Encoding and decoding in fmri. *NeuroImage* 56, 400–410. doi: 10.1016/j.neuroimage.2010.07.073
- Nelson, L. D., and Morrison, E. L. (2005). The symptoms of resource scarcity: judgments of food and finances influence preferences for potential partners. *Psychol. Sci.* 16, 167–173. doi: 10.1111/j.0956-7976.2005.00798.x
- Nicki, R., and Moss, V. (1975). Preference for non-representational art as a function of various measures of complexity. *Can. J. Psychol.* 29, 237–249. doi: 10.1037/h0082029
- Niv, Y. (2009). Reinforcement learning in the brain. *J. Math. Psychol.* 53, 139–154. doi: 10.1016/j.jmp.2008.12.005
- Niv, Y., Edlund, J. A., Dayan, P., and O'Doherty, J. P. (2012). Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. *J. Neurosci.* 32, 551–562. doi: 10.1523/JNEUROSCI.5498-10.2012
- O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., and Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron* 38, 329–337. doi: 10.1016/s0896-6273(03)00169-7
- O'Reilly, R. C., and McClelland, J. L. (1994). Hippocampal conjunctive encoding, storage, and recall: avoiding a trade-off. *Hippocampus* 4, 661–682. doi: 10.1002/hipo.450040605
- Osband, I., Blundell, C., Pritzel, A., and Van Roy, B. (2016). “Deep exploration via bootstrapped DQN,” in *Advances in Neural Information Processing Systems 29 (NIPS 2016)*, 4026–4034.
- Park, K. I. (2018). *Fundamentals of Probability and Stochastic Processes with Applications to Communications*. New York, NY: Springer.
- Park, D. C., and Huang, C.-M. (2010). Culture wires the brain: a cognitive neuroscience perspective. *Perspect. Psychol. Sci.* 5, 391–400. doi: 10.1177/1745691610374591
- Pouget, A., Beck, J. M., Ma, W. J., and Latham, P. E. (2013). Probabilistic brains: knowns and unknowns. *Nat. Neurosci.* 16, 1170–1178. doi: 10.1038/nn.3495
- Qiao, K., Chen, J., Wang, L., Zhang, C., Zeng, L., Tong, L., et al. (2019). Category decoding of visual stimuli from human brain activity using a bidirectional recurrent neural network to simulate bidirectional information flows in human visual cortices. *Front. Neurosci.* 13:692. doi: 10.3389/fnins.2019.00692
- Ramachandran, V. S., and Hirstein, W. (1999). The science of art: a neurological theory of aesthetic experience. *J. Conscious. Stud.* 6, 15–51.
- Reber, R., Schwarz, N., and Winkielman, P. (2004). Processing fluency and aesthetic pleasure: is beauty in the perceiver's processing experience? *Pers. Soc. Psychol. Rev.* 8, 364–382. doi: 10.1207/s15327957pspr0804_3
- Riesz, F., and Szökefalvi-Nagy, B. (1990). *Functional Analysis*. New York, NY: Dover Publications.
- Rousseeuw, P. J., and Leroy, A. M. (2003). *Robust Regression and Outlier Detection*. New York, NY: John Wiley & Sons.
- Ruderman, D. L., and Bialek, W. (1994). Statistics of natural images: scaling in the woods. *Phys. Rev. Lett.* 73, 814–817. doi: 10.1103/PhysRevLett.73.814
- Saklofske, D. H. (1975). Visual aesthetic complexity, attractiveness and diverse exploration. *Percept. Mot. Skills* 41, 813–814. doi: 10.2466/pms.1975.41.3.813
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., et al. (2008). *Global Sensitivity Analysis: The Primer*. New York, NY: John Wiley & Sons.
- Schaul, T., Horgan, D., Gregor, K., and Silver, D. (2015). “Universal value function approximators,” in *Paper Presented at the International Conference on Machine Learning*, Lille, France, 1312–1320.
- Schmidhuber, J. (2010). Formal theory of creativity, fun, and intrinsic motivation. *IEEE Trans. Auton. Ment. Dev.* 2, 230–247. doi: 10.1109/tamd.2010.2056368
- Schultz, W. (1998). Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80, 1–27. doi: 10.1152/jn.1998.80.1.1
- Schultz, W. (2015). Neuronal reward and decision signals: from theories to data. *Physiol. Rev.* 95, 853–951. doi: 10.1152/physrev.00023.2014
- Schultz, W. (2016). Dopamine reward prediction error coding. *Dialogues Clin. Neurosci.* 18, 23–32. doi: 10.31887/DCNS.2016.18.1/wschultz
- Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599. doi: 10.1126/science.275.5306.1593
- Schulz-Hardt, S., Frey, D., Lüthgens, C., and Moscovici, S. (2000). Biased information search in group decision making. *J. Pers. Soc. Psychol.* 78, 655–669. doi: 10.1037//0022-3514.78.4.655
- Selikhov, A. (1997). *Cellular Algorithm for Isoline Extraction From a 2d Image (Vol. 6)*. Joint Bulletin of the Novosibirsk Computer Center and the Institute of Informatics Systems. National Curriculum Council.
- Senzaki, S., Masuda, T., and Nand, K. (2014). Holistic versus analytic expressions in artworks: cross-cultural differences and similarities in drawings and collages by canadian and japanese school-age children. *J. Cross Cult. Psychol.* 45, 1297–1316. doi: 10.1177/0022022114537704
- Shen, W., Yang, Z., Ying, Y., and Yuan, X. (2020). Stability and optimization error of stochastic gradient descent for pairwise learning. *Anal. Appl.* 18, 887–927. doi: 10.1142/s0219530519400062
- Silvia, P. J., Henson, R. A., and Templin, J. L. (2009). Are the sources of interest the same for everyone? Using multilevel mixture models to explore individual differences in appraisal structures. *Cogn. Emot.* 23, 1389–1406. doi: 10.1080/02699930902850528
- Skov, M. (2010). “The pleasure of art,” in *Pleasures of the Brain*, eds M. L. Kringelbach and K. C. Berridge (New York, NY: Oxford University Press), 270–283.
- Somasundaram, J., and Diecidue, E. (2016). Regret theory and risk attitudes. *J. Risk Uncertain.* 55, 1–29. doi: 10.1007/s11166-017-9268-9
- Spall, J. C. (2003). *Introduction to Stochastic Search and Optimization: Estimation, Simulation, and Control*. Hoboken, NJ: Wiley.
- Stone, G. O. (1986). “An analysis of the delta rule and the learning of statistical associations,” in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. I*, eds D. E. Rumelhart and J. L. McClelland (Cambridge, MA: The MIT Press), 444–459.
- Strutz, T. (2016). *Data Fitting and Uncertainty: A Practical Introduction to Weighted Least Squares and Beyond*, 2nd Edn. Wiesbaden, Germany: Springer Vieweg.
- Sutton, R. S., and Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. Cambridge, MA: The MIT Press.
- Sutton, R. S., Modayil, J., Delp, M., Degris, T., Pilarski, P. M., White, A., et al. (2011). “Horde: a scalable real-time architecture for learning knowledge from unsupervised sensorimotor interaction,” in *Paper Presented at the International Conference On Autonomous Agents and Multi-Agent Systems*, Taipei, Taiwan, 761–768.
- Szpunar, K. K., Schellenberg, E. G., and Pliner, P. (2004). Liking and memory for musical stimuli as a function of exposure. *J. Exp. Psychol. Learn. Mem. Cogn.* 30, 370–381. doi: 10.1037/0278-7393.30.2.370
- Tarr, M. J. (1995). Rotating objects to recognize them: a case study on the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychon. Bull. Rev.* 2, 55–82. doi: 10.3758/BF03214412
- Tesauro, G. (2005). “Online resource allocation using decompositional reinforcement learning,” in *Paper Presented at the Association for the Advancement of Artificial Intelligence*, Pittsburgh, PA, 886–891.
- Tsitsiklis, J. N., and Van Roy, B. (1997). “Analysis of temporal-difference learning with function approximation,” in *Paper Presented at the Advances in Neural Information Processing Systems (NIPS)*, Vancouver, Canada, 1–32.
- Van de Cruys, S., and Wagemans, J. (2011). Putting reward in art: a tentative prediction error account of visual art. *Iperception* 2, 1035–1062. doi: 10.1068/i0466aap
- Vartanian, O., and Skov, M. (2014). Neural correlates of viewing paintings: evidence from a quantitative meta-analysis of functional magnetic resonance imaging data. *Brain Cogn.* 87, 52–56. doi: 10.1016/j.bandc.2014.03.004
- Vessel, E. A., and Rubin, N. (2010). Beauty and the beholder: highly individual taste for abstract, but not real-world images. *J. Vis.* 10, 18.1–18.14. doi: 10.1167/10.2.18
- Vidal, J., Karplus, W. J., and Kaludjian, G. (1966). “Sensitivity coefficients for the correction of quantization errors in hybrid computer systems,” in *Sensitivity Methods in Control Theory. Proceedings of the International Symposium, Dubrovnik*, (Pergamon Press), 197–208.

- Wang, T., Mo, L., Mo, C., Tan, L. H., Cant, J. S., Zhong, L., et al. (2015). Is moral beauty different from facial beauty? Evidence from an fmri study. *Soc. Cogn. Affect. Neurosci.* 10, 814–823. doi: 10.1093/scan/nsu123
- Weinberger, N. M. (1995). Dynamic regulation of receptive fields and maps in the adult sensory cortex. *Annu. Rev. Neurosci.* 18, 129–158. doi: 10.1146/annurev.ne.18.030195.001021
- White, A., and White, M. (2016). Investigating practical linear temporal difference learning. *arXiv* [Preprint]. Available online at: <https://arxiv.org/abs/1602.08771>.
- Widrow, B., and Hoff, M. E. (1960). *Adaptive Switching Circuits* (No. TR-1553-1). Stanford Electronics Labs.
- Wu, C.-C., and Chen, C.-C. (2020). Symmetry modulates the amplitude spectrum slope effect on visual preference. *Symmetry* 12:1820. doi: 10.3390/sym12111820
- Xu, X., Hu, D., and Lu, X. (2007). Kernel-based least squares policy iteration for reinforcement learning. *IEEE Trans. Neural Netw.* 18, 973–992. doi: 10.1109/TNN.2007.899161
- Yanchang, Z., and Junde, S. (2001). “Gdilc: a grid-based density-isoline clustering algorithm,” in *Paper Presented at the 2001 International Conferences on Info-Tech and Info-Net*, Beijing, China, 140–145.
- Yu, X. L., and Lewis, E. R. (1989). Studies with spike initiators: linearization by noise allows continuous signal modulation in neural networks. *IEEE Trans. Biomed. Eng.* 36, 36–43. doi: 10.1109/10.16447
- Zomaya, A. Y. (2006). *Handbook of Nature-Inspired and Innovative Computing: Integrating Classical Models with Emerging Technologies*. New York, NY: Springer Science & Business Media.

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.

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

APPENDICES

Optimal Value Function

Claim 1

The expected least-squares error of the prediction of fully motivated reward by Equations (1–3), (5) and (6) is minimized by

$$\mu_{opt}(\vec{u}:\vec{w},\vec{k}) = \langle r^* \rangle (\vec{u}:\vec{I}_u = [\vec{w},\vec{k}]), \quad (27)$$

where $\langle r^* \rangle (\vec{u}:\vec{I}_u = [\vec{w},\vec{k}])$ indicates the mean of r^* given the sampled sensory inputs, and the free (\vec{w}) and constant (\vec{k}) parameters of the value function.

Proof

We start from the expected least-squares error in Equation (8), namely,

$$E = \iiint_{\vec{u}, r^*, m} P(\vec{u}, r^*) P(m|\vec{u}) (m(\mu(\vec{u}) - r^*))^2, \quad (28)$$

where we drop both the dependence on t and the parameters for the sake of conciseness. This equation indicates that the error is a functional of $\mu(\vec{u}(t))$. To calculate the optimal function, we fix $\vec{u}(t)$ and calculate

$$\mu_{opt}(\vec{u}) = \underset{r^*, m}{\operatorname{argmin}} \iint P(r^*|\vec{u}) P(m|\vec{u}) (m(\mu^* - r^*))^2. \quad (29)$$

To find this minimum, we differentiate the integrals by μ^* and set the result to 0, yielding

$$\begin{aligned} \iint_{r^*, m} P(r^*|\vec{u}) P(m|\vec{u}) m^2 (\mu_{opt} - r^*) &= 0, \\ \mu_{opt} \langle m^2 \rangle (\vec{u}) - \langle r^* \rangle (\vec{u}) \langle m^2 \rangle (\vec{u}) &= 0, \end{aligned}$$

where $\langle \rangle (\vec{u})$ indicates average given \vec{u} . This last equation proves our claim.

Comments on Claim 1

- The implication of Equation (27) is to tell us the optimal value function in the least-squares sense.
- Consequently, to find out what this function is, one must know the statistics of reward given the sensory stimuli.
- This conclusion suggests that studying the statistics of reward may be as important as investigating the statistics of natural stimuli.

Minimization of Regret Under Optimal Value Functions and the Delta Rule

Claim 2

If for every τ there is a $t > \tau$ such that $m(t) > 0$, then the learning process minimizes

$$E(\vec{w}) = \left\langle m(t)(r^*(t) - \mu(\vec{u}(t):\vec{w}(t)))^2 \right\rangle_t, \quad (30)$$

where $\langle \rangle_t$ stands for time average.

Proof

The gradient of E with respect to the components of \vec{w} obeys

$$\begin{aligned} \nabla_w E(\vec{w}) &\propto -\langle m(t)(r^*(t) - \mu(\vec{u}(t):\vec{w}(t))) \nabla_w \mu(\vec{u}(t):\vec{w}(t)) \rangle_t, \\ \nabla_w E(\vec{w}) &\propto -\langle r(t) - v(t) \nabla_w \mu(\vec{u}(t):\vec{w}(t)) \rangle_t, \end{aligned}$$

or

$$\nabla_w E(\vec{w}) \propto -\langle \delta(t) \nabla_w \mu(\vec{u}(t):\vec{w}(t)) \rangle_t, \quad (31)$$

Hence, the process governed by Equation (4) minimizes $E(\vec{w})$ by performing a gradient descent (Strutz, 2016).

