Psychology. Journal of the Higher School of Economics. 2026. Vol. 23. N 3. P. 512–527.
Психология. Журнал Высшей школы экономики. 2026. Т. 23. № 3. С. 512–527.
DOI: 10.17323/1813-8918-2026-3-512-527

CONTENT, CONSUMER, CORRECTIONS: A NEUROCOGNITIVE MODEL OF MISINFORMATION PROCESSING

E. MONAHHOVA[a], J.M. GORODNICHEVA[a], A.N. MOROZOVA[a], V.A. KLUCHAREV[a]

[a] *HSE University, 20 Myasnitskaya Str., Moscow, 101000, Russian Federation*

Контент, потребитель, коррекции: нейрокогнитивная модель обработки дезинформации

Э. Монахова[a], Ю.М. Городничева[a], А.Н. Морозова[a], В.А. Ключарев[a]

[a] *Национальный исследовательский университет «Высшая школа экономики», 101000, Россия, Москва, ул. Мясницкая, д. 20*

Abstract

The rapid proliferation of misinformation in digital environments poses a serious threat to public health, political institutions, and public discourse. Behavioral research has accumulated substantial evidence on why people believe and share false information, yet individual studies address cognitive biases, social factors, or emotional processes largely in isolation. Additionally, the influential integrative frameworks that do exist operate exclusively at the behavioral level, overlooking the neurophysiological mechanisms of misinformation processing. This work addresses that gap by proposing a neurocognitive model that synthesizes behavioral and neuroimaging findings within a unified framework. Following an

Резюме

Стремительное распространение дезинформации в цифровой среде представляет серьезную угрозу для общественного здоровья, политических институтов и публичного дискурса. Поведенческие исследования накопили значительный массив данных о том, почему люди доверяют ложной информации и распространяют ее, однако отдельные работы рассматривают когнитивные искажения, социальные факторы или эмоциональные процессы преимущественно изолированно. Кроме того, существующие влиятельные интегративные модели остаются на поведенческом уровне, не учитывая нейрофизиологические механизмы обработки дезинформации. Настоящая работа восполняет этот пробел, предлагая нейрокогнитивную модель, синтезирующую поведенческие и нейровизуализационные данные в рамках единой теоретической структуры. Следуя методологии

The study was conducted within the framework of the Russian Science Foundation (RSF), grant No. 24-18-00432, https://rscf.ru/project/24-18-00432/

Исследование проведено в рамках гранта Российского научного Фонда (РНФ) № 24-18-00432, https://rscf.ru/project/24-18-00432/

integrative review methodology, we organize the literature around three domains: message characteristics (Content: emotionality, novelty, repetition), individual consumer traits (Consumer: analytical thinking, need for cognition, working memory, conformity, emotional state, and prior beliefs), and corrective interventions (Corrections: reactive debunking and pre-emptive prebunking, inoculation, and warnings). For each domain, principal behavioral factors are mapped onto their fMRI and EEG/MEG/ERP correlates. Cross-domain analysis reveals that these behaviorally distinct phenomena rely on a shared neural infrastructure of four functional hubs, each performing a core computational operation: conflict detection (anterior cingulate cortex), belief valuation (ventromedial prefrontal cortex), memory encoding and persistence (hippocampal–amygdala circuit), and executive control (dorsolateral prefrontal cortex and inferior frontal gyrus). Framed as a mechanism sketch, the model generates testable neuroimaging predictions and points toward neuroscience-informed strategies for countering misinformation.

Keywords: misinformation, neurocognitive model, misinformation correction, neural correlates.

Eliana Monahhova — Junior Research Fellow, Centre for Cognition and Decision Making, Institute for Cognitive Neuroscience, Making, HSE University.
Research Area: cognitive neuroscience, disinformation, EEG.
E-mail: e.monakhova@hse.ru

Julia M. Gorodnicheva — Research Assistant, International Laboratory of Social Neurobiology, Institute for Cognitive Neuroscience, HSE University.
Research Area: cognitive neuroscience, disinformation, EEG.
E-mail: ymgorodnicheva@hse.ru

интегративного обзора, мы организуем литературу вокруг трёх доменов: характеристики сообщения (Контент: эмоциональность, новизна, повторение), индивидуальные характеристики потребителя (Потребитель: аналитическое мышление, потребность в познании, рабочая память, конформность, эмоциональное состояние и предшествующие убеждения) и корректирующие интервенции (Коррекция: реактивное опровержение, а также упреждающие пребанкинг, инокуляция и предупреждения). Для каждого домена основные поведенческие факторы сопоставляются с их коррелятами, выявленными методами фМРТ и ЭЭГ/МЭГ/ВП. Кросс-доменный анализ показывает, что эти поведенчески различные феномены опираются на общую нейронную инфраструктуру из четырех функциональных узлов, каждый из которых выполняет базовую вычислительную операцию: детекция конфликта (передняя поясная кора), оценка убеждений (вентромедиальная префронтальная кора), кодирование и устойчивость памяти (гиппокампально-амигдалярный контур) и исполнительный контроль (дорсолатеральная префронтальная кора и нижняя лобная извилина). Сформулированная как «эскиз механизма» (англ. mechanism sketch), модель порождает тестируемые нейровизуализационные предсказания и указывает на нейронаучно обоснованные стратегии противодействия дезинформации.

Ключевые слова: дезинформация, нейрокогнитивная модель, коррекция дезинформации, нейронные корреляты.

Монахова Элиана — младший научный сотрудник, Центр нейроэкономики и когнитивных исследований, Институт когнитивных нейронаук, Национальный исследовательский университет «Высшая школа экономики».
Сфера научных интересов: когнитивная нейронаука, дезинформация, ЭЭГ.
Контакты: e.monakhova@hse.ru

Городничева Юлия Михайловна — стажер-исследователь, Международная лаборатория социальной нейробиологии, Институт когнитивных нейронаук, Национальный исследовательский университет «Высшая школа экономики».
Сфера научных интересов: когнитивная нейронаука, дезинформация, ЭЭГ.
Контакты: ymgorodnicheva@hse.ru

Alexandra N. Morozova — Research Assistant, International Laboratory of Social Neurobiology, Institute for Cognitive Neuroscience, HSE University. Research Area: cognitive neuroscience, disinformation, EEG.
E-mail: an.morozova@hse.ru

Vasily A. Klucharev — Head of Laboratory, International Laboratory of Social Neurobiology, Institute for Cognitive Neuroscience, HSE University. DSc in Cognitive Sciences, Professor.
Research Area: cognitive neuroscience, conformity, disinformation.
E-mail: vklucharev@hse.ru

Acknowledgments
The authors would like to thank Oksana Zinchenko, Senior Research Fellow at the Centre for Cognition and Decision Making, HSE University, for her valuable input in shaping the conceptual framework of this study.

