Spatiotemporal Signatures of Sense of Agency in Neural Representations

Authors: Ke Zhao, Jingjin Gu, Huang Guanhua, Zheng Shuang, Fu Xiaolan, Zhao Ke

Date: 2021-05-16T00:00:00+00:00

Abstract

Sense of agency refers to the subjective experience of controlling one' s own actions and thereby controlling the external environment that arises during voluntary action. The core elements constituting the sense of agency for actions are subjective intention and outcome feedback. This study aims to manipulate different attributes of these two core elements, utilizing magnetoencephalography (MEG) and other techniques, to investigate the feedforward and feedback mechanisms of sense of agency within the frontoparietal-dominant brain network and its spatiotemporal-specific markers, and to construct a novel cognitive neuroscience theoretical model. This will facilitate understanding of the generation and aftereffects of human action, and provide more objective reference criteria for the clinical diagnosis of related psychiatric disorders.

Full Text

Spatiotemporal Markers of the Sense of Agency in the Human Brain

ZHAO Ke[1,2], **GU Jingjin**[1,2], **HUANG Guanhua**[1,2], **ZHENG Shuang**[1,2], **FU Xiaolan**[1,2]

[1] State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China
[2] Department of Psychology, University of Chinese Academy of Sciences, Beijing 100049, China

Abstract: The sense of agency (SoA) refers to the subjective experience of controlling one' s own actions and, through them, events in the external world. Intention of action and outcome feedback constitute the two core components of SoA. This research aims to manipulate different attributes of these core components using magnetoencephalography (MEG) and other techniques to investigate the feedforward and feedback mechanisms and spatiotemporal specificity of SoA in a fronto-parietal brain network. We will construct a new cognitive-neural theoretical model that will enhance understanding of human action generation and its aftereffects, and provide more objective reference standards for clinical diagnosis of related psychiatric disorders.

Keywords: sense of agency, action intention, outcome feedback, temporal binding

Classification Code: B842

1. Problem Statement

Humans interact with and modify their environment through intentional actions that anticipate outcomes (Elsner & Hommel, 2004; Reznik & Mukamel, 2019; Ziessler et al., 2004). The sense of agency (SoA) represents the subjective experience of controlling one' s own actions and, by extension, external events (Haggard, 2017; Reznik & Mukamel, 2019; Zhang et al., 2018; Wu et al., 2019). As a crucial concept in self-consciousness, the feeling that "I" am the cause of an outcome constitutes an important conscious experience that directly influences how individuals define their own behavior (Gu et al., 2020). SoA also serves as an important indicator of mental health, with its abnormalities associated with numerous psychiatric disorders including schizophrenia, depression, autism spectrum disorder, and obsessive-compulsive disorder (Frith et al., 2000; Lindner et al., 2005; Msetfi et al., 2012; Oren et al., 2019; Sperduti et al., 2014; Voss et al., 2010).

Over the past two decades, SoA has gradually attracted widespread attention from researchers (Kalogeras et al., 2002; Li et al., 2019; Schwarz et al., 2018; Stephenson et al., 2018; Figure 1), yet the cognitive and neural mechanisms underlying its generation remain frontier scientific questions in this field. This project will focus on the cognitive-neural mechanisms of SoA, distilling the two core components of action intention and outcome feedback. Based on the comparator model framework and existing evidence regarding cognitive-neural mechanisms, we propose feedforward and feedback modes of operation, and will employ high spatiotemporal resolution techniques such as MEG to explore specific markers of SoA. Elucidating these mechanisms will facilitate understanding of how individual actions are generated and produce aftereffects, while also providing potential objective criteria for SoA applications in clinical, moral, and judicial contexts.

Figure 1 [Figure 1: see original paper]. Number of publications (Figure 1A) and citations (Figure 1B) related to sense of agency (SoA) in psychology and cognitive neuroscience from 2000-2019. Data from Web of Science citation analysis report. Search date: April 12, 2020. Keywords: "sense of agency."