Comments on Claim 2

- The minimization of $E(\vec{w})$ with respect to the components of \vec{w} in Equation (30) implies that $\mu(\vec{u}(t):\vec{w}(t))$ becomes statistically close to $r^*(t)$. Equivalently, $v(t) = m(t) \mu(\vec{u}(t):\vec{w}(t))$ becomes statistically close to $r(t) = m(t) r^*(t)$. Therefore, the process optimizes value by making it as close as possible to reward.
- However, $v(t)$ may not converge exactly to $r(t)$; “Minimization of Regret” section.
- The requirement that for every τ there is a $t > \tau$ such that $m(t) > 0$ is necessary to give the process enough time to reach optimization. If $m(t) = 0$ for every $t > \tau$, then the learning process freezes after τ as shown by Equations (1–4).

Perpendicularity Condition Under the Phi Rule

Claim 3

If one uses the Phi rule [Equations (13–16)] to update reinforcement learning, then Equation (18) holds.

Proof

The Phi rule calls for finding the point in the target isoline (\vec{w}_{opt}) that is closest to the vector of free parameters (\vec{w}). We do this by using the Lagrange-multiplier method (Bertsekas, 1982). We thus build the Lagrangian function

$$\mathcal{L}(\vec{w}_{opt}, \lambda) = (\vec{w}_{opt} - \vec{w})^2 - \lambda (\mu(\vec{u}:\vec{w}) - r^*), \quad (32)$$

where the first term of the right-hand side is the square of the distance between \vec{w}_{opt} and \vec{w} , and the second term is the constraint of the target isoline ($\mu(\vec{u}:\vec{w}) - r^* = 0$) times the multiplier λ . To minimize the distance, we must find the minimum of the Lagrangian function with respect to both \vec{w}_{opt} and λ . We find this minimum by calculating the respective partial derivatives and setting them to zero:

$$\begin{aligned} \nabla_{\vec{w}_{opt}} \mathcal{L}(\vec{w}_{opt}, \lambda) &= 2(\vec{w}_{opt} - \vec{w}) - \lambda \nabla_{\vec{w}_{opt}} \mu(\vec{u}:\vec{w}) = 0, \\ \nabla_{\lambda} \mathcal{L}(\vec{w}_{opt}, \lambda) &= r^* - \mu(\vec{u}:\vec{w}) = 0, \end{aligned}$$

which yields

$$\vec{w}_{opt} - \vec{w} = \frac{\lambda}{2} \nabla_{\vec{w}_{opt}} \mu(\vec{u}:\vec{w}), \quad (33)$$

$$\mu(\vec{u}:\vec{w}) = r^*. \quad (34)$$

These equations prove our claim.

Comments on Claim 3

- The meaning of Equations (33) and (34) is straightforward: Begin with the target isoline [Equation (34)] and find the points in it whose gradients are parallel to the line connecting \vec{w} to \vec{w}_{opt} .
- These gradients are perpendicular to the isoline. Hence, we must calculate the directions perpendicular to the target isoline and find those that are parallel to the line connecting \vec{w} to \vec{w}_{opt} .
- Sometimes, we may have multiple such directions, but this situation is rare.



Imagining How Lines Were Drawn: The Appreciation of Calligraphy and the Facilitative Factor Based on the Viewer's Rating and Heart Rate

Kazuki Matsumoto^{*†} and Takeshi Okada[†]

Department of Educational Psychology, The University of Tokyo, Tokyo, Japan

OPEN ACCESS

Edited by:

Claudio De Stefano,
University of Cassino, Italy

Reviewed by:

Christian Dieter Schunn,
University of Pittsburgh, United States
Jesus Gabriel Cruz-Garza,
Cornell University, United States

*Correspondence:

Kazuki Matsumoto
k-matsumoto@g.ecc.u-tokyo.ac.jp

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

Specialty section:

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

Received: 16 January 2021

Accepted: 07 June 2021

Published: 30 June 2021

Citation:

Matsumoto K and Okada T (2021)
Imagining How Lines Were Drawn:
The Appreciation of Calligraphy
and the Facilitative Factor Based on
the Viewer's Rating and Heart Rate.
Front. Hum. Neurosci. 15:654610.
doi: 10.3389/fnhum.2021.654610

For this study, we examined how recognizing the writing process of calligraphy influences the cognitive and affective processes related to appreciating it, with the aim of contributing to both graphonomics and the psychology of aesthetics. To this end, we conducted two Web-based experiments in which some participants were instructed to view calligraphy by tracing it with their eyes (the tracing method), while others were told to feel free to think and imagine whatever they wanted. Study 1 ($N = 103$) revealed that the tracing method elicits stronger admiration, inspiration, and empathy in viewers. Study 2 ($N = 87$) showed that the tracing method decreases the average heart rate of those who do not frequently engage in calligraphy appreciation as they gaze at calligraphy for a minute-and-a-half (during the second half of the stimulus duration); this suggests that the tracing method could keep viewers from becoming bored while looking at calligraphy. In sum, the tracing method has positive effects on viewing calligraphy. From a broader perspective, the results imply that how in detail viewers recognize the process of creating an artwork will be a key determinant of art appreciation. In addition, our findings demonstrate how we can measure cardiac activities using the emerging technology of the photoplethysmogram (PPG).

Keywords: graphonomics, art viewing, empirical aesthetics, heart rate, smartphone-based PPG, recognition of the process of creation, paralinguistic, human communication

INTRODUCTION

The way people communicate with each other is one of the most important research topics in human science, and is tied to many disciplines including psychology, neuroscience, anthropology, sociology, linguistics, information science, and evolutionary biology (cf. Shannon, 1948; Wiley, 1983; Craig, 1999; Littlejohn and Foss, 2010). Although calligraphy is not the most frequently used form of human communication, as we will see below, studying calligraphy is valuable. It contributes to clarifying written, non-verbal, or artistic communication, areas of communication that remain largely unexplored. In this introduction, we must first see how calligraphy can be positioned within these communication subfields and how it shares certain characteristics with other types of communication.

It is common to distinguish between verbal and non-verbal communication. Needless to say, verbal communication is crucial in social activities, and people have valued it since ancient times;

this is reflected in the fact that many kinds of verbal activities have been refined and made more sophisticated from generation to generation, eventually becoming so-called “art” forms like literature or rhetoric.

How about non-verbal communication? Verbal communication is inseparable from non-verbal communication (cf. Jones and LeBaron, 2002; Hall et al., 2019); if we think about paralinguistic phenomena, we can readily grasp the connection. Paralanguage is normally defined as vocal behavior accompanied by aspects of words (such as pitch and volume), or, more broadly, the aggregation of “vocal, kinesics (gestural), and proxemics (spatial) channels” (Loveday, 1982; cf. Pennycook, 1985; Hall et al., 2019). These definitions connote that any linguistic communication inevitably contains paralinguistic features. Further, paralanguage plays a substantial role in relaying information such as a speaker’s affective state (e.g., Scherer et al., 1973; Johnstone and Scherer, 2000; Scherer, 2003) or intention (e.g., Hellbernd and Sammler, 2016); it is far from a mere peripheral occurrence. Paralinguistic information can be an important part of some artistic activities in the same way that verbal information is. Singing is a good example. When we listen to someone singing, the lyrics convey verbal messages, but we usually pay more attention to how (or in what tone) they are sung, which is the counterpart of paralanguage in everyday face-to-face conversations (for the similarity between everyday vocal expression and musical performance (see Juslin, 2013; Juslin and Laukka, 2003).

Whereas “paralinguistic” research only centers on phenomena in vocal communication or textual simulations, such as emoticon (Luangrath et al., 2017), written linguistic elements are seldom studied as a kind of paralanguage (for exceptions, see Kilyeni, 2009 and Oshiki et al., 2010). However, many examples show that people receive, as well as frequently and actively gather, information from “paralinguistic” components of written communication, in the sense that linguistic content (what is written) always accompanies visible characteristics (how something is written). This influences how readers form impressions. For instance, many organizations in countries such as France and the United States use techniques from graphology for personnel selection (King and Koehler, 2000). Among the aristocracy of pre-modern Japan, the quality of handwriting was seen as fundamental to spousal selection, in addition to literature skills (Gatten, 1986). As for the art of calligraphy, it can be found in virtually any culture with letters. In terms of handwriting, “paralanguage” is too crucial to numerous human behaviors to be left unexamined. If we can psychologically clarify how we receive “paralinguistic” information in written communication, we can make a significant contribution to the entire field of communication research. Based on the above points, we experimentally examined how handwritten objects mediate social interactions from the perspective of the perceiver or viewer, with a focus on Japanese calligraphy. In the remaining part of this introduction, we review the literature on graphonomics and the psychology of aesthetics, argue how our study theoretically

contributes to both fields, and illustrate the purpose of the experiments.

Graphology and Graphonomics

Historically, people in literate societies have been interested in the individuality of handwriting. Yang Xiong, an ancient Chinese philosopher, stated, “The spoken is voice of spirit. The written is picture of spirit” (言心聲也書心畫也). This idea is also familiar in the contemporary West, with many methods of assessing personality through handwriting systematized as graphology. However, although graphology has a long history and is widely popular, most scientific research to date has failed to support its validity (Simner and Goffin, 2003). In the first half of the 20th century, some psychologists perceived graphology as a pseudoscience, similar to phrenology or palmistry (Allport and Vernon, 1933); this critical stance has become more broadly accepted – but not dismissed – in the field of psychology.

Unlike graphology, graphonomics is a more recent and empirical discipline; it refers to the “scientific and technological effort involved in identifying relationships between the planning and generation of handwriting and drawing movements, the resulting spatial traces of writing and drawing instruments (either conventional or electronic), and the dynamic features of these traces” (van Gemmert and Teulings, 2004). As implied in the above definition, existing graphonomic research focuses on written or drawn traces, or the within-individual processes in which they are produced. Notwithstanding, words are written for communication in the first place; thus, we cannot understand entire systems of writing behavior if we ignore what handwriting (including “paralinguistic” features) expresses to readers. Given that a writer might (perhaps unconsciously) modify her/his handwriting so that it looks good to readers (based on her/his own reading experience) during the stage of movement planning or right in the middle of writing, even in a situation where the intrapersonal handwriting process is the object of research interest, it will only be partially revealed, without discussing how readers perceive handwriting. Hence, it is vital for graphonomic research to explore the cognitive processes that underlie perceiving someone’s handwriting.

For this study, we considered an aesthetically valued style of handwriting; that is, calligraphy. There are two advantages to this approach. First, calligraphy is written with the pursuit of an ideal visual appearance of characters (both for the calligrapher and the viewer), and has rich implications for studying people’s handwriting preferences. The process of forming a preference for another person’s characters is likely rooted in the same basis as the process of developing a goal when writing. This signifies that clarifying the mental processes of aesthetic impression formation in viewing calligraphy is meaningful for graphonomics, in the sense that we can investigate the higher-order cognitive processes underlying general writing behavior (such as planning how to make letters look better). Second, in exploring calligraphy as art, we can be informed by theories in the psychology of aesthetics for a deep discussion. Grounded in the theories in question, we expect that calligraphy works convey some information to the viewer other than semantic

content, which is not limited to speculative calligraphers' personalities graphologists have suggested so far, but may include recognition of calligraphers' skills, or of the process of writing calligraphy (Matsumoto and Okada, 2019), as described in the next section.

As we review related studies in the psychology of aesthetics in the following section, we should keep in mind that many of them focus on visual arts, not verbal arts. In the context of the psychology of aesthetics, while diverse investigations deal with visual and literal arts, virtually no research has covered calligraphic works. Since we want to shed light on non-verbal or "paralinguistic" functions in visual features of calligraphy (rather than purely verbal ones), we expect theories of visual art to have useful implications, to which we primarily refer.

Art-Viewing and Viewers' Recognition of the Process of Creating Artworks

When viewing calligraphy, what determines our evaluation, and what kinds of cognitive and affective processes are involved? Findings from the psychology of aesthetics provide a framework for addressing this question. Recently, numerous studies in this area have shown that the mental process by which viewers integrate an artwork's physical features with their own memories or knowledge (whether consciously or not) is essential in establishing their impressions. Leder et al. (2004) information processing model contains five sequential stages in individual art appreciation: (1) perceptual analysis, (2) implicit memory integration, (3) explicit classification, (4) cognitive mastering, and (5) evaluation. This model continues to be updated. Today, a lot of researchers agree that the "evaluation" stage – after "cognitive mastering" in the original model – is not necessarily located at the end of the actual art-viewing process. That is, during relatively long-term viewing, reconsideration and re-evaluation of artwork can be observed; hence, we can expect impressions of art to change as time goes on (Pelowski and Akiba, 2011; Pelowski et al., 2017, 2020).

While the information processing model is concerned with classifying and segmenting components of art appreciation in terms of cognitive psychology, another approach gaining interest centers on viewers' internal representations, generated through a series of processing stages. Following the influential work of Tinio (2013), art creation and viewing are in a symmetrical relationship in which viewers mentally trace artists' process of creation in reverse order. Bulot and Reber's (2013) psycho-historical framework also suggests that art can convey causal-historical information, although without emphasis on the processing order underscored by Tinio's (2013) mirror model. These frameworks are similar in that a communicative aspect of art creation and viewing is elucidated. From the standpoint of communication theory, various empirical studies have shown that how a viewer (i.e., the "receiver" in the communication model) evaluates an artwork (receiving the "message" from the "signal") depends on her/his knowledge of its creator (the "transmitter") or the process of creation, as in the typical communication model, where reception of a message depends on the encoding/decoding rules. For example, Jucker et al. (2014) demonstrated that how

lay people define an object as art (or not) and how they like it is affected by artist-related instructions given prior to viewing, such as regarding whether the artist intentionally created the work.