Морозова Александра Николаевна — стажер-исследователь, Международная лаборатория социальной нейробиологии, Институт когнитивных нейронаук, Национальный исследовательский университет «Высшая школа экономики».
Сфера научных интересов: когнитивная нейронаука, дезинформация, ЭЭГ.
Контакты: an.morozova@hse.ru

Ключарев Василий Андреевич — заведующий лабораторией, Международная лаборатория социальной нейробиологии, Институт когнитивных нейронаук, ,Национальный исследовательский университет «Высшая школа экономики», доктор когнитивных наук, профессор.
Сфера научных интересов: когнитивная нейронаука, конформизм, дезинформация.
Контакты: vklucharev@hse.ru

Благодарности
Авторы выражают благодарность Оксане Зинченко, старшему научному сотруднику Центра нейроэкономики и когнитивных исследований НИУ ВШЭ, за ценный вклад в разработку концептуальной структуры исследования.

Supplementary Materials are available at: https://psy-journal.hse.ru/article/view/42723/33378

Introduction

The rapid proliferation of misinformation in digital media environments poses a serious threat to public health, political institutions, and public discourse (Lazer et al., 2018; Pennycook & Rand, 2021). Behavioral research on misinformation susceptibility has accumulated substantial evidence, yet individual studies focus on cognitive biases, social factors, or emotional processes, largely in isolation. Influential frameworks have attempted to integrate these strands: for example, false belief formation has been modelled in terms of its psychological drivers (Ecker et al., 2022), susceptibility has been reframed as a failure of analytical reasoning (Pennycook & Rand, 2021), and principles of effective correction have been formulated (Lewandowsky et al., 2020). However, this exclusively behavioural scope leaves a critical gap. A growing body of neuroimaging studies has examined individual components of misinformation processing (van der Linden & Cohen, 2025), yet these findings remain disconnected from one another and from a unified neurocognitive framework.

To bridge this gap, we reviewed the relevant literature and outlined a pathway toward a neurocognitive model that synthesizes behavioral and neurophysiological findings on misinformation processing. To our knowledge, this study represents the first attempt to systematically integrate neuroscientific evidence on misinformation processing within a unified theoretical framework, with the explicit goal to identify

cross-domain interactions and their potential causal pathways. The literature review and the model are organized around three domains: message characteristics (Content), consumer characteristics (Consumer), and correction mechanisms (Corrections) — reflecting complementary determinants of misinformation processing: properties of the message, characteristics of the recipient, and interventions targeting belief updating. For each domain, we identify principal behavioral factors and map them onto their neurophysiological correlates identified in functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) studies. This approach allows us to consolidate disparate findings and identify cross-domain neural convergences.

In this mini-review, "misinformation" serves as an umbrella term denoting information that diverges from the actual state of affairs (adapted from Ecker et al., 2022). Given the nascent state of neuroscientific research in this domain, this inclusive operationalization facilitates a more comprehensive synthesis of the available empirical evidence. Importantly, we employ the term "model" to denote a process-level mechanistic framework rather than a formal computational model, representing a distinction in the degree of formalization rather than theoretical validity. The proposed construct constitutes a mechanism sketch (Piccinini & Craver, 2011): a partial mechanistic explanation specifying functional components (four cross-domain neurocognitive hubs), their respective operations (conflict detection, valuation, memory encoding/persistence, and executive control), and their organizational configurations (see the model structure in Figure 2). Crucially, algorithmic-level details are left open for subsequent specification (Craver, 2007).

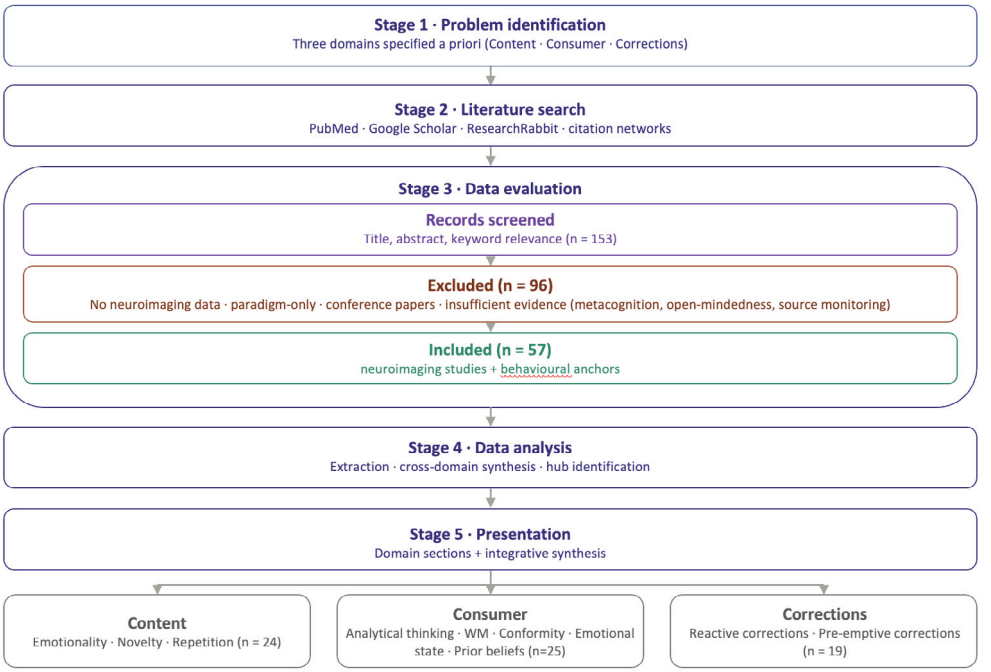
Methodology

The present review follows an integrative review methodology (Whittemore & Knafl, 2005). It provides a structured five-stage procedure — problem identification, literature search, data evaluation, data analysis, and presentation — suited to synthesizing findings across methodologically diverse research traditions (see Figure 1). This approach was chosen because our objective required integrating behavioral and neuroimaging evidence within a single theoretical framework (for a similar approach, see Scholz et al., 2020). A literature search was conducted in PubMed, Google Scholar, and ResearchRabbit. The search strategy covered three domains — content characteristics, consumer characteristics, and correction of misinformation — using "misinformation" and related terms (disinformation, fake news, false information) combined with domain-specific and neuroimaging keywords, e.g., fMRI, EEG, and ERP (see Table 1 in the Supplementary Materials).