2. Research Status

SoA can be measured through two approaches: subjective reports of control (explicit measurement) and perceptual differences (implicit measurement). Explicit measurement typically asks participants questions such as "Did you cause this outcome?" or "How much control did you have over the action outcome?" Since voluntary actions are usually accompanied by two perceptual effects—action-outcome temporal compression and sensory attenuation—these can serve as implicit indicators of SoA. Kalogeras et al. (2002) found that when participants' voluntary actions triggered auditory outcomes, they reported the action occurrence later and the sound earlier compared to control conditions, demonstrating subjective compression of the perceived interval between action and outcome. The action-outcome temporal compression effect correlates with explicitly reported SoA (Imaizumi & Tanno, 2019) and is considered the primary implicit measure of SoA (Buehner, 2012; Ebert & Wegner, 2010; Zhao et al., 2013). Sensory attenuation refers to the phenomenon where individuals perceive the intensity of self-generated outcomes as reduced (Blakemore, Frith & Wolpert, 1999). A classic example is that "being tickled by others feels ticklish, but self-tickling does not" (Blakemore, Wolpert & Frith, 1999; Schafer & Marcus, 1973). Research has shown that voluntary actions reduce the N100 component amplitude and decrease activation in sensory cortices (Blakemore, Wolpert & Frith, 1999), making sensory attenuation another implicit measure of SoA (Burin et al., 2017; Dewey & Knoblich, 2014).

Explicit measurement is straightforward but susceptible to cognitive biases in subjective reporting; for instance, participants tend to exaggerate their SoA (Haggard, 2017). Implicit measurement is relatively covert and objective but may dissociate from explicit measures (Dewey & Knoblich, 2014; Saito et al., 2015; Schwarz et al., 2019). Consequently, researchers have proposed combining both implicit and explicit indicators (Pfister et al., 2021). Among the two implicit measures, temporal compression focuses on the temporal offset between action and outcome, emphasizing the cognitive process linking them, whereas sensory attenuation focuses on perceptual intensity of the outcome, emphasizing cognitive processing of outcome feedback.

2.1. Components of Action-Related Sense of Agency

Action intention in SoA. Subjective intention refers to the conscious experience of wanting to perform an action (Vinding et al., 2013). Whether and when we act reflects the role of intention. Under conditions involving intention, individuals "bind" actions and outcomes along the temporal dimension, producing action-outcome temporal compression (Kalogeras et al., 2002; Zhao et al., 2016). Absence, reduction, or termination of intention all alter SoA. Passive key-pressing or key-releasing actions without subjective intention fail to produce temporal compression (Zhao et al., 2016). Hypnotically induced actions reduce temporal compression due to diminished intention (Lush et al., 2017). Coerced actions with reduced intention produce smaller temporal compression effects compared to non-coerced conditions (Caspar et al., 2016). Intentional termination also reduces temporal compression, such as when generated intentions are subsequently inhibited (Haggard et al., 2009).

Outcome feedback in SoA. Action outcome valence, action-outcome causal association, and outcome presentation modality all influence SoA. Yoshie and Haggard (2013, 2017) used emotionally valenced auditory stimuli (positive, neutral, negative) as outcome feedback and found that negative outcomes produced smaller temporal compression effects. Barlas et al. (2018) similarly found weaker subjective SoA when action outcomes were negative sounds. When two sounds (T1 and T2) were presented sequentially, with T2 causally linked to the action, stronger temporal compression occurred (Buehner & Humphreys, 2009). Outcome modality also affects SoA, though few studies have examined unimodal versus multimodal outcomes with inconsistent conclusions. Kawabe et al. (2013) and van Kemenade et al. (2017) found lower SoA with multimodal versus unimodal outcomes, whereas Thanopoulos et al. (2018) found that consistency across multimodal outcomes could enhance SoA.