Whereas many studies have revealed the effect of directly presented information about the creator upon the viewer's evaluation, viewers can adopt a new way of viewing art following a change in their own cognitive structure, even without such a direct presentation of information as the one used by Jucker et al. (2014; Matsumoto and Okada, 2019). By comparing viewers with and without prior experience of creating origami works in a laboratory, Matsumoto and Okada (2019) found that viewers' own creative experiences enabled them to discern and imagine the process of creating artworks (creative origami works) by others in more detail, followed by the promotion of the aesthetic experience in a positive direction, including admiration elicited by upward social comparison. Matsumoto and Okada (2019) discovered that individual differences in cognitive processes of art-viewing – which are especially salient in the comparison between experts and novices (cf. Leder et al., 2004; Leder et al., 2012; Bulot and Reber, 2013) – are reproducible to some degree with novices' acquired experiences of creation. On the other hand, whether the key determinant of the aesthetic experience is the creative experience, or the way in which the process of creating a work is perceived, remains unresolved; Matsumoto and Okada (2019) imply the latter based on post-hoc correlational analysis.

Considering the above, for the current study, we utilized another approach to establish whether how we perceive the process of creating artworks can influence the cognitive process of appreciation, including the overall evaluation. More specifically, we investigated whether the simple cognitive orientation of "how to look at artwork" – without requiring any special equipment or training – can change one's impression of art in diverse ways. Not only did we examine causal relationships that prior research did not fully test; we also indicated the generalizability of the finding in Matsumoto and Okada (2019) due to adopting different types of works (compared to our previous study), which is considered a contribution to the psychology of aesthetics. We had to choose materials suitable for our purpose, and Chinese/Japanese calligraphic works sufficed (mentioned next), as well as graphonomic interest (described previously).

Chinese/Japanese Calligraphy and the "Tracing Method"

Chinese calligraphy dates back to the beginning of the use of Chinese characters (more than 3,000 years ago at least). Since Japanese calligraphy is a branch of Chinese calligraphy and has the same historical origins, they are quite similar. While the most striking difference is in the characters used (*kana* are sometimes used along with Chinese characters in Japanese calligraphy), as for aspects relevant to the current study (such as the process of creation or standards of the value of works), they are very similar, since Chinese traditions (including calligraphy) have been regarded as role models by Japanese intellectuals, for whom writing calligraphy has been possible throughout almost

all of Japanese history (considering this similarity, hereafter, we will not distinguish between them and will refer to both styles as “calligraphy” unless otherwise noted). Unlike modern Western art, traditional calligraphy is not necessarily intended to be exhibited in public. In addition, conventionally speaking, calligraphy is not a fully independent genre, but is inextricably related to diverse forms of written communication. Thus, people today perceive copies of poems or sutras, and personal writings such as letters or diaries, as valuable forms of calligraphy.

In virtue of using calligraphy as experimental material, we were able to effectively investigate what role the viewer's recognition of the process of creating artworks plays in art-viewing, which is an important research topic in the psychology of aesthetics (as mentioned earlier). The reason for this is that the cognitive orientation regarding the recognition of the writing process is attainable by giving a relatively simple instruction on perceptual performance, even when a non-expert contemplates calligraphy. This is because although calligraphy has a static form (like painting), it more directly and vividly presents traces of creation on paper to viewers compared to other forms of visual art, who can refer to the rules of character stroke order and can imagine how (and in which order) lines were drawn if they scrutinized the visual features of lines (such as a blur or gradation). Moreover, our experiments benefitted from the fact that Japanese elementary and junior high schools offer a class on penmanship; almost all Japanese people have had the experience of writing imitatively in a way similar to traditional calligraphy. Due to these circumstances, the Japanese participants likely had a cognitive foundation for recognizing the calligrapher's writing process, without any additional creative experiences. As such, by guiding viewers' perceptual process with the textual instructions, such as “Please pay attention to how lines were drawn,” we were able to experimentally manipulate their recognition of the process of creating artworks in a different way from that of Matsumoto and Okada (2019), thereby providing new evidence for the importance of that kind of recognition in art-viewing.

In addition to stressing that viewers should pay attention to the process of writing while appreciating calligraphy, much of the literature recommends specific methods for viewing. In particular, imitative writing of the same characters in a work of calligraphy (whether with an ink brush or by tracing them with one's finger) is considered effective (e.g., Ishikawa, 2011; Shi, 2020). Below, we will discuss how this technique can facilitate the beginner's detailed recognition of the calligrapher's writing process. Under normal reading conditions, letters or characters are processed quickly enough on average for 4 to 5 letters to be covered by a single eye fixation when reading English sentences (Samuels et al., 2010), where it is neither necessary nor common to pay close attention to the fine features of letters or characters. This is considered true for Japanese to some extent, which transmits a comparable amount of information per fixation versus English and German (Fukuda and Fukuda, 2009). This is also the case in situations where novices are exposed to calligraphy. Even if they are told to be “aware” of the process of writing, they might still construct limited mental representations based on the relatively small amount of

information available from their accustomed habit of reading. In contrast, by spending longer time for imitative writing or tracing characters with one's finger, the viewer can pay close attention to the physical features that convey rich information about how the brush was used – which may also be related to the mental aspect of the calligrapher's writing process (cf. Matsumoto and Okada, 2019) – and thereby think about them at a deeper level (see Okada and Ishibashi, 2017 for a similar discussion in the context of art creation).

Although using a brush or finger is usually recommended for appreciating calligraphy, the previous paragraph implies that we can derive similar benefits by moving only our eyes, as if tracing each stroke, without any other body movement. If employing this method while contemplating calligraphy, viewers are expected to construct a detailed recognition of the writing process based on the work's paralinguistic features. Otherwise – especially without any cognitive orientation for recognition of the writing process through an instruction – novice viewers are likely to compress visual information into a rather simple impression, such as “strong” or “delicate.” For the present study, based on the discussion in this section, we used Japanese calligraphy as experimental material and manipulated an instruction (the independent variable) between participants in terms of how to view calligraphy. We expected this to determine the direct factor in changing their impressions; that is, the cognitive processes involved in recognizing the writing process. In order to influence recognition of the writing process so as to promote their impressions of calligraphy, we instructed the participants to move their eyes as if tracing the characters (hereafter called the “tracing method”) under one condition.

The Current Study

Through experiments using Japanese calligraphy, we explored how recognizing the process of creating artworks may affect the cognitive and affective processes of viewing them. Unlike Matsumoto and Okada (2019), we directly manipulated the viewer's perception through an instruction instead of a creative experience. More specifically, we compared two conditions within this paradigm: one scenario involved a group of participants whom we instructed to use the tracing method; the participants whom we subjected to the second condition were not given any cognitive orientation. Further, to more accurately identify pertinent factors, we added a third condition in the first experiment: We instructed this group of participants to pay attention to the writing process of calligraphy without using the tracing method. By doing so, we were able to establish whether novice participants could construct detailed mental representations of the writing process, so as to update their impressions via a single instruction, without any specific method or procedure (such as the tracing method).

Although we primarily employed the same measurements as Matsumoto and Okada (2019) for the sake of theoretical continuity (namely liking and admiration as aesthetic impressions), we also introduced two new measurements. First, we gauged the degree to which the participants felt inspired by calligraphy using a slightly modified version of a questionnaire developed by Ishiguro and Okada (2015), which was originally

based on Thrash and Elliot's (2003) Inspiration Scale (see **Supplementary Table 1** for details). According to Ishiguro and Okada (2020), inspiration – which mediates between art-viewing and creative behavior – is related to the social comparative processes, and can be encouraged by the viewer's "dual focus"; that is, attention paid to both others (e.g., "His approach was to take numerous photos") and self (e.g., "I only took a few shots to obtain the best photo"). Based on Smith's (2000) discussion of inspiration and other social comparative emotions – as a crucial element of the theoretical foundation of both Ishiguro and Okada (2020) and Matsumoto and Okada (2019) – inspiration is fairly likely to be promoted when the recognition of the process of creating artworks changes enough to elicit admiration. This is because inspiration is akin to admiration in Smith's (2000) classification. Both emotions are "upward" and "assimilative"; the only essential difference is whether they are "dual" focused or "other" focused. Thus, when admiration is elicited, a large part of the basis for inspiration is already satisfied. If the duration of viewing is not too brief, viewers' attention may shift to themselves or others from time to time, and inspiration and admiration may co-occur when contemplating a single work. Therefore, we measured how viewers were inspired by calligraphy and expected that when admiration was encouraged, inspiration would be also fostered.

Second, we used a physiological indicator: heart rate value estimation by photoplethysmogram (PPG). PPG is a non-invasive optical method for gauging the relative changes in blood volume in an area of tissue with blood capillaries on the skin surface (such as a finger or earlobe); it can be used to detect heartbeat or pulse and has been increasingly adopted in recent years (Gil et al., 2010; Kamshilin et al., 2015; Lohani et al., 2019). Cardiovascular activity is a commonly employed physiological parameter in the psychology of aesthetics due to its ease of use (e.g., Libby et al., 1973; Nell, 1988; Tschacher et al., 2012). Many studies have explored the link between autonomic nervous system activity (e.g., heartbeats) and affective feelings (mediated by brain regions such as the anterior cingulate cortex, Critchley et al., 2013; see Kreibitz, 2010 for a review). Among affective states, boredom is tied to art-viewing and is associated with heart rate; that is, people tend to exhibit higher heart rates when bored (London et al., 1972; Merrifield and Danckert, 2014; Raffaelli et al., 2018).

Recent research suggests that PPG can be obtained via commercial smartphones without any specialized equipment (Kurylyak et al., 2012; Garcia-Agundez et al., 2017; Guede-Fernández et al., 2020). Smartphone-based PPG allows us to conduct experiments remotely, which is especially valuable in the recent situation of COVID-19 pandemic. This technique has not yet been perfected and is rarely introduced, particularly in psychology. Thus, our findings offer useful insights for practical application. One thing to note here is that smartphone-based PPG in a remote experiment is likely to increase the cognitive load on participants and to make experiments complicated. To address this issue, we conducted two experiments. We designed the first one to be simple; we only included psychological rating scales as dependent variables. The second experiment contained all measurements.

We posited that the tracing method would have a positive impact on participants' admiration and inspiration based on the discussion so far. Related to that, it was possible for the average heart rate to differ between conditions, influenced by changes in affective states elicited by the instructions. In addition, we examined effects on other psychological variables: "the degree of empathy for the calligrapher" (empathy) and "the degree to which imagination is triggered" without any specific hypothesis.

STUDY 1

Method

Participants

A total of 103 participants took part via their own Web-connected computers instead of in a laboratory. We recruited them through a crowdsourcing service and paid them each 1,000 JPY for completing the tasks, which took less than an hour. Only native adult speakers of Japanese with no visual impairments were allowed to participate.

Stimuli

We chose four works of Japanese calligraphy (**Figure 1**) based on the following criteria: (a) they are regarded as classic, established works; (b) the phrases are brief, and contain characters that are easy to read; and (c) they do not differ too much from modern styles, and are considered readable for beginners in calligraphy. In the end, each image only had Chinese characters, the number of which ranged from 1 to 4 (see **Supplementary Table 2** for detailed information about each work).

Experimental Conditions

We randomly assigned the participants to one of three conditions: the orientation with the tracing method (the "Tracing Group"; $n = 34$); the orientation without the tracing method (the "No-Tracing group"; $n = 34$); and the non-orientation (the "Control Group"; $n = 35$). For the Tracing Group, we asked the participants to view the calligraphy by tracing each line in the order in which it was supposed to have been drawn, and by imagining both the mental and physical processes of writing. As for the meaning of "tracing," we only told them to move their eyes; there was no mention about hands or fingers. We instructed the participants in the No-Tracing Group in the same way as the Tracing Group, except for the part about the "tracing method." We told them to imagine the process of creating artworks while viewing it without describing any specific method. For the Control Group, we did not give the participants any orientation and told them to feel free to think about and imagine whatever they wanted. Also, for all participants, there were explicit statements that any kind of thought or imagination that is not suggested in our instruction is not prohibited at all (see **Supplementary Table 3** for detail).

Procedure

After reading a broad description of the tasks (see **Figure 2** for a schematic representation) in an online document, each participant provided informed consent and began the tasks using

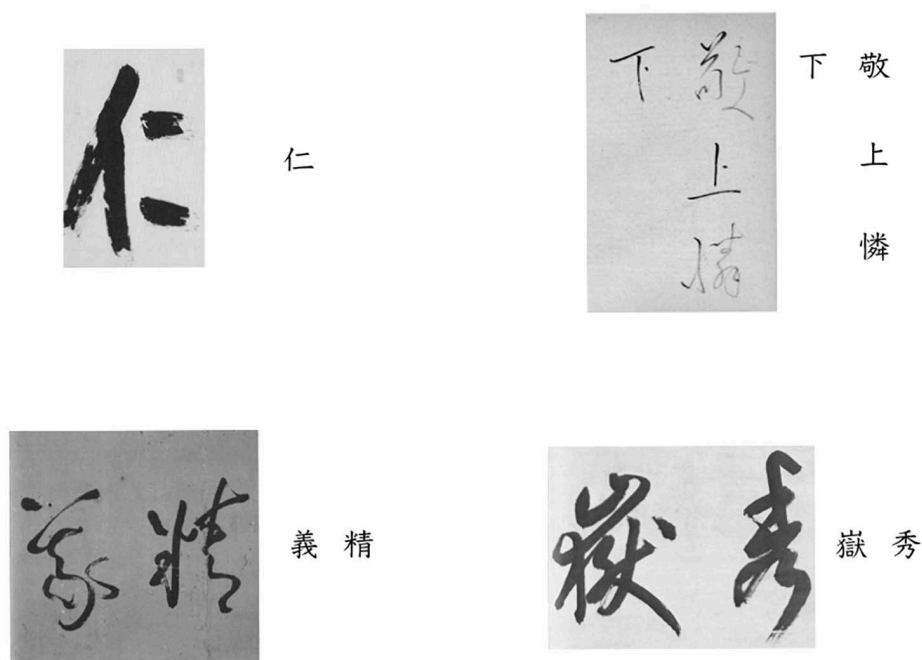


FIGURE 1 | The four calligraphy works used as stimuli (with printed font).