We employed specific strategies to search the literature across three key domains. First, studies on message characteristics (Content) focused on emotionality, novelty, and repetition. Second, research into individual characteristics (Consumer) addressed analytical thinking, need for cognition, working memory, conformity, emotional state, and attitudes. Finally, studies investigating correction mechanisms (Corrections) distinguished between reactive (debunking), applied after misinformation exposure, and pre-emptive (prebunking) interventions, applied before it (Ecker et al., 2022). The final corpus is centered on neuroscience

Figure 1

A Flow Diagram of the Literature Search and Inclusion Procedure across Five Methodological Stages



evidence, complemented by key behavioral studies and major meta-analytic reviews where these provided essential theoretical or empirical grounding. Studies were identified through database searches, reference lists, and citation networks and screened iteratively for relevance. The final corpus comprised 57 unique studies (for details, see Table 1 in the Supplementary Materials).

Model Domains

Content: Message Characteristics Affecting Misinformation Processing

In contemporary information ecosystems, susceptibility to misinformation is increasingly understood in terms of how message features modulate attention, valuation, memory, and metacognitive control. Through these cognitive mechanisms, the content of misinformation shapes belief, confidence, information seeking, and sharing. The Content domain is organized around three message-level features most consistently identified in the literature: emotionality, novelty, and repetition (van der Linden & Cohen, 2025). Additional message-level characteristics, such as presentation format and endorsement by an expert (e.g., Blinova & Shcherbakova, 2022; Turcotte et al., 2015), have been investigated in behavioral research but are

not included in the present model due to limited neuroimaging evidence in the misinformation context.

Table 1 summarizes the behavioral and neurocognitive mechanisms underlying each content factor. Emotionality increases attentional salience and favors heuristic processing, with misinformation often exhibiting high arousal and negative skew that captures attention and drives sharing (Beauvais, 2022; Carrasco-Farré, 2022; Damstra et al., 2021; Paschen, 2020). Novelty, in turn, acts as a prediction error signal that prioritizes false news for processing and transmission independently of veracity (Photiou et al., 2021; Vosoughi et al., 2018; Nevo & Horne, 2022). Repetition drives susceptibility via the illusory truth pathway, wherein processing fluency is misattributed to veracity (Beauvais, 2022; Wang et al., 2016), although it can paradoxically increase perceived "fakeness" when the context prompts fake-news categorization (Corneille et al., 2020).

Consumer: individual characteristics affecting misinformation processing

The Consumer domain encompasses individual-level characteristics that modulate how people evaluate, accept, or resist misinformation [i.e., analytical thinking,

Table 1

Neurocognitive Mechanisms of Content Characteristics: Behavioral and Neural Pathways Underlying Emotionality, Novelty, and Repetition Effects in Misinformation Processing

Factor	Behavioral Mechanism	Neural Substrates	Key References
Emotionality (valence/arousal) and emotional framing	Emotional salience increases heuristic processing; emotional framing guides interpretation and belief; negative skew captures attention and drives sharing; headline reframing alters engagement	Affective valuation: amygdala activation with reduced PFC control (frame-consistent choices); EEG: persistent LPP amplitudes post-credibility assessment (emotional influence on late evaluative processing)	Baum & Abdel Rahman, 2021; De Martino et al., 2006; Jeng et al., 2024; Martel et al., 2020; Nabi, 2003; Song et al., 2023; van der Linden & Cohen, 2025
Novelty	Novel stimuli act as prediction errors; drives transmission independently of veracity; curiosity promotes engagement and memorability; topic novelty moderates veracity guidance	Novelty encoding: SN/VTA dopaminergic signaling (novelty detection), hippocampus (memory encoding); Reward: nucleus accumbens (curiosity-driven engagement)	Bunzeck & Düzel, 2006; Gruber et al., 2014; Song et al., 2023; van der Linden & Cohen, 2025; Wittmann et al., 2007
Repetition	Processing fluency misattributed to veracity (illusory truth effect); reduces moral condemnation of sharing; can paradoxically increase perceived "fakeness" in fake-news contexts	Familiarity processing: PRC tracks perceived truth (repetition-dependent); familiarity-driven judgment rather than recollection-based verification	Orchinik et al., 2025; van der Linden & Cohen, 2025; Wang et al., 2016

Note. PFC = prefrontal cortex; LPP = late positive potential; SN/VTA = substantia nigra/ventral tegmental area; PRC = perirhinal cortex.

the need for cognition, working memory, conformity, emotional state, and internal attitudes (beliefs)].

Analytical thinking and NFC modulate the balance between heuristic (System 1) and deliberate (System 2) processing, with higher CRT scores predicting better discernment (Pennycook & Rand, 2021; Bago et al., 2020). Working memory capacity constrains the ability to detect inconsistencies and integrate corrections, with less efficient updating linked to the continued influence effect (Brydges et al., 2018; McIlhiney et al., 2023). Social conformity modulates acceptance through perceived group norms (Ecker et al., 2022), while emotional states shift processing toward heuristic reliance (Martel et al., 2020). Prior beliefs remain among the strongest predictors, with individuals more likely to accept belief-congruent claims and resist incongruent corrections (Ecker et al., 2022; Kaplan et al., 2016; see details about the behavioral and neurocognitive mechanisms underlying each consumer factor in Table 2).

Other individual-level factors, for example, age and educational level (e.g., Brashier & Schacter, 2020; Kalogeropoulos et al., 2019), are recognized behavioral predictors of misinformation susceptibility. However, the absence of neuroimaging evidence linking these characteristics to misinformation processing currently precludes their inclusion in the model. Nevertheless, we consider the investigation of such demographic factors a critical direction for future neurophysiological research.

Corrections: Correction Mechanisms

The third domain of the model, corrective interventions, functions not as deletion mechanisms but as regulatory operations that modulate conflict detection, memory competition, valuation of belief change, and source-based encoding, thereby shaping belief revision and resistance to future exposure (Ecker et al., 2022; Lewandowsky et al., 2020).