2.2. Theoretical Models of Action-Related Sense of Agency

The most influential theoretical framework is the comparator model (Frith et al., 2000). This model proposes that SoA arises when perceived outcomes match predicted outcomes generated during sensorimotor processing. This conceptualization of the action system as an internal comparator process traces back to Helmholtz (1866). The comparator model of action control helps explain SoA generation (Kawato, 1999; Synofzik et al., 2008; Wolpert et al., 1995). Before action execution, the efference copy of the intention predicts the action outcome. When predicted sensory feedback is compared with actual outcome feedback, a predictive error (PE) is generated. If an event is self-caused (the sensory prediction is correct), actual feedback perfectly matches prediction and PE equals zero; otherwise, PE is non-zero. According to the comparator model, SoA emerges when predicted outcomes match actual sensory feedback (Haggard, 2017; Moore & Obhi, 2012). Any mismatch in this comparison reduces or eliminates SoA. The comparator model successfully explains phenomena of "lost agency," such as when operation failure generates predictive error and produces a feeling of "I didn' t do that; something went wrong."

2.3.1. Spatial Characteristics of Brain Processing for Sense of Agency

Direct research on the cognitive-neural mechanisms of SoA remains limited. Existing studies indicate that the supplementary motor area (SMA), pre-supplementary motor area (pre-SMA), parietal cortex, insula, and dorsolateral prefrontal cortex (DLPFC) play important roles in SoA generation (Crivelli & Balconi, 2017). We therefore summarize SoA brain mechanisms from the perspectives of intention and outcome feedback.

First, subjective intention may be closely related to SMA and pre-SMA activity. An early study found that when participants planned and imagined executing hand movements, only the SMA showed significant activation among 254 preselected regions (Roland et al., 1980). Compared to passive movements, monkeys showed significant SMA and pre-SMA activation during active movements (Halsband et al., 1994; Picard & Strick, 2003). After lesioning this region, monkeys exhibited reduced active movements and action initiation deficits (Thaler et al., 1995), suggesting SMA/pre-SMA involvement in encoding action commands preceding active movements (Goldberg, 1985; Orgogozo & Larsen, 1979). Fried et al. (1991) stimulated different parts of epileptic patients' SMA with weak currents, finding that pre-SMA stimulation generated the urge to move, while SMA stimulation produced involuntary simple movements. Both cathodal and anodal tDCS over pre-SMA reduced temporal compression effects (Cavazzana et al., 2015). Moore et al. (2010) used theta-frequency stimulation to disrupt local brain function and found that pre-SMA dysfunction reduced temporal compression between active key presses and subsequent tactile stimulation, whereas stimulation of sensorimotor regions, especially the hand area, did not produce this effect. Clinical studies show that basal syndrome patients with pre-SMA damage exhibit greater action-outcome compression, with more severe symptoms showing stronger compression (Wolpe et al., 2014). Another region related to subjective intention is DLPFC, which participates primarily in intention selection. Research shows DLPFC, like pre-SMA, is involved in planning and executing active movements (Teuchies et al., 2016). Frith (2000) proposed that prefrontal regions including DLPFC are closely related to response selection processes. By constructing a "response space" that includes selecting correct responses while inhibiting alternatives, this region participates in accurate selection of non-automatic responses.

Second, outcome feedback may be associated with posterior parietal cortex (PPC) and insula activity. Farrer and Frith (2002) required participants to navigate a circle through a T-shaped path, with prior instruction that circle movement was controlled either by themselves or another person. When participants believed they controlled the movement, the insula showed significant activation; when they believed movement was unrelated to their actions, the inferior parietal lobe activated. Another study manipulated participants' control over a virtual hand on screen, finding that reported SoA levels correlated negatively with right inferior parietal lobule (especially angular gyrus) activation and positively with insula activation (Farrer et al., 2003). The insula also serves as an important region for multimodal sensory integration and monitors the mismatch between internal action expectations and external outcome feedback (Farrer et al., 2003; Farrer & Frith, 2002; Fink et al., 1999).