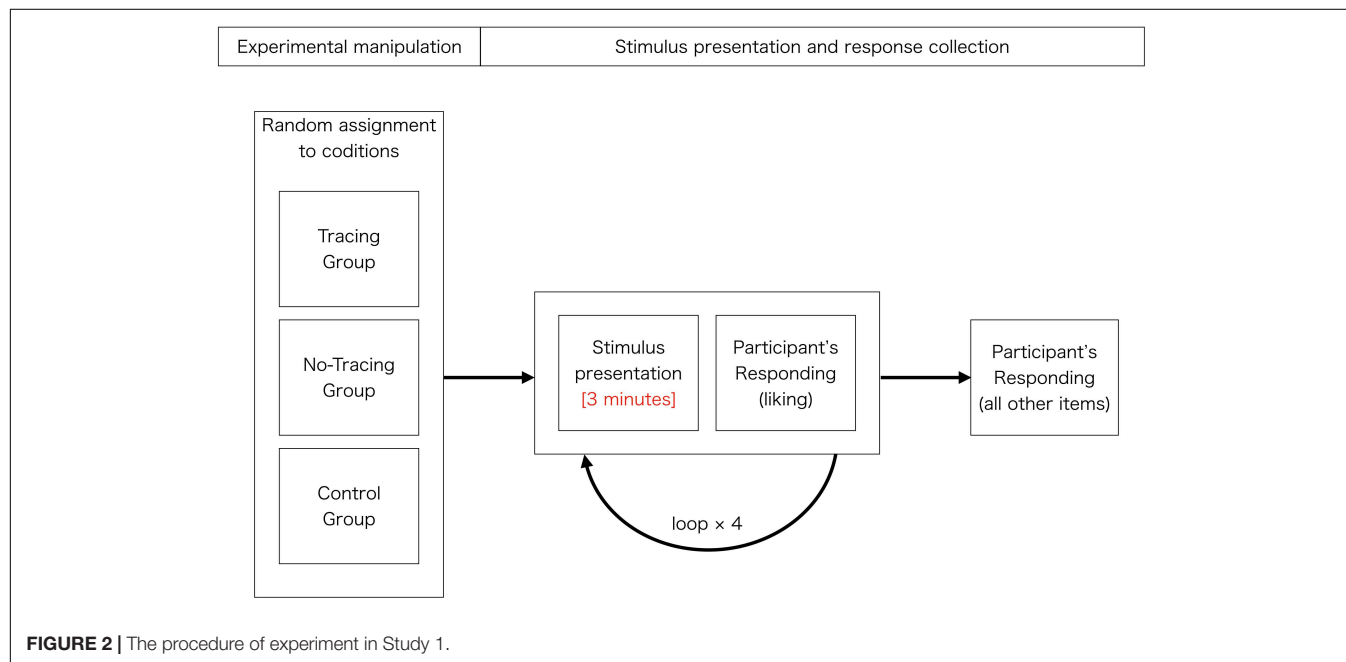


FIGURE 2 | The procedure of experiment in Study 1.

their favored browser, accessing our web server. We required them to complete the tasks using a computer with a stable connection to a network in a quiet, non-distracting environment. In order to let them behave as naturally as possible, we did not strictly control their physical or software-related conditions.

After the participants adjusted their settings, they read a text that suggested how they should view the calligraphy (the manipulation on this was described in the previous section and

Supplementary Table 3). The experimental stimuli were then presented sequentially for 3 min each. With the aim of enabling participants to read older calligraphy stress-free, each stimulus consisted of a set of three images for a single artwork: (1) calligraphy only; (2) calligraphy with the same characters using printed font; and (3) only a short description of the meaning of the written word and information about the author's name, as well as the author's year of birth and death. These images

were displayed one by one; the participants were allowed to switch between the three images at any time by pressing certain keys. Following each stimulus presentation, they were asked to answer two 7-point Likert items measuring the degree to which they liked each presented work, and the degree to which they liked meaning of the written word in each work. After all the stimuli were shown once, they were asked to answer the other Likert items: admiration; empathy; imagination; recognition of the process of creation (awareness of each mental and physical aspect); inspiration; filler items which were not included in any analysis but contributed to avoiding participants being overly aware of what indicators we theoretically focused on (such as “I couldn’t see what it said”); and measurement of their personal characteristics (including the frequency they view calligraphy and length of time spent learning calligraphy). Because of the assumed high correlation between the two personal variables, only the frequency of viewing (with its interaction with the condition) was used as an individual difference factor in every analysis in Study 1. Except for some cases such as inspiration, these items consisted of simple statements such as “I felt admiration” and scales of agreement ranging from 0 (do not agree at all) to 6 (very strongly agree) and they were asked for each stimulus and thus repeated as many times as the number of stimuli, namely four times (see **Figure 2** and **Supplementary Table 1**). There was no time limit for any item. We implemented the main program with jsPsych (de Leeuw, 2015), a JavaScript library for creating behavioral experiments in a Web browser.

Results and Discussion

For each variable measured more than once, we calculated each participant’s score by averaging the stimuli in each session, which, in turn, we used for statistical analysis as a data point. Unless otherwise stated, we compared each dependent variable between groups in analysis of covariance (ANCOVA) models with Tukey’s correction (the “glht” function of the package “multcomp” of R; Hothorn et al., 2008). We included a dummy variable for the condition of instruction, the score of the frequency of viewing calligraphy, and the interaction term of both variables as predictor variables. Following a recommendation from the existing literature on multiple comparison (Hsu, 1996; Wilkinson, 1999), we did not consider rejecting the global null-hypothesis as a prerequisite for pairwise comparisons; nor did we test the global null-hypothesis beforehand in those models, since we were interested in identifying the differences between each of the two groups in those cases.

Validation of Assignments and Experimental Manipulation

To check whether the participants were appropriately assigned in terms of their attitudes toward calligraphy, we calculated the mean and standard deviations (SD) of their scores for the frequency of viewing calligraphy for each condition (see **Table 1**). As one-way analysis of variance (ANOVA) revealed no significant differences among them, $F(2, 100) = 0.03$, $p = 0.966$, there is no evidence to suggest that the assignments were biased regarding the frequency of viewing calligraphy.

Further, to verify whether the manipulation of the instructions influenced the participants’ cognitive processes as expected, we calculated the mean and SD of their scores for (a) the degree to which they were aware of the physical process of writing calligraphy (awareness of physical creation) and (b) the extent to which they were aware of the mental process of writing calligraphy (awareness of mental creation; see **Table 1**). We expected that participants in the Tracing Group and No-Tracing Group would have higher scores than the Control Group for both (a) and (b). As for (a) awareness of physical creation, there was a significant difference between the Tracing Group and the Control Group, $b = 1.48$, $t(97) = 5.20$, $p < 0.001$, and also between the No-Tracing Group and the Control Group, $b = 0.89$, $t(97) = 3.14$, $p = 0.006$. We did not detect any significant differences between the Tracing Group and the No-Tracing Group, $b = 0.59$, $t(97) = 2.07$, $p = 0.101$. Multiple regression analysis indicated no other significant effect for the frequency of viewing and the interactions (**Table 2**). As for (b) awareness of mental creation, like the former results, there was a significant difference between the Tracing Group and the Control Group, $b = 1.00$, $t(97) = 3.15$, $p = 0.006$, and also between the No-Tracing Group and the Control Group, $b = 1.18$, $t(97) = 3.75$, $p < 0.001$. We did not witness any significant differences between the Tracing Group and the No-Tracing Group, $b = -0.18$, $t(97) = -0.57$, $p = 0.835$. In addition, multiple regression analysis only showed a significant effect for the frequency of viewing ($p = 0.031$, **Table 2**) other than the above. These findings were consistent with our prediction. Hence, we considered the participants’ cognitive processes to be appropriately orientated.

Aesthetic Impression

We calculated the mean and SD of participants’ scores for admiration of works of calligraphy (**Table 1** and **Figure 3**). Tukey pairwise comparisons indicated a significant difference only between the Tracing Group and the Control Group, $b = 0.79$, $t(97) = 2.42$, $p = 0.045$, and no significant differences between the other pairs; $b = 0.66$, $t(97) = 2.02$, $p = 0.114$ (tracing–no-tracing), $b = 0.13$, $t(97) = 0.40$, $p = 0.917$ (no-tracing–control). Multiple regression analysis demonstrated no other significant effect for the frequency of viewing and the interactions (**Table 2**). Likewise, we computed the mean and SD of the participants’

TABLE 1 | Descriptive statistics for all rating values in Study 1.

Variable	Tracing group	No-tracing group	Control group
Frequency of viewing	0.41 (0.74)	0.41 (0.78)	0.37 (0.69)
Awareness of physical creation	4.66 (0.87)	4.07 (1.02)	3.33 (1.16)
Awareness of mental creation	4.16 (1.16)	4.42 (0.94)	3.36 (1.36)
Admiration	3.58 (1.05)	3.15 (1.18)	2.94 (1.40)
Liking of a work	3.35 (1.12)	3.12 (1.01)	3.04 (1.16)
Liking of meaning of word	4.15 (0.84)	4.04 (0.78)	4.01 (0.82)
Inspiration	3.51 (1.38)	3.38 (1.24)	2.87 (1.67)
Empathy	3.43 (0.98)	3.30 (1.03)	2.90 (1.23)
Imagination	3.85 (1.23)	4.05 (0.97)	3.42 (1.27)

Standard deviations appear in parentheses.

TABLE 2 | Multiple regression analyses in Study 1.

Dependent variable	Predictor	<i>b</i>	<i>t</i>	<i>p</i>
Awareness of physical creation	Intercept	3.17***	15.99	<0.001
	Condition (with the baseline of control group)			
	Tracing group	1.48***	5.20	(<0.001)
	No-tracing group	0.89**	3.14	(0.006)
	Frequency of viewing	0.42	1.62	0.108
	Interaction (condition × frequency)			
	Tracing condition	−0.39	−1.11	0.271
Awareness of mental creation	Intercept	2.13***	14.11	<0.001
	Condition (with the baseline of control group)			
	Tracing group	1.00**	3.15	(0.006)
	No-tracing group	1.18***	3.75	(<0.001)
	Frequency of viewing	0.63*	2.19	0.031
	Interaction (condition × frequency)			
	Tracing condition	−0.53	−1.35	0.181
Admiration	Intercept	2.75***	12.06	<0.001
	Condition (with the baseline of control group)			
	Tracing group	0.79*	2.42	(0.045)
	No-tracing group	0.13	0.40	(0.917)
	Frequency of viewing	0.50	1.70	0.093
	Interaction (condition × frequency)			
	Tracing condition	−0.40	−0.98	0.328
Liking of a work	Intercept	2.91***	13.82	<0.001
	Condition (with the baseline of control group)			
	Tracing group	0.51	0.30	(0.212)
	No-tracing group	0.06	0.30	(0.980)
	Frequency of viewing	0.37	1.35	0.179
	Interaction (condition × frequency)			
	Tracing condition	−0.55	−1.43	0.155
Inspiration	Intercept	2.53***	9.54	<0.001
	Condition (with the baseline of control group)			
	Tracing group	0.93*	2.46	(0.041)
	No-tracing group	0.55	1.47	(0.309)
	Frequency of viewing	0.92**	2.68	0.009
	Interaction (condition × frequency)			
	Tracing condition	−0.82	−1.75	0.084
Empathy	Intercept	2.72***	13.32	<0.001
	Condition (with the baseline of control group)			
	Tracing group	0.70*	2.40	(0.048)
	No-tracing group	0.38	1.32	(0.388)

(Continued)

TABLE 2 | Continued

Dependent variable	Predictor	<i>b</i>	<i>t</i>	<i>p</i>
Imagination	Frequency of viewing	0.50	1.89	0.062
	Interaction (condition × frequency)			
	Tracing condition	−0.47	−1.30	0.198
	No-tracing condition	−0.00	−0.01	0.995
	Intercept	3.20***	14.40	<0.001
	Condition (with the baseline of control group)			
	Tracing group	0.67	2.11	(0.093)
	No-tracing group	0.74	2.33	(0.056)
	Frequency of viewing	0.59*	2.05	0.043
	Interaction (condition × frequency)			
	Tracing condition	−0.64	−1.62	0.108
	No-tracing condition	−0.32	−0.82	0.415

Parenthesized values were corrected for multiple comparisons.
p* < 0.05; *p* < 0.01; ****p* < 0.001.

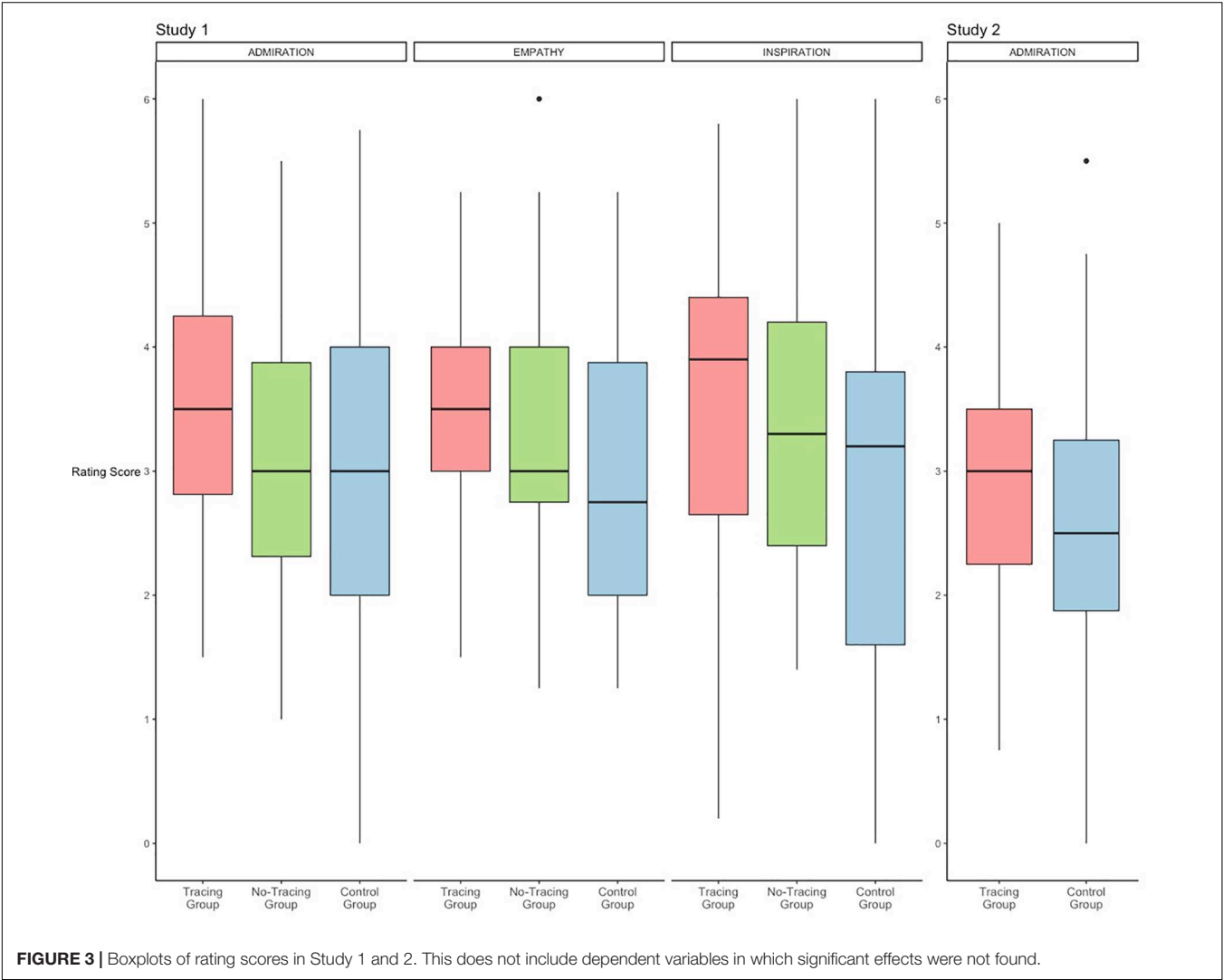


FIGURE 3 | Boxplots of rating scores in Study 1 and 2. This does not include dependent variables in which significant effects were not found.