Table 3 describes the neurocognitive mechanisms of reactive and pre-emptive correction strategies. Reactive corrections (debunking) involve post-exposure interventions constrained by the continued influence effect, explained by two accounts: mental-model-updating, in which retractions create coherence gaps, and selective-retrieval, in which misinformation and corrections compete during retrieval (Lewandowsky et al., 2012; Ecker et al., 2022). Pre-emptive corrections (prebunking, inoculation, warnings) operate at encoding, grounded in inoculation theory: prior exposure to weakened misleading content builds resistance to future, more potent versions (Compton, 2013; Cook et al., 2017; Lewandowsky & van der Linden, 2021).

Cross-Domain Neural Convergences

The neuroimaging evidence reviewed above reveals that the three domains — Content, Consumer, Corrections — are not neurally independent. Across behaviorally distinct phenomena, the same brain structures recur with striking regularity.

Table 2

Neurocognitive Mechanisms of Consumer Characteristics: Individual-Level Factors Modulating Misinformation Susceptibility

Factor	Behavioral Mechanism	Neural Substrates	Key References
Analytical thinking and need for cognition (NFC)	CRT scores predict fake news discernment; NFC predicts habitual System 2 engagement; confirmation bias may override analytical processing	Default processing: DMN (System 1 heuristic mode); Conflict detection: dACC, left AI, IFG (inhibitory control in dual-process tasks); Effortful cognition: frontal cortex engagement (NFC); Valuation: vmPFC activation (NFC-matched messages)	Aquino et al., 2020; Gronchi & Giovannelli, 2018; Gronchi et al., 2024; Kahneman, 2011; Moravec et al., 2019; Mussel et al., 2016
Working memory	WM capacity constrains inconsistency detection and correction integration; less efficient WM updating associated with CIE	Encoding: disrupted PrC/PCC (coherence-building); Concurrent storage: left AG, bilateral PrC (correct and incorrect information); Executive control: left IFG modulates temporal regions; Retrieval: ACC recruited during corrected information retrieval	Gordon et al., 2017, 2019; Jia et al., 2024
Conformity	Social consensus increases false belief; judgments adjusted toward group opinion; politically motivated reasoning engages distinct neural networks	Conflict signaling: ventral striatum deactivation, pMFC/AI (group-deviation detection); Memory distortion: amygdala-hippocampal modulation (false memory formation); Motivated reasoning: vmPFC, ventral striatum (value encoding), dmPFC/AI (error detection), bilateral TPJ (mentalizing)	Edelson et al., 2011; Lois et al., 2024; Wu et al., 2016
Emotional state	Negative affect reduces analytical processing; emotional valence overrides source trustworthiness; motivated reasoning activates emotional regulation circuits	Emotional regulation: vmPFC, ACC, PCC (activated over lateral PFC reasoning networks); Affect-biased belief: AI, amygdala, VTA/SN (belief formation); Belief updating: AI connectivity tracks emotional valence	Baum & Abdel Rahman, 2021; Müller-Pinzler et al., 2022; Westen et al., 2006
Attitudes (prior beliefs)	Belief-congruent claims accepted; incongruent corrections rejected; repeated misinformation reshapes beliefs; confirmation bias has a neural substrate	Belief defense: DMN activation (political challenges), amygdala/AI (resistance to change); Bias-driven processing: vmPFC, dmPFC (high conspiracy believers), hippocampus/PrC (low conspiracy believers); Confirmation bias: pMFC tracks confirming over disconfirming opinions	Kappes et al., 2020; Kossowska et al., 2026; Zhao et al., 2025

Note. CRT = Cognitive Reflection Test; NFC = need for cognition; DMN = default mode network; dACC = dorsal anterior cingulate cortex; AI = anterior insula; IFG = inferior frontal gyrus; vmPFC = ventromedial prefrontal cortex; PrC = precuneus; PCC = posterior cingulate cortex; AG = angular gyrus; pMFC = posterior medial frontal cortex; TPJ = temporoparietal junction; VTA/SN = ventral tegmental area/substantia nigra.

These cross-domain convergences indicate that misinformation processing relies on a set of domain-general neurophysiological hubs, each performing a core computational operation: conflict detection, valuation, memory encoding, or executive

Table 3

Neurocognitive Mechanisms of Correction Strategies: Neural Substrates of Reactive (Debunking) and Pre-Emptive (Prebunking, Inoculation, Warnings) Interventions

Correction Type	Behavioral Mechanism	Neural Substrates	Key References
Reactive corrections (Debunking)	Post-exposure identification of misinformation + alternative explanation; constrained by CIE; mental-model-updating vs. selective-retrieval accounts	Encoding: reduced PrC/PCC (coherence), reduced MTG/IFG (semantic integration); Retrieval: reduced MFG/ACC/dlPFC (control), enhanced AG/PrC (concurrent storage); EEG/ERP: frontal conflict signals, left-parietal recollection, P300 (recollection with alternative explanation), frontal N400 (familiarity persistence); Belief revision: alPFC tracks value of updating; amygdala-hippocampal coupling constrains revision	Brydges & Ecker, 2018; Brydges et al., 2020; Edelson et al., 2014; Gordon et al., 2017, 2019; Guo et al., 2025; Hua et al., 2022; Jia et al., 2024; Metcalfe et al., 2012
Pre-emptive corrections (Prebunking, Inoculation, Warnings)	Weakened misleading content builds resistance (inoculation theory); warnings increase source-tagging; modulates reconstructive memory; increases neural complexity	IFG, ACC, dlPFC (conflict detection, source monitoring); decreased auditory reactivation (misinformation), increased visual reactivation (original event); stronger visual reinstatement predicts reduced susceptibility; increased frontal/temporal neural complexity	Karanian et al., 2020, 2024; Molina et al., 2025

Note. CIE = continued influence effect; PrC = precuneus; PCC = posterior cingulate cortex; MTG = middle temporal gyrus; IFG = inferior frontal gyrus; MFG = middle frontal gyrus; ACC = anterior cingulate cortex; dlPFC = dorsolateral prefrontal cortex; AG = angular gyrus; alPFC = anterior-lateral prefrontal cortex.