The angular gyrus of the inferior parietal lobule also processes outcome feedback. Blanke et al. (2002) found that electrical stimulation of epileptic patients' right angular gyrus produced an experience of actions not belonging to oneself. Other studies found that during explicit SoA judgments, the angular gyrus activated when visual feedback was attributed to non-self actions (Farrer & Frith, 2002). TMS suppression of the angular gyrus reduced reported SoA (Chambon et al., 2015), and another study found stronger angular gyrus activation when SoA ratings were low (Chambon et al., 2013). Meta-analyses similarly identify the temporo-parietal junction (including angular gyrus) as fundamental to SoA (Sperduti et al., 2011). Schizophrenia patients show stronger SoA and greater temporal compression, with fMRI revealing abnormal functional connectivity between prefrontal cortex and angular gyrus when cues and actions mismatch (Voss et al., 2017). In a study of apraxic patients with parietal cortex lesions, participants performed key presses while viewing either their own hand or another person' s hand on screen, reporting each trial whether the observed hand belonged to them. When the experimenter' s gestures matched their own, patients' recognition accuracy decreased significantly, suggesting that parietal damage impairs effective evaluation of internal and external action feedback (Sirigu et al., 1999). MacDonald and Paus (2003) used TMS to stimulate participants' posterior parietal cortex, finding that inhibition of left posterior parietal cortex significantly reduced temporal compression in active action conditions.

2.3.2. Temporal Characteristics of Brain Processing for Sense of Agency

Fewer studies have used electroencephalography (EEG) or magnetoencephalography (MEG) to decode temporal features of SoA processing. Within the action-outcome window, components likely involved in action monitoring include correct-related negativity (CRN) and stimulus-preceding negativity (SPN). CRN occurs approximately 100 ms after action execution, reflecting subjective confidence in actions (Boldt & Yeung, 2015; Scheffers & Coles, 2000) or objective outcome uncertainty and task difficulty (Endrass et al., 2012; Pailing & Segalowitz, 2004). Sidarus et al. (2017) found that CRN amplitude correlated negatively with SoA, with lower control ratings corresponding to larger amplitudes.

SPN is a slow cortical wave occurring after action execution but before outcome feedback (Böcker et al., 1994; Damen & Brunia, 1987; Seidel et al., 2015; Simons et al., 1979). SPN typically appears before meaningful outcome feedback (Chwilla & Brunia, 1991), especially before high-arousal outcomes such as monetary rewards, erotic images, or electric shocks (Brunia et al., 2010). During initial learning, SPN amplitude is large due to the importance of outcome feedback. As tasks are mastered and outcome feedback becomes less critical, SPN amplitude decreases (Hirao et al., 2017; Morís et al., 2013; Ren et al., 2017). Moreover, high reward and complex feedback conditions elicit larger SPN amplitudes than no-reward or simple feedback conditions (Kotani et al., 2003). Kotani et al. (2001) investigated SPN changes under positive emotion (reward) and negative emotion (noise) conditions, finding larger SPN amplitudes when rewards or noise served as action outcomes compared to control conditions. Additionally, Böcker et al. (2001) recorded SPN preceding electric shocks induced by fear, demonstrating that SPN reflects anticipation of emotional stimuli. SoA typically emerges when participants can freely choose actions, with greater choice range producing higher SoA levels (Barlas et al., 2017). Correspondingly, SPN amplitude is larger when participants believe they can choose actions compared to non-choice conditions (Masaki et al., 2010). Zhao et al. (2014) found that during the blank interval of time perception tasks, active key pressing produced larger anterior-mid scalp P1 components than active key releasing, corresponding to delta-theta frequency components that are thought to be inversely related to temporal compression; both conditions produced identical P2 components.

For the outcome time window, major EEG/MEG components include N100/M100 and feedback-related negativity (FRN). The auditory N100 component appears approximately 100 ms after stimulus presentation, typically generated in primary sensory cortex (Godey et al., 2001; Näätänen & Michie, 1979; Zouridakis et al., 1998) or frontal cortex (Näätänen & Picton, 1987). Compared to externally generated sounds, self-generated sounds produce reduced N100 amplitudes (Ford et al., 2001; Horváth, 2015; Klaffehn et al., 2019). When auditory feedback mismatches expected outcomes, N100 amplitude increases (Behroozmand et al., 2009; Behroozmand & Larson, 2011). When participants pressed keys to trigger sounds or listened to sound sequences at identical intervals, auditory cortex N100 amplitudes were significantly lower in the self-triggered condition (Martikainen et al., 2005). FRN occurs approximately 250-300 ms after outcome feedback (Dehaene et al., 1994; Miltner et al., 1997), with larger amplitudes when outcomes are unexpected. Sidarus et al. (2017) also found that FRN amplitude correlated negatively with SoA ratings.