scores for liking works of calligraphy (**Table 1**). For this analysis, we omitted one participant due to missing data. As opposed to the results of admiration, Tukey pairwise comparisons signaled no significant differences between any two groups; $b = 0.45$, $t(96) = 1.50$, $p = 0.295$ (tracing–no-tracing), $b = 0.51$, $t(96) = 1.70$, $p = 0.212$ (tracing–control), $b = 0.06$, $t(96) = 0.19$, $p = 0.980$ (no-tracing–control). Multiple regression analysis revealed no other significant effect for the frequency of viewing and the interactions (**Table 2**).

The former outcome is in line with our hypothesis based on Matsumoto and Okada (2019), and supports the notion that the degree of admiration we have for artworks depends on how we recognize the process of creating them. On the other hand, the latter result does not support the idea that liking art depends highly on one's recognition of the process of creation. Since Matsumoto and Okada (2019) showed that whether viewers have had creative experiences significantly affects both their admiration and liking, it may seem contradictory at first glance that our results are different for admiration and liking. However, given the differences between the characteristics of admiration and liking, they can be interpreted in a way consistent with our previous study. Admiration is a type of social comparative emotion, and that is seen as the reason why our recognition of the process of creation affects it. According to Matsumoto and Okada (2019), when viewers compare themselves to outstanding others in terms of recognizing others' processes of creation in a detailed way, they should feel intense admiration, as they become confident of the difficulty or "unachievability" of creation by others. In contrast, liking, a typical measurement of aesthetic judgment by viewers, is formed through more varied mental processes (cf. Leder et al., 2004). For example, people may become attached to a painting because it depicts a scene they like, or because it was drawn by a beloved child, rather than due to its outstanding achievement. Therefore, although liking and admiration are likely to correlate, since liking is determined by factors other than how we recognize the writing process, the impact of the instruction seems to be diluted.

For additional analysis, for each condition, we calculated the correlation coefficient between liking of the work itself and liking of meaning of the word in it. The Pearson's r was 0.68, 0.64, and 0.70 for the Tracing Group, the No-Tracing Group, and the Control Group, respectively (all $p < 0.001$). Since the degree of liking the meanings of words are formed quite independently of the recognition of the writing process, our results align with the perspective that other factors determine the overall extent to which one likes a work of calligraphy.

Regarding the effect of instruction in the No-Tracing Group, there were no obvious results, with no significant differences compared to either the Tracing Group or the Control Group. Although our findings do not directly support the existence of a clear difference between the Tracing Group and the No-Tracing Group, the fact that the difference of the Control Group only occurred in the Tracing Group suggests the importance of the perceptual intervention of the tracing method, which should lead to a detailed recognition of the writing process (as described in the Introduction). Because this discussion can be applied to any other dependent variable, hereafter, we will omit the same kind

of discussion related to the No-Tracing Group, unless there is a noteworthy outcome.

Inspiration

We calculated the mean and SD of the participants' scores for inspiration by viewing calligraphy (Cronbach's alpha coefficient was 0.93; **Table 1** and **Figure 3**). Tukey pairwise comparisons only revealed a significant difference between the Tracing Group and the Control Group, $b = 0.93$, $t(97) = 2.46$, $p = 0.041$. We did not find any significant differences between the other pairs; $b = 0.38$, $t(97) = 0.99$, $p = 0.583$ (tracing–no-tracing), $b = 0.55$, $t(97) = 1.47$, $p = 0.309$ (no-tracing–control). In addition, multiple regression analysis only showed a significant effect for the frequency of viewing ($p = 0.008$, **Table 2**), other than the above.

These results are in line with our hypothesis based on the dual focus model of inspiration (Ishiguro and Okada, 2020), which suggests that a comparative process with attention paid to both others and oneself leads to inspirational experiences. Further, we found that the frequency of viewing calligraphy correlates with the intensity of inspiration, which is consistent with Ishiguro and Okada (2019), who showed that art experience (including the frequency of art appreciation) correlates with inspiration.

Other Measurements

We calculated the mean and SD of participants' scores for empathy toward calligraphers (**Table 1** and **Figure 3**). Tukey pairwise comparisons only indicated a significant difference between the Tracing Group and the Control Group, $b = 0.70$, $t(97) = 2.40$, $p = 0.048$, but no significant differences between the other pairs; $b = 0.32$, $t(97) = 1.08$, $p = 0.527$ (tracing–no-tracing), $b = 0.38$, $t(97) = 1.32$, $p = 0.388$ (no-tracing–control). Multiple regression analysis revealed no other significant effect for the frequency of viewing and the interactions (**Table 2**).

These outcomes imply that the tracing method encourages viewers to empathize with calligraphers more than when no orientation is given. This seems to be mediated by reinforcement of the viewer's "dual focus" on self and others' writing process in the Tracing Group, which is also considered as a cause of inspiration. Given that the dual focus process (particularly the mental aspect of creation) inevitably involves assuming the artist's perspective, it is natural that the dual focus would foster empathy, which is closely associated with perspective-taking (Healey and Grossman, 2018).

We also calculated the mean and SD of the participants' scores for imagination (**Table 1**). Tukey pairwise comparisons showed no significant difference between any pair; $b = 0.67$, $t(97) = 2.11$, $p = 0.093$ (tracing–control), $b = -0.06$, $t(97) = -0.20$, $p = 0.978$ (tracing–no-tracing), $b = 0.74$, $t(97) = 2.33$, $p = 0.056$ (no-tracing–control). Multiple regression analysis only indicated a significant effect for the frequency of viewing ($p = 0.043$, **Table 2**).

Although these outcomes do not support the idea that differences in instruction and cognitive orientations influence the degree to which a viewer's imagination is triggered, since the differences in both pairs of tracing–control and no-tracing–control approach significance, it would be difficult to conclude their independence solely from those outcomes. In addition, we found that people who appreciate calligraphy frequently are more

likely to have richer imaginations. As imagination is one of the typical ways to enjoy art, this correlation is very natural.

STUDY 2

In Study 2, retaining the same hypotheses as Study 1, we added the heart rate analysis using smartphone-based PPG.

Method

Participants

A total of 81 participants took part through their own Web-connected computers instead of in a laboratory. We recruited them through a crowdsourcing service or via recruitment statements on social networking sites (SNS) and paid them 1,000 JPY (for participants from the crowdsourcing service) or a 1,500 JPY electronic gift certificate (for participants from the SNS) for completing the tasks, which took about an hour. Since physiological measures are sensitive to gender and age, we limited the participants to males between the ages of 18 and 39 ($M = 25.00$, $SD = 2.50$). In addition, all participants were native speakers of Japanese and had no visual or cardiac impairments. We only excluded one participant in advance from all analyses because of overly extreme responses for the psychological Likert items (only reporting either end of the range of the Likert scale for all items).

Stimuli

We chose the same four works of Japanese calligraphy as in Study 1 (Figure 1).

Experimental Conditions

While there were three conditions in Study 1, we excluded the No-Tracing Group from Study 2 because we did not find any major differences between the No-Tracing and the Control Group. Hence, we had the remaining two conditions: the Tracing Group and the Control Group. The instructions for each condition were the same as those used in Study 1; that is, we told participants in the Tracing Group to view calligraphy by tracing each line and imagining the writing process, and we told those in the Control Group to think and imagine whatever they wanted, without any orientation.

There was some difference in the sample size of each condition, despite random assignment ($n = 37$ in the Tracing Group; $n = 43$ in the Control Group). Further, this was intertwined with the place from where they were recruited (crowdsourcing service: $n_t = 8$, $n_c = 18$; SNS: $n_t = 29$, $n_c = 25$; the “t” or “c” suffix indicates the “Tracing Group” or the “Control Group”). This difference was caused by the discrepancy between conditions for number of participants who had agreed to participate in the experiment, but did not actually finish (not counted in the values above). Since we did not detect such cases in Study 1, the procedures for measuring the PPG in Study 2 may have been complicated beyond their expectations so as to lead to this situation. In order to partial out these effects, in the regression analyses, we included a dummy variable for the place of recruitment as a covariate (see the Results and discussion section for details).

Procedure

The procedure as a whole was very similar to that of Study 1. The participants read a broad description of the tasks, provided informed consent, and started the Web-based experiment, which included instructions for calligraphy, the stimulus presentation, and completing the Likert items. The only important difference from Study 1 was that for Study 2, the participants were required to use their own smartphones with cameras to record a video of the skin of the tip of their left index finger. Specifically, for the first task of the experiment, we asked them to first free up space on their device’s memory to record a video, and then asked them to record their skin for 30 s as a trial using a diagram (Figure 4 depicts some examples). Upon finishing the trial recording, they uploaded the video file to a server specified by the experimenter without pause. The experimenter checked the uploaded file for any flaws in the video as soon as possible. If there was no problem, he sent a message to continue the experiment; otherwise, he requested that the participant re-take the video until there were no more issues with it (the first recording needed to have been completed correctly, because it was also used as a baseline in the following analysis). Subsequently, the participants read a text suggesting how they should view calligraphy (depending on the conditions), in the same way as Study 1. From this point, the sequence of the experiment returned to the same form as in Study 1. When the experimental stimuli were presented sequentially for 3 min after that, the participants recorded their fingers over the entire period of stimulus presentation, stopping and re-taking the recording each time. Since the participants had to keep their left hand on the camera lens of their smartphones, they used their right hand to operate the keyboard to switch between images. We synchronized the timing of data acquisition between stimuli presentation and PPG recordings using electronic sounds from each participant’s computer (which ran the program for the experiment); the sounds were heard at the beginning and end of each stimulus presentation and recorded in the video of the skin.

Because there is a circadian rhythm of heart rate (e.g., Massin et al., 2000) – which should be controlled for the sake of data quality – we limited the time that participants could start the

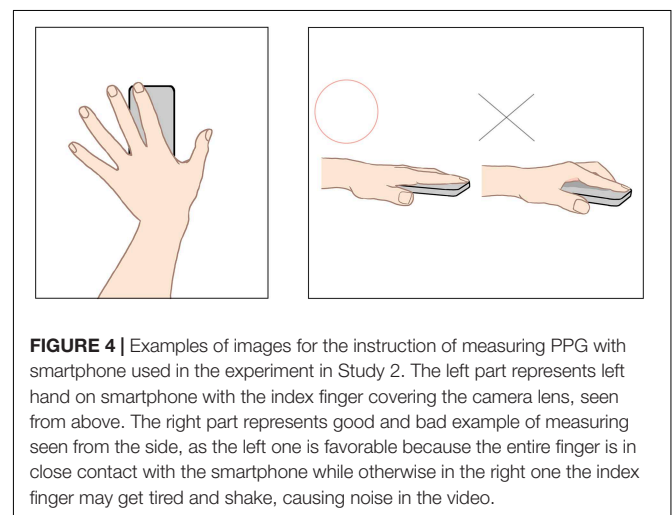


FIGURE 4 | Examples of images for the instruction of measuring PPG with smartphone used in the experiment in Study 2. The left part represents left hand on smartphone with the index finger covering the camera lens, seen from above. The right part represents good and bad example of measuring seen from the side, as the left one is favorable because the entire finger is in close contact with the smartphone while otherwise in the right one the index finger may get tired and shake, causing noise in the video.

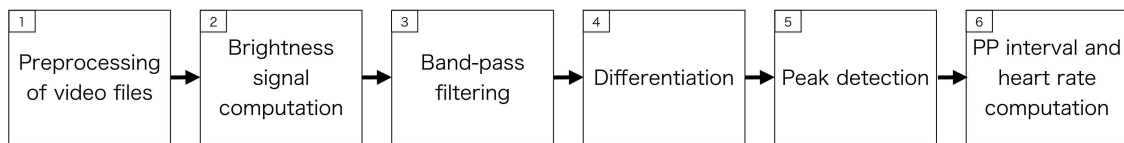


FIGURE 5 | The main processing blocks of the heart rate estimation.