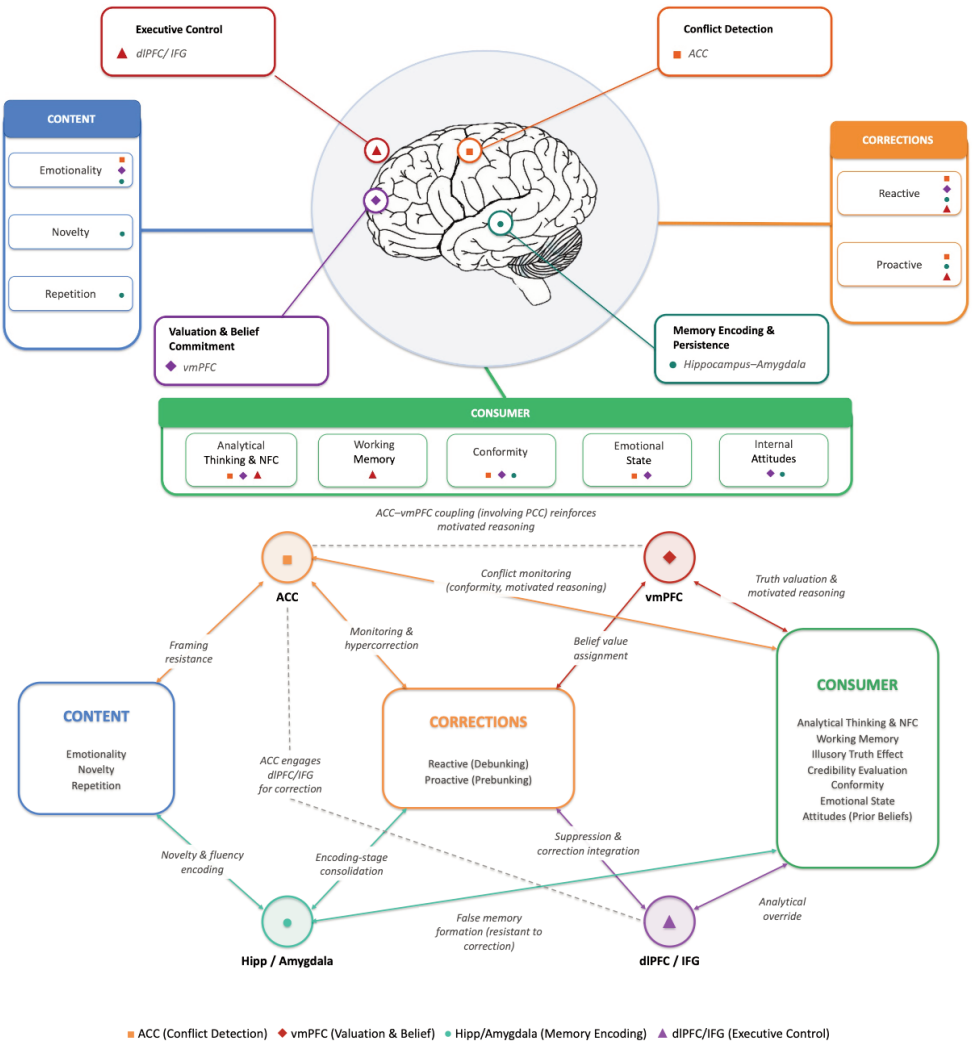
control (see Figure 2). These functional attributions are proposed as candidate computational roles consistent with predominantly correlational neuroimaging evidence. Accordingly, the directional relationships described below represent theoretically motivated interpretations rather than established causal mechanisms (Poldrack, 2010).

Conflict Detection: The Anterior Cingulate Cortex (ACC)

The ACC emerges as a single structure most consistently implicated across all three domains. In the Content domain, the ACC detects discrepancies between affective responses and evaluative standards during resistance to emotional framing (De Martino et al., 2006). In the Consumer domain, the dACC supports conflict monitoring across analytical engagement, opinion confirmation, social conformity, and motivated reasoning (Gronchi et al., 2024; Kappes et al., 2020; Wu et al., 2016; Westen et al., 2006). Finally, in the Corrections domain, reduced ACC activation indexes monitoring failure (Jia et al., 2024), while ACC recruitment

Figure 2

A Neurocognitive Model of Misinformation Processing



Note. (A) Structural overview: the Content, Consumer, and Corrections domains converge on four neurocognitive hubs – conflict detection (■, ACC), valuation (◆, vmPFC), memory encoding (●, hippocampus–amygdala), and executive control (▲, dlPFC/IFG). (B) Cross-domain convergences: bidirectional arrows indicate that domain characteristics engage hub computations and, reciprocally, hub activation patterns shape downstream behavioral outcomes (e.g., framing resistance, belief updating, correction efficacy). Dashed lines denote functional inter-hub connections.

drives hypercorrection (Metcalfe et al., 2012) and the tagging of discrepant information under prewarning (Karanian et al., 2024).

These manifestations converge on a single computational role: the ACC operates as a domain-general conflict monitor that flags mismatches between competing representations. The downstream behavioral effect depends on which circuits the ACC engages — lateral prefrontal regions (dlPFC/IFC) facilitate correction, whereas vmPFC and PCC coupling reinforces motivated reasoning. This dual role explains why the ACC is active both when people successfully revise false beliefs and when they resist belief change.

Valuation and Belief Commitment: The vmPFC

The vmPFC functions as the central valuation hub within the proposed model. In the Consumer domain, vmPFC activity tracks truth judgements, NFC-matched messaging, politically motivated reasoning and conspiracy belief (Aquino et al., 2020; Lois et al., 2024; Westen et al., 2006; Zhao et al., 2025). In the Corrections domain, the adjacent alPFC similarly tracks the subjective value of switching to a corrected representation (Edelson et al., 2014).

This convergence indicates that belief acceptance engages a valuation computation in which the brain assigns subjective worth to a candidate belief; revision requires the expected gain from updating to exceed the cost of abandoning the prior representation. This explains why corrections merely negating misinformation are less effective than those providing causal alternatives (Lewandowsky et al., 2020) and why affectively charged misinformation remains resistant to correction.

Memory Encoding and Persistence: The Hippocampus and Amygdala

The hippocampus and amygdala jointly function as the primary gateway through which misinformation enters long-term memory. In the Content domain, midbrain-hippocampus circuits encode novelty signals (Bunzeck & Düzel, 2006), while the PrC tracks repetition-driven fluency, contributing to familiarity-based truth judgements (Wang et al., 2016). In the Consumer domain, amygdala-hippocampus coupling during misleading social influence produces false memories resistant to corrections (Edelson et al., 2011), while amygdala activation predicts resistance to belief change (Kaplan et al., 2016). Similarly, in the Corrections domain, encoding-stage amygdala-hippocampus coupling predicts the subsequent failure of belief revision (Edelson et al., 2014).