In summary, within the action-outcome window, action monitoring involves CRN and SPN components. CRN occurs around 100 ms post-action and correlates negatively with SoA, while SPN appears before meaningful outcomes and shows stronger amplitude with greater SoA. In the outcome window, research has primarily identified N100/M100 and FRN components. However, whether these components have one-to-one correspondence with intention and outcome feedback requires further investigation.

3. Research Proposal

In summary, existing research on the cognitive-neural mechanisms of SoA remains limited and has several shortcomings. Investigations of SoA influencing factors have been fragmented, lacking systematic research perspectives on subjective intention and outcome feedback. Existing explanatory models emphasize cognitive processes but lack corresponding cognitive-neural empirical support. Moreover, current research has focused primarily on spatial aspects, necessitating more comprehensive and in-depth investigation of spatiotemporal specificity markers in SoA generation.

To address these issues, this project will focus on the cognitive-neural mechanisms of action-related SoA by manipulating different attributes of subjective intention and outcome feedback to examine feedforward and feedback modes of operation. Using high spatiotemporal resolution MEG, we will investigate specific markers of SoA in a fronto-parietal brain network within both action-outcome intervals and outcome time windows.

The project comprises two main components: the feedforward influence of subjective intention on SoA and the feedback influence of outcome feedback on SoA. Intention serves as the compass for action, influencing pre-action expectations about actions and outcomes, and its effect on SoA is typically feedforward. By manipulating the presence and type of subjective intention, we will systematically investigate intention' s role in SoA and reveal the cognitive-neural mechanisms underlying its feedforward effects (Study 1). Outcome feedback occurs post-action and influences SoA generation and strength through feedback mechanisms. By manipulating outcome valence, action-outcome causal association, and outcome presentation modality, we will systematically examine how outcome feedback affects SoA and reveal the cognitive-neural mechanisms underlying its feedback effects (Study 2). Using SoA subjective ratings, temporal compression effects, and MEG time-domain and time-frequency components, we will identify specific markers of action-related SoA and distinguish the neural processing of feedforward intention versus feedback outcome mechanisms.

3.1. Study 1: Effects of Subjective Intention on Sense of Agency

Study 1 will use two experiments to investigate the feedforward mode of subjective intention and its cognitive-neural mechanisms. Behaviorally, we will examine differences in reported action-outcome intervals and subjective SoA ratings under conditions of intention presence (active vs. passive, Experiment 1) and intention type (emerging vs. disappearing intention, Experiment 2). With constant outcome feedback, we can isolate the feedforward modulatory effects of subjective intention on SoA. Using MEG, we will examine magnetic components in the action-outcome window (CRN, SPN, etc.) and outcome feedback window (M100, FRN, etc.), focusing on time-domain, time-frequency components, and functional connectivity in a fronto-parietal network including (pre-)SMA and DLPFC.

Experiment 1 manipulates intention presence to investigate its role in SoA and its neural manifestations in fronto-parietal networks during action-outcome and outcome windows. Experiment 2 manipulates intention type to examine how different intention types affect SoA and corresponding neural indicators in fronto-parietal networks across both time windows.