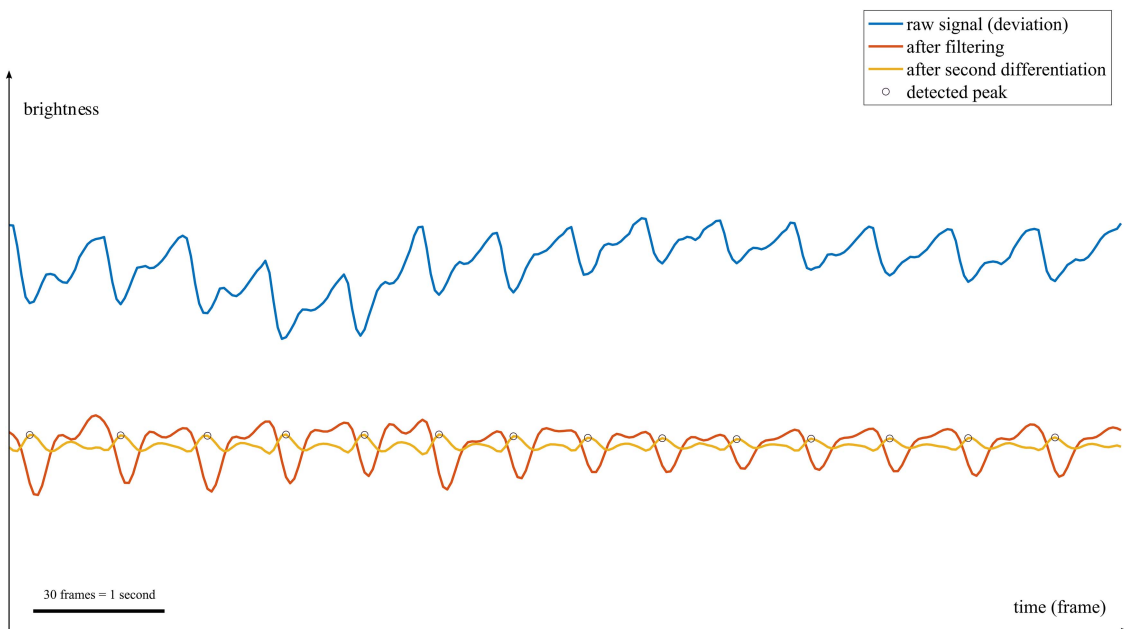


FIGURE 6 | An example of PPG waveforms. The horizontal axis represents time (the sequence of frames in a video file), and the vertical axis represents image brightness reflecting blood flow. The length of the black horizontal bar at the bottom left of the figure corresponds to 30 frames (=1 s). The blue line shows the raw signal in the second block of **Figure 5** (each value of brightness on the vertical axis is a deviation from the mean value, that is, the value from which we subtracted its mean value). The red line shows the band-pass-filtered signal in the third block. The yellow line shows the second derivative signal in the fourth block. The black circles show every detected peaks in the fifth block.

experiment to between 13:00 and 18:00. All other parts of the procedure were the same as in Study 1.

Analysis of PPG

Figure 5 portrays the processing for estimating the average heart rate. In each block, the process was as follows: (1) We first preprocessed all video files submitted by the participants to convert them (the videos were originally shot on different smartphone models and using different frame rate settings) to 30 frames per second. At the same time, we identified the frame numbers in each video where the electronic sounds, emanating from the experimental program, were recorded. We then used the frame numbers to trim the videos. (2) For every frame, we calculated the mean value of all pixels of red intensity in RGB, regarding it as an index of “brightness” at each time point. The change in brightness over time is thought to reflect the change in blood flow. (3) We band-pass filtered the obtained signal over 0.67–3.83 Hz (\approx 40–230 beats per minute) using a second-order Butterworth filter to reduce noise. (4) We differentiated the waveform of brightness twice to make the peak derived

from the heartbeat more prominent (**Figure 6** displays this contrast; see Guede-Fernández et al., 2020 for an example of using differentiation). (5) We detected the peaks using the “findpeaks” function in MATLAB. (6) Based on the detected peaks, we calculated the pulse-to-pulse intervals (PPI). Because there were often problems with skipping or double counting peaks (that is, our algorithm often failed to find a peak, or counted two peaks at one cycle of heartbeat in the previous process due to noisy data), we had to calculate an average value that took these variations into account, instead of the mean value. To address this, we defined M' for each period such that $f(M')$ below reached the minimum, with M' ranging from 15 to 40 frames in increments of 0.01.

$$f(M') = \sum_k \min_{a \in \{2, 1, 0.5\}} (aPPI_k - M')^2$$

Subsequently, we computed the average heart rate h (beats per minute; identified with the pulse rate in this study) as the inverse of M' .

We determined the average heart rate for (a) the baseline period; (b) the first time period; and (c) the second time period. (a) The baseline period corresponds to the first 30 s of trial measurement, during which the participants were not presented with stimuli of calligraphy works. The (b) first time period and (c) second time period correspond to the first and second halves of each 3-min presentation of the calligraphy work, respectively. We cut off the first 5 s of (b) and the last 5 s of (c) to remove artifacts resulting from bandpass filtering.

After calculating the average heart rate for each period, we excluded any clearly inaccurate or doubtful data from the analysis of heart rate (for the criteria, see **Supplementary Table 4**). As a result, we did not include 40.00% of participants in that analysis (see **Supplementary Text 1** and **Supplementary Table 5** for the reliability of the analytical procedures).

Results and Discussion

For each variable measured more than once, we calculated each participant's score by averaging the stimuli in each session, which, in turn, we used for statistical analysis as a data point. Unless otherwise stated, we examined each dependent variable using multiple regression analysis. We included a dummy variable for the condition of instruction, the score for the frequency of viewing calligraphy, the interaction term of both variables, and (additionally in Study 2) the place of recruitment as predictor variables. Different from Study 1, we did not include scores of liking in the analysis, while this time the length of time spent learning calligraphy was included as a predictor variable in some post-hoc analyses.

Validation of Assignments and Experimental Manipulation

To check whether we appropriately assigned the participants in terms of their attitudes toward calligraphy, we calculated the mean and SD of their scores for the frequency of viewing calligraphy for each condition (see **Table 3**). As the independent *t*-test showed no significant differences among them, $t(77) = 0.813$, $p = 0.419$, there is no evidence to suggest that the assignments were biased.

Further, to confirm whether the manipulation of the instructions influenced the participants' cognitive processes as expected, we calculated the mean and SD of their scores for (a) the degree to which they were aware of the physical process of writing calligraphy (awareness of physical creation) and (b) the degree to which they were aware of the mental process of writing calligraphy (awareness of mental creation; see **Table 3**). As for (a) awareness of physical creation, the outcomes showed that the effects were significant for the condition, $b = 0.76$, $t(74) = 2.46$, $p = 0.016$, and the frequency of viewing calligraphy, $b = 0.76$, $t(74) = 2.55$, $p = 0.013$, but the other effects were insignificant (**Table 4**). As for (b) awareness of mental creation, the results indicated that the effects were insignificant for all predictors (**Table 4**). These findings imply that the manipulation in the current study had a similar function to that of Study 1, and changed the viewer's recognition of the writing process of calligraphy to some extent, but did not change as radically as that of Study 1, since there was no significant difference between the

conditions in terms of mental creation (at least consciously). The reason why the effect was weakened (despite using the same texts regarding calligraphy) may be that the relatively complicated procedures of using a smartphone to measure PPG distracted the participants from contemplating the calligraphy. Therefore, we considered the manipulation to be validated enough to create some differences between conditions, but less effective than Study 1, perhaps making some effects unobservable.

Aesthetic Impression

We calculated the mean and SD of the participants' scores for the admiration of calligraphy (**Table 3** and **Figure 3**). Multiple regression analysis only revealed significant effects for the condition, $t(74) = 2.01$, $p = 0.048$ and the frequency of viewing, $t(74) = 2.10$, $p = 0.039$ (**Table 4**).

This outcome is largely consistent with our hypothesis and the result in Study 1 in terms of supporting the idea that admiration is elicited by the change in recognition of the writing process of calligraphy. Although there is another significant effect of the frequency of viewing, since that effect in Study 1 was also near the significant level ($p = 0.093$), there seems to be no essential difference between Study 1 and 2. In addition, because expertise often has a positive impact on the viewer's aesthetic impression (cf. Leder et al., 2012; van Paasschen et al., 2015; Matsumoto and Okada, 2019), the fact that the frequency of viewing positively correlates with admiration is not surprising.

Inspiration and Other Rating Items

We calculated the mean and SD of participants' scores on inspiration by viewing calligraphy (Cronbach's alpha coefficient was 0.84; **Table 3**). Multiple regression analysis only revealed a significant effect for the frequency of viewing, $t(74) = 3.44$, $p < 0.001$ (**Table 4**). Likewise, we calculated the mean and SD of participants' scores for empathy and imagination (**Table 3**). For both variables, multiple regression analysis showed no significant effect for all predictors (**Table 4**).

Based on these results, we can conclude that the manipulation of instructions and orientation of cognitive processes for viewing calligraphy had no observable effects other than admiration, the construct most closely related to the recognition of the writing process of calligraphy among all items in Study 2. The other effects of the tracing method suggested by Study 1 were not shown in Study 2; this seems to be because they were diluted by the complicated procedures of

TABLE 3 | Descriptive statistics for all rating values in Study 2.

Variable	Tracing group	Control group
Frequency of viewing	0.42 (0.65)	0.30 (0.60)
Awareness of physical creation	4.36 (1.11)	3.74 (1.27)
Awareness of mental creation	3.37 (1.56)	3.09 (1.78)
Admiration	2.89 (1.07)	2.50 (1.27)
Inspiration	2.64 (1.22)	2.74 (1.45)
Empathy	2.57 (1.14)	2.33 (1.34)
Imagination	3.22 (1.10)	2.95 (1.15)

Standard deviations appear in parentheses.

TABLE 4 | Multiple regression analyses in Study 2.

Dependent variable	Predictor	<i>b</i>	<i>t</i>	<i>p</i>
Awareness of physical creation	Intercept	3.61***	14.08	<0.001
	Condition (Tracing: 1, Control: 0)	0.76*	2.46	0.016
	Frequency of viewing	0.76*	2.55	0.013
	Interaction (condition × frequency)	−0.56	−1.31	0.194
	Place of recruitment	−0.18	−0.65	0.521
Awareness of mental creation	Intercept	3.04***	8.33	<0.001
	Condition (Tracing: 1, Control: 0)	0.11	0.24	0.812
	Frequency of viewing	0.63	1.48	0.143
	Interaction (condition × frequency)	0.31	0.52	0.607
	Place of recruitment	−0.24	−0.59	0.555
Admiration	Intercept	2.61***	10.44	<0.001
	Condition (Tracing: 1, Control: 0)	0.60*	2.01	0.048
	Frequency of viewing	0.61*	2.10	0.039
	Interaction (condition × frequency)	−0.33	−0.81	0.422
	Place of recruitment	−0.50	−1.81	0.074
Inspiration	Intercept	2.70***	10.12	<0.001
	Condition (Tracing: 1, Control: 0)	0.10	0.30	0.761
	Frequency of viewing	1.07***	3.44	<0.001
	Interaction (condition × frequency)	−0.39	−0.89	0.375
	Place of recruitment	−0.48	−1.63	0.106
Empathy	Intercept	2.50***	9.44	<0.001
	Condition (Tracing: 1, Control: 0)	0.19	0.60	0.551
	Frequency of viewing	0.50	1.64	0.106
	Interaction (condition × frequency)	0.22	0.49	0.623
	Place of recruitment	−0.54	−1.86	0.067
Imagination	Intercept	2.92***	11.86	<0.001
	Condition (Tracing: 1, Control: 0)	0.33	1.11	0.271
	Frequency of viewing	0.53	1.86	0.067
	Interaction (condition × frequency)	−0.15	−0.37	0.713
	Place of recruitment	−0.24	−0.87	0.386
Heart rate change ratio for first time period: $\log(h_{p1} / h_b)$	Intercept	0.00	0.18	0.859
	Condition (Tracing: 1, Control: 0)	−0.01	−0.54	0.592
	Frequency of viewing	−0.03	−1.06	0.297
	Interaction (condition × frequency)	0.05	1.33	0.191
	Place of recruitment	0.01	0.51	0.610
Heart rate change ratio for second time period: $\log(h_{p2} / h_b)$	Intercept	0.00	0.20	0.844
	Condition (Tracing: 1, Control: 0)	−0.02	−0.67	0.505
	Frequency of viewing	−0.05*	−2.13	0.039
	Interaction (condition × frequency)	0.08*	2.28	0.028
	Place of recruitment	0.01	0.59	0.559
(Post hoc analysis for low-frequency group)	Intercept	0.02	0.92	0.366
	Condition (Tracing: 1, Control: 0)	−0.05*	−2.08	0.047
	Time spent learning calligraphy	−0.05*	−2.91	0.007

(Continued)

TABLE 4 | Continued

Dependent variable	Predictor	<i>b</i>	<i>t</i>	<i>p</i>
(Post hoc analysis for high-frequency group)	Interaction (condition × time)	0.08	1.93	0.064
	Place of recruitment	0.02	1.04	0.310
	Intercept	0.03	0.58	0.573
	Condition (Tracing: 1, Control: 0)	0.07	1.05	0.320
	Time spent learning calligraphy	−0.05	−2.15	0.057
	Interaction (condition × time)	0.00	0.00	0.998
	Place of recruitment	−0.00	−0.07	0.949

p* < 0.05; *p* < 0.01; ****p* < 0.001.