This reveals a critical asymmetry: novelty, emotional arousal, social consensus, and fluency all converge on hippocampal-amygdala circuits to consolidate misinformation, whereas corrections necessitate effortful prefrontal reprocessing. This asymmetry underscores the importance of pre-emptive interventions that modulate encoding-stage dynamics before a false memory is consolidated.

Executive Control: The Lateral Prefrontal Network

The dlPFC and IFG constitute the executive control axis. In the Consumer domain, the IFG serves as a core component of analytical override (Gronchi et al.,

2024). In the Corrections domain, the right dlPFC suppresses outdated beliefs during hypercorrection (Metcalfe et al., 2012), whereas reduced dlPFC activity under prewarning reflects an attenuated processing strategy (Karanian et al., 2024). Finally, IFG-to-temporal connectivity modulates correction integration (Jia et al., 2024).

The effectiveness of these operations is capacity-limited and depends on working memory resources (Brydges et al., 2018; McIlhiney et al., 2023). This explains why misinformation encountered under cognitive load, time pressure, or emotional arousal — conditions that deplete prefrontal resources — is more readily accepted and more resistant to correction.

Conclusion

This mini-review systematizes the neurophysiological fingerprints of misinformation perception within a unified theoretical framework. Integrating evidence across the Content, Consumer, and Corrections domains, we identify four convergent neurocognitive processes — conflict detection (ACC), valuation and belief commitment (vmPFC), memory encoding and persistence (hippocampus–amygdala), and executive control (dlPFC/IFG) — that constitute the shared neural infrastructure underlying misinformation processing. Although this framework targets a universal neurocognitive model, the nascent state of the field precludes definitive conclusions regarding its applicability across all misinformation variants. Consequently, future research must evaluate whether these identified processes extend uniformly across different misinformation types, consumer profiles, and intervention contexts beyond our current scope. As required for a tangible model, it generates testable predictions. For instance, tDCS/TMS upregulation of the ACC prior to misinformation exposure should enhance fake-news detection accuracy. Similarly, prebunking should increase right inferior frontal gyrus (rIFG) and ACC activity during subsequent exposure — verifiable via EEG N400/P600 components, while ventromedial prefrontal cortex (vmPFC) sensitivity to belief-update value should predict debunking efficacy. Future research should test these predictions through targeted neuroimaging experiments.

References

Aquino, A., Alparone, F. R., Pagliaro, S., Haddock, G., Maio, G. R., Perrucci, M. G., & Ebisch, S. J. H. (2020). Sense or sensibility? The neuro-functional basis of the structural matching effect in persuasion. *Cognitive, Affective, & Behavioral Neuroscience*, *20*(3), 536–550. https://doi.org/10.3758/s13415-020-00784-7

Bago, B., Rand, D. G., & Pennycook, G. (2020). Fake news, fast and slow: Deliberation reduces belief in false (but not true) news headlines. *Journal of Experimental Psychology: General*, *149*(8), 1608–1613. https://doi.org/10.1037/xge0000729

Baum, J., & Abdel Rahman, R. (2021). Emotional news affects social judgments independent of perceived media credibility. *Social Cognitive and Affective Neuroscience*, *16*(3), 280–291. https://doi.org/10.1093/scan/nsaa164

Beauvais, C. (2022). Fake news: Why do we believe it? *Joint Bone Spine*, *89*(4), Article 105371. https://doi.org/10.1016/j.jbspin.2022.105371

Blinova, E. N., & Shcherbakova, O. V. (2022). Vliyanie formata teksta na kognitivnuyu otsenku dostovernosti ego soderzhaniya [The influence of text format on the cognitive assessment of content credibility]. *Psikhologicheskii Zhurnal*, *43*(3), 89–101. https://doi.org/10.31857/S020595920020499-4 (in Russian)

Brashier, N. M., & Schacter, D. L. (2020). Aging in an era of fake news. *Current Directions in Psychological Science*, *29*(3), 316–323. https://doi.org/10.1177/0963721420915872

Brydges, C. R., & Ecker, U. K. H. (2018). *Updating misinformation in memory after correction: An event-related potentials (ERP) study*. https://doi.org/10.17605/OSF.IO/VKZXM

Brydges, C. R., Gignac, G. E., & Ecker, U. K. H. (2018). Working memory capacity, short-term memory capacity, and the continued influence effect: A latent-variable analysis. *Intelligence*, *69*, 117–122. https://doi.org/10.1016/j.intell.2018.03.009

Brydges, C. R., Gordon, A., & Ecker, U. K. H. (2020). Electrophysiological correlates of the continued influence effect of misinformation. *Journal of Cognitive Psychology*, *32*(8), 771–784. https://doi.org/10.1080/20445911.2020.1849226

Bunzeck, N., & Düzel, E. (2006). Absolute coding of stimulus novelty in the human substantia nigra/VTA. *Neuron*, *51*(3), 369–379. https://doi.org/10.1016/j.neuron.2006.06.021

Carrasco-Farré, C. (2022). The fingerprints of misinformation: How deceptive content differs from reliable sources in terms of cognitive effort and appeal to emotions. *Humanities and Social Sciences Communications*, *9*(1), 162. https://doi.org/10.1057/s41599-022-01174-9

Compton, J. (2013). Inoculation theory. In J. P. Dillard & L. Shen (Eds.), *The SAGE handbook of persuasion* (2nd ed., pp. 220–236). Sage. https://doi.org/10.4135/9781452218410.n14

Cook, J., Lewandowsky, S., & Ecker, U. K. H. (2017). Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence. *PLoS ONE*, *12*(5), Article e0175799. https://doi.org/10.1371/journal.pone.0175799

Corneille, O., Mierop, A., & Unkelbach, C. (2020). Repetition increases both the perceived truth and fakeness of information: An ecological account. *Cognition*, *205*, Article 104470. https://doi.org/10.1016/j.cognition.2020.104470

Craver, C. F. (2007). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. Clarendon Press.