3.2. Study 2: Effects of Outcome Feedback on Sense of Agency

Study 2 will use three experiments to investigate the feedback mode of outcome feedback characteristics on SoA from the perspectives of outcome-action causal association, outcome presentation modality, and outcome valence. Behaviorally, we will examine differences in reported action-outcome intervals and subjective

SoA ratings under conditions of high versus low causal association (Experiment 3), unimodal versus multimodal outcomes (Experiment 4), and positive versus neutral versus negative valence (Experiment 5). Using the difference between intentional and non-intentional conditions as an implicit SoA index, we will examine feedback modulatory effects of outcome feedback changes. Using MEG, we will examine neural activity differences between intentional and non-intentional conditions as indicators of magnetic components in the outcome feedback window (M100, FRN) and action-outcome window (CRN, SPN) under different outcome feedback conditions. We will focus on time-domain, time-frequency components, and functional connectivity in a fronto-parietal network including the angular gyrus, insula, pre-SMA, and DLPFC.

Experiment 3 manipulates causal association strength to investigate its influence on SoA and behavioral and neural indicators in both time windows. Experiment 4 manipulates unimodal versus multimodal outcome features to examine how presentation modality affects SoA and its behavioral and neural manifestations. Experiment 5 manipulates outcome valence (positive/negative) to investigate its effects on SoA and corresponding behavioral and neural indicators.

4. Theoretical Construction and Innovation

The cognitive-neural mechanisms of SoA remain a critical unresolved scientific question in this field. This project distills the two core components of intention and outcome feedback and, based on the comparator model framework, attempts to investigate spatiotemporal specificity markers of SoA from feedforward-feedback perspectives. Integrating existing research, the fronto-parietal network likely plays a crucial role in SoA. The SMA, pre-SMA, and DLPFC may be associated with intention generation and selection, while the insula and angular gyrus monitor mismatches between expected and actual feedback. The prefrontal cortex generates action selection, intention, and initiation signals (Desmurget & Sirigu, 2009), whereas the parietal cortex monitors these signals and compares intention signals with sensory feedback (Haggard, 2017). Frontal midline regions operate feedforward—selecting and initiating voluntary actions (pre-action)—then transmit information to parietal cortex that monitors intentions, actions, and outcomes, thereby operating feedback—evaluating whether actions achieved expected outcomes (Figure 2 [Figure 2: see original paper]). Previous research has not clarified how feedforward and feedback mechanisms operate or how they are interconnected.

This project' s distinctive feature lies in exploring the feedforward mechanism of intention and feedback mechanism of outcomes, systematically investigating specific markers of SoA in a fronto-parietal network during action-outcome and outcome time windows. High spatiotemporal resolution MEG can noninvasively measure real-time changes in brain magnetic signals, making it suitable for measuring correlations between different brain regions and synchrony of population neuronal activity in cortical networks. We will comprehensively investigate signal transmission in the fronto-parietal network during SoA using time-domain analysis, time-frequency analysis, and Granger causality analysis. Furthermore, this field requires a more comprehensive cognitive-neural theory to explain SoA at levels of subjective intention, sensorimotor control and monitoring, and recognition. This project will refine the cognitive-neural theoretical model of SoA through a series of experimental studies.

Figure 2 [Figure 2: see original paper]. Schematic diagram of feedforward-feedback brain networks for action-related sense of agency (red indicates brain regions involved in feedforward processing, blue indicates regions involved in feedback processing).

References

Gu, J. J., Zhao, K., & Fu, X. L. (2020). Sense of agency in behavior and responsibility attribution. *Chinese Science Bulletin*, 65(19), 1902–1911. doi:10.1360/TB-2019-0715

Wu, D., Gu, J. J., Li, M., Zhang, M., Zhang, M., Zhao, K., & Fu, X. L. (2019). Sense of agency for actions and sense of agency for causality: The mechanism underlying temporal binding of voluntary actions. *Advances in Psychological Science*, 27(5), 804–810. doi:10.3724/SP.J.1042.2019.00804

Zhang, M., Wu, D., Li, M., Ling, Y. B., Zhang, M., & Zhao, K. (2018). Measurement and cognitive-neural mechanisms of sense of agency. *Advances in Psychological Science*, 26(10), 1787–1793. doi:10.3724/SP.J.1042.2018.01787

[The remaining references from the original text are preserved in their original format and order]

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv —Machine translation. Verify with original.