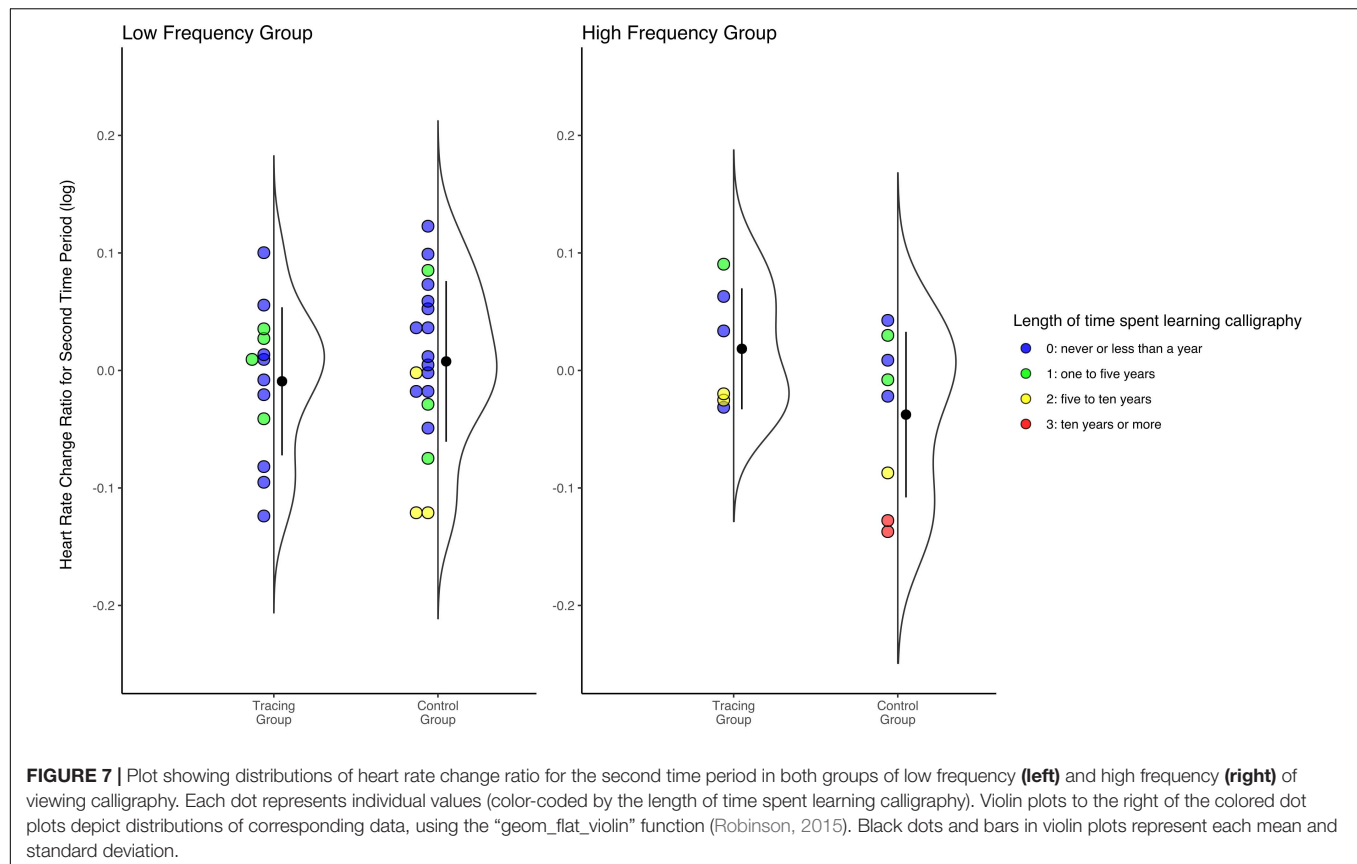
measuring PPG, as well as the index of mental creation described in “Validation of assignments and experimental manipulation.” In addition, any psychological variables used in this study except inspiration are measured by single-item scales, and their reliability cannot be assessed by Cronbach’s alpha coefficient but may possibly be low. This lack of reliability may potentially account for the difference between the two studies. Therefore, although the effects on inspiration and empathy observed in Study 1 are not so robust (at least as much as admiration), and the effects on imagination suggested in Study 1 are not supported either, whether or not these variables are influenced by the way one recognizes the writing process of calligraphy should not be determined based on the current findings alone.

Heart Rate

We calculated the mean and SD of the participants’ average heart rate from PPG signals for the baseline (h_b), the first time period (h_{p1}), and the second time period (h_{p2}). In the Tracing Group, the mean (SD) of h_b , h_{p1} , and h_{p2} was 81.52 (10.50), 82.37 (13.10), and 82.44 (12.41) beats per minute, respectively. In the Control Group, this was 79.27 (8.47), 79.40 (8.24), and 78.73 (7.31) beats per minute, respectively. Subsequently, we calculated the ratio of h_{p1} and h_{p2} to h_b for each participant. We used the natural logarithmic value of this ratio as the response variable in the multiple regression analysis, with the same predictor variables as other psychological measurements. While there was no significant effect for all predictors for the first period, for the second time period, the effect were significant for the frequency of viewing, $b = -0.05$, $t(42) = -2.13$, $p = 0.039$ and the interaction, $b = 0.08$, $t(42) = 2.28$, $p = 0.028$, but the other effects were insignificant (Table 4). To examine the nature of the interaction effect via post-hoc analysis, we further divided the participants into two groups based on their scores for the frequency of viewing (the low-frequency group of participants scoring 0: those who “almost never” have the opportunity to appreciate works of calligraphy, $n_l = 13$, $n_c = 19$; and the high-frequency group of participants scoring 1 or greater: those who have the opportunity to appreciate works of calligraphy “once every few years” or more, $n_l = 7$, $n_c = 8$; see also **Supplementary Table 1**). We carried out multiple regression analysis for each group with four predictor variables (length of time spent learning

calligraphy and its interaction term with condition were added to condition and place of recruitment) and the same dependent variable. While the condition had a significant effect in the low-frequency group, $b = -0.05$, $t(27) = -2.08$, $p = 0.047$, it had no significant effect in the high-frequency group, $b = 0.07$, $t(10) = 1.05$, $p = 0.320$ (Figure 7). In addition, there was significant effect of length of time spent learning calligraphy only in the low-frequency group, $b = -0.05$, $t(27) = -2.91$, $p = 0.007$. The effect of place of recruitment was insignificant for both groups (Table 4). This indicates that the influence of the condition in all samples was especially derived from differences among participants who did not have the habit of appreciating calligraphy.

If we focus on the simple main effect for the low-frequency group, the current result can be interpreted in a way that is straightforwardly consistent with previous studies. The extant literature suggests that heart rate during task execution is positively correlated with boredom (London et al., 1972; Merrifield and Danckert, 2014; Raffaelli et al., 2018). This trend also applies to the presentation of aesthetic objects; that is, attention to (and interest in) stimuli that are associated with lower heart rate responses (Libby et al., 1973). Therefore, the currently observed effect of the changing ratio of heart rate between conditions may mean that the participants in the Tracing Group remained interested in the art without becoming bored, which resulted from the tracing method. Additionally, there was also the effect of time spent learning calligraphy, which affected the low-frequency group in the same way that the tracing method did. That is, the longer a participant spent learning calligraphy, the lower the participant’s heart rate, which resulted from the participant’s maintaining interest in the stimuli. To put it in another way, the current results of the PPG suggest that, in terms of how calligraphy is viewed, the tracing method may, to some extent, substitute for the effects obtained by the long-term study of calligraphy. We should also note that condition had no effect during the first time period. One possibility is that participants in both the Tracing Group and Control Group viewed works without boredom in the beginning, but those in the Control Group gradually became bored because they lacked of an appropriate strategy for viewing the art, whereas those in the Tracing Group remained interested.



There are likely a few reasons why the results concerning the difference in change in heart rate ratio between conditions differed between the low-frequency and high-frequency groups. First, there were not enough participants in the high-frequency group (because the majority of participants had seldom engaged in calligraphy appreciation), which may be the main reason why the post-hoc analysis in the high-frequency group did not yield any significant effect. Second, there is the possibility that, especially for people who are accustomed to regular calligraphy appreciation, introducing a new technique such as the tracing method has a neutral or even obstructive effect. The tracing method might interfere with the viewer's usual behavior patterns, which may have been established by frequent exposure to calligraphy. Even though the lack of significance for the high-frequency group in the post-hoc analysis may owe to the small sample size, based on the above view concerning the interaction between frequency of appreciation and experimental intervention, it is also possible that even with more participants in the high-frequency group, the heart rate of the Tracing Group might be nearly the same or even higher (rather than lower) than that of the Control Group because of the decreased interest of experienced participants.

By reflecting on the current analysis of PPG, we can see that the smartphone-based PPG can be used effectively in a Web experiment. At the same time, however, there are limitations regarding the accuracy of this measurement. The fact that we

could not use about 40% of the data of all participants for the analysis of heart rate, due to various noises, clearly shows the magnitude of this problem. Hence, future research should refine the smartphone-based PPG method and involve follow-up testing of the results using more reliable equipment.

GENERAL DISCUSSION

We will summarize the findings of Studies 1 and 2 as follows (see Table 5 for a comparison of results). By orientating viewers' cognition about the recognition of the process of creating artworks through manipulated instructions about how to perceive the art (i.e., suggesting the “tracing method,” the procedure of viewing calligraphy in the order in which lines were drawn as if by “tracing” with one's eyes), viewers will have deeper admiration. This outcome is not only consistent with both of the experiments, but also supports the findings of Matsumoto and Okada (2019), whereby detailed recognition of the process of creating artworks leads to admiration. As for inspiration and empathy, Study 1 showed significant effects of the tracing method, while Study 2 did not. Since it is likely that the complicated procedures of Study 2 distracted the participants from appreciating the calligraphy, the relationship between those variables and recognition of the process of creating artworks requires further investigation. Study 2 further probed viewers' physiological responses related to the autonomic

TABLE 5 | Comparison of results of Likert scales in Studies 1 and 2.

Variable	Study 1	Study 2
Frequency of viewing	n.s.	n.s.
Awareness of physical creation	Condition	Condition frequency of viewing
Awareness of mental creation	Condition frequency of viewing	n.s.
Admiration	Condition	Condition frequency of viewing
Liking of a work	n.s.	(Not included in analysis)
Inspiration	Condition frequency of viewing	Frequency of viewing
Empathy	Condition	n.s.
Imagination	Frequency of viewing	n.s.

n.s. = not significant for all predictor variables. In cases where there is any significant effect in each result of regression analysis, only significant predictor variables are listed. See **Tables 1–4** and **Figures 3, 7** for details of each result.

nervous system by smartphone-based PPG. As a result – especially for participants who view calligraphy relatively less frequently – we found a significant difference between conditions in the ratio of the average heart rate in the second half of the duration of viewing stimuli compared to the baseline, with lower ratio for those who were given tracing method. This suggests that the tracing method may allow viewers to continue contemplating calligraphy without growing bored, even during the latter half of a fairly long viewing time (although this is only one possibility).

This study contributes to the literature on the psychology of aesthetics in the following ways. First, like previous research (Matsumoto and Okada, 2019), this study demonstrates that the role played in the art-viewing process by recognition of the process of creating artworks – which has not been highlighted in the dominant models of art-viewing (e.g., Leder et al., 2004; Pelowski et al., 2017) – is large and robust in both studies. Our investigation into recognition of the process of creating artworks was made possible in large part by the characteristics of Chinese/Japanese calligraphy, as they enable the method of viewing by tracing, which is hard to apply directly to other kinds of art (such as painting). This does not mean that application of our findings must be limited to the field of calligraphy, because the mental mechanism of feeling admiration – imagining the process of another person creating an artwork, making a social comparison between oneself and others, and then feeling admiration – can be assumed in any kind of art-viewing (Matsumoto and Okada, 2019). Rather, this study is an example of how psychological research on the arts in relatively atypical artistic domains can provide deeper insight into vital aspects of the cognitive process of art-viewing, which have yet to be adequately addressed. Second, this study revealed the inspirational process caused by viewing art. The way in which art appreciation leads to inspiration has rarely been dealt with in experimentally controlled situations; our study is also valuable in that regard (see Ishiguro and Okada, 2019, for an example of a questionnaire survey). Up until now, in virtually all models of art appreciation, interest has been exclusively focused on the

range of time between viewers' initial perceptions of art and their impression formation. In the current study, we made a first step to empirically explore the mental process, starting from art appreciation and expanding to other activities of viewers. Third, from a practical perspective, this study has implications for how to display art to allow viewers to contemplate it in such a way that even novices can deeply appreciate it. Although recognition of the process of creating artworks has an important role in art-viewing, the results of Study 1 imply that simply thinking about that process while looking at them will not make any notable difference. On the other hand, at least for calligraphy, we can expect that an instruction as simple as the “tracing method” will satisfy viewers to some degree, even in a practical situation. Since this is not only effective but also simple and short, our findings are practically significant for art education or museum management. Moreover, it is quite possible that as with the tracing method in calligraphy, guidance orientating viewers toward a detailed recognition of the process of creation is generally effective for viewing other types of art. Specific methods can be addressed in future research.

Further, this study has significance in that it introduces a new perspective to graphonomics. By applying the theory of the psychology of aesthetics and conducting experiments, we clarified a part of the process of forming aesthetic impressions (especially admiration) of other people's handwriting, which remains almost unexamined in graphonomics. Since admiration and inspiration are classified as assimilative emotions rather than contrastive ones (Smith, 2000), when people aim to write well, they might tend to model themselves after the handwriting of others whom they admire, whether consciously or not. Therefore, starting from this study and performing more detailed analysis to find out how the physical characteristics of handwriting are associated with the formation of mental representations about the creative process, it may become clearer what kinds of handwriting we admire, and what we internally represent as a goal when writing. This will have important implications for graphonomic research, which addresses the process of human writing. The possibility that written language can have “paralinguistic” cues and communicate information about the writer has not received much attention in recent graphonomic research. This may be due to the longstanding criticism of graphology that the image of the writer that the reader constructs, using the letters as clues, does not reflect reality (King and Koehler, 2000; Simner and Goffin, 2003). However, this study suggests that – especially in the art of calligraphy – the information about the author conveyed by the characters is closely linked to aesthetic feelings. Hence, by introducing a communicative perspective, future graphonomic studies are likely to reveal new aspects of human activities linked to written language. At the same time, this study shows the communicative aspect of art-related activities, which has been drawing increasing attention (e.g., Dolese, 2015). Our findings span the two subfields of communication research – written language and art – and should contribute to the holistic understanding of human communication.

This study is also very valuable as a practical example of measuring a physiological parameter, without any specialized equipment, in a remote environment. The outcome of average

heart rate in this study is consistent with existing findings. In this respect, our study supports the usefulness of smartphone-based PPG in empirical aesthetics or other kinds of psychophysiological research. This will be especially effective when a researcher cannot conduct face-to-face experiments with participants or examine their daily activities in vivo. On the other hand, smartphone-based PPGs tend to be noisy and require caution in their use in situations without face-to-face supervision by an experimenter. Future research should consider elements such as the type of device suitable for measurement and the wording of instructions to ensure that participants can perform the measurement accurately.

Finally, future research could go in several directions: (1) examining different forms of characters and letters from the ones dealt with in this study, namely, handwriting in ordinary situations or printed fonts; (2) confirming whether presenting the actual process of creating artworks through media (such as video) will have a similar effect to that observed in this study; (3) adding a rating scale of boredom and examining correlations between it and heart rate to verify the current discussion on heart rate; (4) measuring other indices (such as skin conductance or eye movements) for a more multifaceted understanding of art appreciation and impression formation for letters.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and

accession number(s) can be found below: <https://github.com/psychologyKM/2020experiments>.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

KM designed and performed the experiments, analyzed the data, and co-wrote the manuscript. TO supervised the research and co-wrote the manuscript.