Damstra, A., Boomgaarden, H. G., Broda, E., Lindgren, E., Strömbäck, J., Tsfati, Y., & Vliegenthart, R. (2021). What does fake look like? A review of the literature on intentional deception in the news and on social media. *Journalism Studies*, *22*(14), 1947–1963. https://doi.org/10.1080/1461670X.2021.1979423

De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science*, *313*(5787), 684–687. https://doi.org/10.1126/science.1128356

Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., Kendeou, P., Vraga, E. K., & Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, *1*(1), 13–29. https://doi.org/10.1038/s44159-021-00006-y

Edelson, M. G., Dudai, Y., Dolan, R. J., & Sharot, T. (2014). Brain substrates of recovery from misleading influence. *Journal of Neuroscience*, *34*(23), 7744–7753. https://doi.org/10.1523/JNEUROSCI.4720-13.2014

Edelson, M. G., Sharot, T., Dolan, R. J., & Dudai, Y. (2011). Following the crowd: Brain substrates of long-term memory conformity. *Science*, *333*(6038), 108–111. https://doi.org/10.1126/science.1203557

Gordon, A., Ecker, U. K. H., & Lewandowsky, S. (2017). Exploring the neural substrates of misinformation processing. *Neuropsychologia*, *106*, 216–224. https://doi.org/10.1016/j.neuropsychologia.2017.10.003

Gordon, A., Ecker, U. K. H., & Lewandowsky, S. (2019). Keeping track of 'alternative facts': The neural correlates of processing misinformation corrections. *NeuroImage*, *193*, 46–56. https://doi.org/10.1016/j.neuroimage.2019.03.014

Gronchi, G., & Giovannelli, F. (2018). Dual process theory of thought and default mode network. *Frontiers in Psychology*, *9*, Article 1237. https://doi.org/10.3389/fpsyg.2018.01237

Gronchi, G., Gavazzi, G., Viggiano, M. P., & Giovannelli, F. (2024). Dual-process theory of thought and inhibitory control: An ALE meta-analysis. *Brain Sciences*, *14*(1), Article 101. https://doi.org/10.3390/brainsci14010101

Gruber, M. J., Gelman, B. D., & Ranganath, C. (2014). States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron*, *84*(2), 486–496. https://doi.org/10.1016/j.neuron.2014.08.060

Guo, S., Chen, D., & Hu, X. (2025). Providing an alternative explanation improves misinformation rejection and alters event-related potentials during veracity judgements. *Brain and Cognition*, *186*, Article 106290. https://doi.org/10.1016/j.bandc.2025.106290

Jeng, J. H., Kasangu, G. A. B., Starke, A. D., Knudsen, E., & Trattner, C. (2024, October). *Negativity Sells? Using an LLM to Affectively Reframe News Articles in a Recommender System*. Paper presented at the International Workshop on News Recommendation and Analytics (INRA 2024), co-located with the 18th ACM Conference on Recommender Systems, Bari, Italy. https://mediafutures.no/inra_jeng/

Jia, L., Jin, H., & Jin, X. (2024). Neural mechanisms of the continued influence effect of misinformation: Analysis based on fMRI causal connectivity. *Neuroscience Letters*, *836*, Article 137861. https://doi.org/10.1016/j.neulet.2024.137861

Jin H., Jia, L., Yin, X., Yan, S., Wei, S., & Chen, J. (2022). The neural basis of the continued influence effect of misinformation. *Acta Psychologica Sinica*, *54*(4), 343–354. https://doi.org/10.3724/SP.J.1041.2022.00343

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Kalogeropoulos, A., Suiter, J., Udris, L., & Eisenegger, M. (2019). News media trust and news consumption: Factors related to trust in news in 35 countries. *International Journal of Communication*, *13*, 3672–3693.

Kaplan, J. T., Gimbel, S. I., & Harris, S. (2016). Neural correlates of maintaining one's political beliefs in the face of counterevidence. *Scientific Reports*, *6*, Article 39589. https://doi.org/10.1038/srep39589

Kappes, A., Harvey, A. H., Lohrenz, T., Montague, P. R., & Sharot, T. (2020). Confirmation bias in the utilization of others' opinion strength. *Nature Neuroscience*, *23*(1), 130–137. https://doi.org/10.1038/s41593-019-0549-2

Karanian, J. M., Rabb, N., Wulff, A. N., Torrance, M. G., Thomas, A. K., & Race, E. (2020). Protecting memory from misinformation: Warnings modulate cortical reinstatement during memory retrieval. *Proceedings of the National Academy of Sciences*, *117*(37), 22771–22779. https://doi.org/10.1073/pnas.2008595117

Karanian, J. M., Thomas, A. K., & Race, E. (2024). Warning before misinformation exposure modulates memory encoding. *Cognitive, Affective, & Behavioral Neuroscience*, *24*, 601–619. https://doi.org/10.3758/s13415-024-01183-y

Kossowska, M., Szwed, P., Czarnek, G., Honcharowa, S., Domagalik, A., & Doradzińska, Ł. (2026). Neural correlates of belief change in political and non-political domains among left-wing individuals confronted with counterarguments. *Scientific Reports*, *16*, Article 35397. https://doi.org/10.1038/s41598-026-35397-6

Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science*, *359*(6380), 1094–1096. https://doi.org/10.1126/science.aao2998

Lewandowsky, S., & van der Linden, S. (2021). Countering misinformation and fake news through inoculation and prebunking. *European Review of Social Psychology*, *32*(2), 348–384. https://doi.org/10.1080/10463283.2021.1876983

Lewandowsky, S., Cook, J., Ecker, U. K. H., Albarracín, D., Amazeen, M. A., Kendeou, P., Lombardi, D., Newman, E. J., Pennycook, G., Porter, E., Rand, D. G., Rapp, D. N., Reifler, J., Roozenbeek, J., Schmid, P., Seifert, C. M., Sinatra, G. M., Swire-Thompson, B., van der Linden, ... Zaragoza, M. S. (2020). *The debunking handbook 2020*. Center for Climate Change Communication. https://doi.org/10.17910/b7.1182

Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, *13*(3), 106–131. https://doi.org/10.1177/1529100612451018

Lois, G., Tsakas, E., Yuen, K., & Riedl, A. (2024). Tracking politically motivated reasoning in the brain. *Social Cognitive and Affective Neuroscience*, *19*(1), Article nsae056. https://doi.org/10.1093/scan/nsae056

Martel, C., Pennycook, G., & Rand, D. G. (2020). Reliance on emotion promotes belief in fake news. *Cognitive Research: Principles and Implications*, *5,* Article 47. https://doi.org/10.1186/s41235-020-00252-3

McIlhiney, P., Gignac, G. E., Ecker, U. K. H., Kennedy, B. L., & Weinborn, M. (2023). Executive function and the continued influence of misinformation: A latent-variable analysis. *PLoS ONE, 18*(4), Article e0283951. https://doi.org/10.1371/journal.pone.0283951

Metcalfe, J., Butterfield, B., Habeck, C., & Stern, Y. (2012). Neural correlates of people's hypercorrection of their false beliefs. *Journal of Cognitive Neuroscience*, *24*(7), 1571–1583. https://doi.org/10.1162/jocn_a_00228

Molina, R., Crespo, Y., Árbol, J. R., Arias-Orduña, A. V., Ibáñez-Molina, A. J., & Iglesias-Parro, S. (2025). Exploring the neurophysiological basis of misinformation: A behavioral and neural complexity analysis. *Behavioural Brain Research, 487*, Article 115592. https://doi.org/10.1016/j.bbr.2025.115592

Moravec, P. L., Minas, R. K., & Dennis, A. R. (2019). Fake news on social media: People believe what they want to believe when it makes no sense at all. *MIS Quarterly*, *43*(4), 1343–1360. https://doi.org/10.25300/MISQ/2019/15505

Müller-Pinzler, L., Czekalla, N., Mayer, A. V., Schröder, A., Stolz, D. S., Paulus, F. M., & Krach, S. (2022). Neurocomputational mechanisms of affected beliefs. *Communications Biology*, *5*, Article 1241. https://doi.org/10.1038/s42003-022-04165-3

Mussel, P., Ulrich, N., Allen, J. J. B., Osinsky, R., & Hewig, J. (2016). Patterns of theta oscillation reflect the neural basis of individual differences in epistemic motivation. *Scientific Reports*, *6*, Article 29245. https://doi.org/10.1038/srep29245

Nabi, R. L. (2003). Exploring the framing effects of emotion: Do discrete emotions differentially influence information accessibility, information seeking, and policy preference? *Communication Research*, *30*(2), 224–247. https://doi.org/10.1177/0093650202250881

Nevo, D., & Horne, B. D. (2022). How topic novelty impacts the effectiveness of news veracity interventions. *Communications of the ACM*, *65*(2), 68-75. https://doi.org/10.1145/3460350

Orchinik, R., Bhui, R., & Rand, D. G. (2025). Replicability and generalizability of the repeated exposure effect on moral condemnation of fake news. *Nature Communications*, *16*, Article 7206. https://doi.org/10.1038/s41467-025-62462-x

Paschen, J. (2020). Investigating the emotional appeal of fake news using artificial intelligence. *Journal of Product & Brand Management*, *29*(2), 223-233. https://doi.org/10.1108/JPBM-12-2018-2179

Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences*, *25*(5), 388–402. https://doi.org/10.1016/j.tics.2021.02.007

Photiou, A., Nicolaides, C., & Dhillon, P. S. (2021). Social status and novelty drove the spread of online information during the early stages of COVID-19. *Scientific Reports*, *11*, Article 20098. https://doi.org/10.1038/s41598-021-99060-y

Piccinini, G., & Craver, C. (2011). Integrating psychology and neuroscience: Functional analyses as mechanism sketches. *Synthese*, *183*(3), 283-311. https://doi.org/10.1007/s11229-011-9898-4

Poldrack, R. A. (2010). Mapping mental function to brain structure: How can cognitive neuroimaging succeed? *Perspectives on psychological science*, *5*(6), 753–761. https://doi.org/10.1177/1745691610388777

Scholz, C., Jovanova, M., Baek, E. C., & Falk, E. B. (2020). Media content sharing as a value-based decision. *Current Opinion in Psychology*, *31*, 83–88. https://doi.org/10.1016/j.copsyc.2019.08.004

Song, H., So, J., Shim, M., Kim, J., Kim, E., & Lee, K. (2023). What message features influence the intention to share misinformation? *Computers in Human Behavior*, *138*, Article 107439. https://doi.org/10.1016/j.chb.2022.107439

Turcotte, J., York, C., Irving, J., Scholl, R. M., & Pingree, R. J. (2015). News recommendations from social media opinion leaders: Effects on media trust and information seeking. *Journal of Computer-Mediated Communication*, *20*(5), 520–535. https://doi.org/10.1111/jcc4.12127

Van der Linden, S., & Cohen, M. S. (2025). The neuroscience of misinformation: A research agenda. *Neuron*, *113*, 2225-2228. https://doi.org/10.1016/j.neuron.2025.05.010

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Wang, W.-C., Brashier, N. M., Wing, E. A., Marsh, E. J., & Cabeza, R. (2016). On known unknowns: Fluency and the neural mechanisms of illusory truth. *Journal of Cognitive Neuroscience*, *28*(5), 739–746. https://doi.org/10.1162/jocn_a_00923

Westen, D., Blagov, P. S., Harenski, K., Kilts, C., & Hamann, S. (2006). Neural bases of motivated reasoning: An fMRI study of emotional constraints on partisan political judgment in the 2004 US Presidential Election. *Journal of Cognitive Neuroscience*, *18*(11), 1947–1958. https://doi.org/10.1162/jocn.2006.18.11.1947

Whittemore, R., & Knafl, K. (2005). The integrative review: Updated methodology. *Journal of Advanced Nursing*, *52*(5), 546–553. https://doi.org/10.1111/j.1365-2648.2005.03621.x

Wittmann, B. C., Bunzeck, N., Dolan, R. J., & Düzel, E. (2007). Anticipation of novelty recruits reward system and hippocampus while promoting recollection. *NeuroImage*, *38*(1), 194–202. https://doi.org/10.1016/j.neuroimage.2007.06.038

Wu, H., Luo, Y., & Feng, C. (2016). Neural signatures of social conformity: A coordinate-based ALE meta-analysis. *Neuroscience & Biobehavioral Reviews*, *71*, 101–111. https://doi.org/10.1016/j.neubiorev.2016.08.038

Zhao, S., Wang, T., & Xiong, B. (2025). Neural correlates of conspiracy beliefs during information evaluation. *Scientific Reports*, *15*, Article 18375. https://doi.org/10.1038/s41598-025-03723-z