ACKNOWLEDGMENTS

We thank Editage (www.editage.jp) for English language editing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnhum.2021.654610/full#supplementary-material>

REFERENCES

- Allport, G. W., and Vernon, P. E. (1933). *Studies in Expressive Movement*. New York, NY: Hafner.
- Bullot, N. J., and Reber, R. (2013). The artful mind meets art history: toward a psycho-historical framework for the science of art appreciation. *Behav. Brain Sci.* 36, 123–137. doi: 10.1017/s0140525x12000489
- Craig, R. T. (1999). Communication theory as a field. *Commun. Theory* 9, 119–161. doi: 10.1111/j.1468-2885.1999.tb00355.x
- Critchley, H. D., Eccles, J., and Garfinkel, S. N. (2013). “Interaction between cognition, emotion, and the autonomic nervous system,” in *Handbook of Clinical Neurology*, eds R. M. Buijs and D. F. Swaab (Amsterdam: Elsevier), 59–77. doi: 10.1016/b978-0-444-53491-0.00006-7
- de Leeuw, J. R. (2015). jsPsych: a javascript library for creating behavioral experiments in a Web browser. *Behav. Res. Methods* 47, 1–12. doi: 10.3758/s13428-014-0458-y
- Dolese, M. J. (2015). *Art as communication: Employing Gricean Principles of Communication as a Model for art Appreciation*. [Ph.D. dissertation]. New York, NY: City University of New York.
- Fukuda, R., and Fukuda, T. (2009). Comparison of reading capacity for Japanese German, and English. *Percept. Mot. Ski.* 108, 281–296. doi: 10.2466/pms.108.1.281-296
- Garcia-Agundez, A., Dutz, T., and Goebel, S. (2017). Adapting smartphone-based photoplethysmography to suboptimal scenarios. *Physiol. Meas.* 38, 219–232. doi: 10.1088/1361-6579/aa51db
- Gatten, A. (1986). Weird ladies: narrative strategy in the Genji monogatari. *J. Assoc. Teach. Jpn.* 20, 29–48. doi: 10.2307/489516
- Gil, E., Orini, M., Bailon, R., Vergara, J. M., Mainardi, L., and Laguna, P. (2010). Photoplethysmography pulse rate variability as a surrogate measurement of heart rate variability during non-stationary conditions. *Physiol. Meas.* 31, 1271–1290. doi: 10.1088/0967-3334/31/9/015
- Guede-Fernández, F., Ferrer-Mileo, V., Mateu-Mateus, M., Ramos-Castro, J., García-González, M. Á., and Fernández-Chimeno, M. (2020). A photoplethysmography smartphone-based method for heart rate variability assessment: device model and breathing influences. *Biomed. Signal. Proces* 57:101717. doi: 10.1016/j.bspc.2019.101717
- Hall, J. A., Horgan, T. G., and Murphy, N. A. (2019). Nonverbal communication. *Annu. Rev. Psychol.* 70, 271–294.
- Healey, M. L., and Grossman, M. (2018). Cognitive and affective perspective-taking: evidence for shared and dissociable anatomical substrates. *Front. Neurol.* 9:491. doi: 10.3389/fneur.2018.00491
- Hellbernd, N., and Sammler, D. (2016). Prosody conveys speaker's intentions: acoustic cues for speech act perception. *J. Mem. Lang.* 88, 70–86. doi: 10.1016/j.jml.2016.01.001
- Hothorn, T., Bretz, F., and Westfall, P. (2008). Simultaneous inference in general parametric models. *Biom. J.* 50, 346–363. doi: 10.1002/bimj.200810425
- Hsu, J. (1996). *Multiple Comparisons: Theory and Methods*. Boca Raton, FL: CRC Press.
- Ishiguro, C., and Okada, T. (2015). “The effects of art experience, competence in artistic creation, and methods of appreciation on artistic inspiration,” in *Proceedings of the 32nd Conference of Japanese Cognitive Science Society 205–213 (JCSS, 2015)*, Chiba.
- Ishiguro, C., and Okada, T. (2019). “How does art appreciation promote artistic inspiration?,” in *Proceedings 41st Annual Meeting of the Cognitive Science Society 3286–3291 (CogSci, 2019)*, Montreal, QC.
- Ishiguro, C., and Okada, T. (2020). How does art viewing inspire creativity? *J. Creat. Behav.* [Epub ahead of print].
- Ishikawa, K. (2011). *Taction: The Drama of the Chisel, Stylus, Brush in Oriental Calligraphy*. Tokyo: International House of Japan.
- Johnstone, T., and Scherer, K. R. (2000). “Vocal communication of emotion,” in *Handbook of Emotions*, 2nd Edn, eds M. Lewis

- and J. M. Haviland-Jones (New York, NY: Guilford Press), 220–235.
- Jones, S. E., and LeBaron, C. D. (2002). Research on the relationship between verbal and nonverbal communication: emerging integrations. *J. Commun.* 52, 499–521. doi: 10.1111/j.1460-2466.2002.tb02559.x
- Jucker, J. L., Barrett, J. L., and Wlodarski, R. (2014). “I just don’t get it”: perceived artists’ intentions affect art evaluations. *Empir. Stud. Arts* 32, 149–182. doi: 10.2190/em.32.2.c
- Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: towards a unified theory of musical emotions. *Phys. Life Rev.* 10, 235–266. doi: 10.1016/j.plrev.2013.05.008
- Juslin, P. N., and Laukka, P. (2003). Communication of emotions in vocal expression and music performance: different channels, same code? *Psychol. Bull.* 129, 770–814. doi: 10.1037/0033-2909.129.5.770
- Kamshilin, A. A., Nippola, E., Sidorov, I. S., Vasilev, P. V., Erofeev, N. P., Podolian, N. P., et al. (2015). A new look at the essence of the imaging photoplethysmography. *Sci. Rep.* 5:10494.
- Kilyeni, A. (2009). Nonverbal communication in print ads. *Prof. Commun. Trans. Stud.* 2, 17–24.
- King, R. N., and Koehler, D. J. (2000). Illusory correlations in graphological inference. *J. Exp. Psychol. Appl.* 6, 336–348. doi: 10.1037/1076-898x.6.4.336
- Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: a review. *Biol. Psychol.* 84, 394–421. doi: 10.1016/j.biopsycho.2010.03.010
- Kurylyak, Y., Lamonaca, F., Grimaldi, D., and Duro, F. J. (2012). “Smartphone-based photoplethysmogram measurement,” in *Digital Image, Signal and Data Processing for Measurement Systems*, eds R. J. Duro and F. López-Peña (Denmark: River Publishers), 135–164.
- Leder, H., Belke, B., and Oeberst, A. (2004). A model of aesthetic appreciation and aesthetic judgments. *Br. J. Psychol.* 95, 489–508. doi: 10.1348/0007126042369811
- Leder, H., Gerger, G., Dressler, S. G., and Schabmann, A. (2012). How art is appreciated. *Psychol. Aesthet. Creat. Arts* 6, 2–10. doi: 10.1037/a0026396
- Libby, W. L. Jr., Lacey, B. C., and Lacey, J. I. (1973). Pupillary and cardiac activity during visual attention. *Psychophysiology* 10, 270–294. doi: 10.1111/j.1469-8986.1973.tb00526.x
- Littlejohn, S. W., and Foss, K. A. (2010). *Theories of Human Communication*. Long Grove, IL: Waveland Press.
- Lohani, M., Payne, B. R., and Strayer, D. L. (2019). A review of psychophysiological measures to assess cognitive states in real-world driving. *Front. Hum. Neurosci.* 13:57. doi: 10.3389/fnhum.2019.00057
- London, H., Schubert, D. S., and Washburn, D. (1972). Increase of autonomic arousal by boredom. *J. Abnorm. Psychol.* 80, 29–36. doi: 10.1037/h0033311
- Loveday, L. (1982). *The Sociolinguistics of Learning and Using a Non-native Language*. Oxford: Pergamon Press.
- Luangrath, A. W., Peck, J., and Barger, V. A. (2017). Textual paralinguistic and its implications for marketing communications. *J. Consum. Psychol.* 27, 98–107. doi: 10.1016/j.jcps.2016.05.002
- Massin, M. M., Maeyns, K., Withofs, N., Ravet, F., and Gérard, P. (2000). Circadian rhythm of heart rate and heart rate variability. *Arch. Dis. Child.* 83, 179–182. doi: 10.1136/adc.83.2.179
- Matsumoto, K., and Okada, T. (2019). Viewers recognize the process of creating artworks with admiration: evidence from experimental manipulation of prior experience. *Psychol. Aesthet. Creat. Arts* 15, 352–362. doi: 10.1037/aca0000285
- Merrifield, C., and Danckert, J. (2014). Characterizing the psychophysiological signature of boredom. *Exp. Brain Res.* 232, 481–491. doi: 10.1007/s00221-013-3755-2
- Nell, V. (1988). The psychology of reading for pleasure: needs and gratifications. *Read. Res. Q.* 23:6. doi: 10.2307/747903
- Okada, T., and Ishibashi, K. (2017). Imitation, inspiration, and creation: cognitive process of creative drawing by copying others’ artworks. *Cogn. Sci.* 41, 1804–1837. doi: 10.1111/cogs.12442
- Oshiki, H., Terashima, N., and Koike, M. (2010). Tegaki bunsho ni okeru paralinguistic teki youso niyoru dentatsu ni kansuru kisoteki kenkyu [A basic study on communication by paralinguistic elements in handwritten documents]. *Shodo Shosha Kyoiku Kenkyu* 24, 21–32.
- Pelowski, M., and Akiba, F. (2011). A model of art perception, evaluation and emotion in transformative aesthetic experience. *New Ideas Psychol.* 29, 80–97. doi: 10.1016/j.newideapsych.2010.04.001
- Pelowski, M., Cabai, G., Brinkmann, H., Mikuni, J., Hegelmaier, L. M., Forster, M., et al. (2020). The kitsch switch—or (when) do experts dislike thomas kinkade art? A study of time-based evaluation changes in top-down versus bottom-up assessment. *Psychol. Aesthet. Creat. Arts* [Epub ahead of print].
- Pelowski, M., Markey, P. S., Forster, M., Gerger, G., and Leder, H. (2017). Move me, astonish me...delight my eyes and brain: the Vienna integrated model of top-down and bottom-up processes in art perception (VIMAP) and corresponding affective, evaluative, and neurophysiological correlates. *Phys. Life Rev.* 21, 80–125. doi: 10.1016/j.plrev.2017.02.003
- Pennycook, A. (1985). Actions speak louder than words: paralinguistic, communication, and education. *TESOL Q.* 19, 259–282. doi: 10.2307/3586829
- Raffaelli, Q., Mills, C., and Christoff, K. (2018). The knowns and unknowns of boredom: a review of the literature. *Exp. Brain Res.* 236, 2451–2462. doi: 10.1007/s00221-017-4922-7
- Robinson, D. (2015). *Geom Flat Violin: Github Repository*. Available online at: <https://gist.github.com/dgrtwo/eb7750e74997891d7c20> (accessed May 13, 2021).
- Samuels, S. J., Hiebert, E. H., and Rasinski, T. V. (2010). “Eye movements make reading possible,” in *Revisiting Silent Reading: New Directions for Teachers and Researchers*, eds E. H. Hiebert and D. R. Reutzel (Newark, DE: International Reading Association), 24–44. doi: 10.1598/0833.02
- Scherer, K. R. (2003). Vocal communication of emotion: a review of research paradigms. *Speech Commun.* 40, 227–256. doi: 10.1016/s0167-6393(02)00084-5
- Scherer, K. R., London, H., and Wolf, J. J. (1973). The voice of confidence: paralinguistic cues and audience evaluation. *J. Res. Pers.* 7, 31–44. doi: 10.1016/0092-6566(73)90030-5
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423.
- Shi, X. (2020). As if one witnessed the creation: rethinking the aesthetic appreciation of Chinese calligraphy. *Philos. East West* 70, 485–505. doi: 10.1353/pew.2020.0031
- Simmer, M. L., and Goffin, R. D. (2003). A position statement by the International Graphonomics Society on the use of graphology in personnel selection testing. *Int. J. Test.* 3, 353–364. doi: 10.1207/s15327574ijt0304_4
- Smith, R. H. (2000). “Assimilative and contrastive emotional reactions to upward and downward social comparisons,” in *Handbook of Social Comparison: Theory and Research*, eds J. Suls and L. Wheeler (New York, NY: Plenum), 173–200. doi: 10.1007/978-1-4615-4237-7_10
- Thrash, T. M., and Elliot, A. J. (2003). Inspiration as a psychological construct. *J. Pers. Soc. Psychol.* 84, 871–889. doi: 10.1037/0022-3514.84.4.871
- Tinio, P. P. (2013). From artistic creation to aesthetic reception: the mirror model of art. *Psychol. Aesthet. Creat. Arts* 7, 265–275. doi: 10.1037/a0030872
- Tschacher, W., Greenwood, S., Kirchberg, V., Wintzerth, S., van den Berg, K., and Tröndle, M. (2012). Physiological correlates of aesthetic perception of artworks in a museum. *Psychol. Aesthet. Creat. Arts* 6, 96–103. doi: 10.1037/a0023845
- van Gemmert, A. W., and Teulings, H. L. (2004). Connecting sciences using graphonomic research. *Motor Control* 8, 367–370. doi: 10.1123/mcj.8.4.367
- van Paasschen, J., Bacci, F., and Melcher, D. P. (2015). The influence of art expertise and training on emotion and preference ratings for representational and abstract artworks. *PLoS One* 10:e0134241. doi: 10.1371/journal.pone.0134241
- Wiley, R. H. (1983). The evolution of communication: information and manipulation. *Anim. Behav.* 2, 156–189.
- Wilkinson, L. (1999). Statistical methods in psychology journals: guidelines and explanations. *Am. Psychol.* 54, 594–604. doi: 10.1037/0003-066x.54.8.594

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

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

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