Cognitive Intelligence Compensation for Aphantasia: A Methodology-Oriented Narrative Review of Technologies, Challenges, and Future Directions

Jiayi Liu[1], Lijuan Chen[1], Luping Wang[1,2,3,*]

[1] School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai, China
[2] School of Gongli Hospital Medical Technology, University of Shanghai for Science and Technology, Shanghai, China
[3] Laboratory of Algorithms for Cognitive Models, School of Computer Science, Fudan University, Shanghai, China

*__Correspondence:__ Luping Wang, School of Mechanical Engineering, University of Shanghai for Science and Technology, Jungong Road 516, Shanghai 200093, China.
E-Mail: 15110240007@fudan.edu.cn

Abstract

Background: Aphantasia is a neurocognitive variation characterized by the inability to voluntarily generate mental imagery, with an estimated prevalence of 2%–5% in the general population.

Methods: This methodology-oriented narrative review presents an examination of representative neurocognitive, behavioral, and computational studies associated with the generation of mental imagery and cognitive compensation for aphantasia. The studies under review are categorized into visual imagery generation, auditory imagery generation, multimodal mental imagery, and cognitive compensation strategy methods.

Results: The reviewed evidence suggests that the subjective absence of visual imagery does not necessarily lead to uniform objective impairment, because some individuals perform tasks using analytical, verbal, spatial, or other nonvisual strategies. A key contribution of this review consists in affirming that computational representations are not replicas of neural representations, but may be functionally connected through encoding, decoding, cross-modal alignment, and feedback. The reviewed evidence inspires us to propose a conceptual framework wherein computational systems may provide substitutive, assistive, or augmentative support for imagery-related tasks.

Conclusions: By connecting computational methods with neurocognitive evidence, this framework offers a conceptual basis for developing individualized and multimodal assistive technologies. Nevertheless, because the majority of neural reconstruction studies have been conducted in participants without aphantasia, the proposed framework still requires direct empirical validation.

Keywords: aphantasia; imagination disorder; ventral pathway; cognitive compensation; mental imagery

Introduction

Mental imagery is an essential component of human cognitive function. It refers to the ability to mentally generate, manipulate, and experience sensory information in the absence of corresponding external sensory input[1–4].

Received: 4 July 2026; **Accepted:** 11 September 2026
DOI: 10.15302/IH.2026.000010
© 2026 The Author(s). Published by Higher Education Press. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0).

Visual mental imagery commonly referred to as the "mind's eye," enables the mental visualization of objects that are not physically present, ranging from familiar faces to imagined scenes, from the mental rotation of geometric figures to the contour completion of occluded objects. Not only does this capacity underpin everyday thinking and creativity, but it also plays an important role in higher-order cognitive processes, including spatial navigation, episodic memory retrieval, emotion regulation, social cognition, and complex problem solving. In 2015, Zeman and colleagues introduced the term 'aphantasia' to describe reduced or absent voluntary visual imagery[5]. Over the next decade, researchers began paying increasing attention to aphantasia across cognitive neuroscience, psychology, and clinical medicine. Epidemiological surveys suggest a 2%–5% prevalence in the general population. Individuals with aphantasia frequently report that they are unable to perform simple tasks such as "imagining an apple in the mind," and encounter difficulties with mental rotation, scene construction, facial imagery, and other visual imagery operations[6]. Notably, aphantasia is not a newly discovered disorder but a long-overlooked cognitive difference; many individuals only realize in adulthood that their "inner visual" experience differs fundamentally from that of others[7,8]. Because impairments in mental imagery may significantly affect cognitive performance across a wide range of everyday and professional activities, developing effective cognitive compensation strategies is important for this population.

The significance of aphantasia research extends beyond its clinical scope. From a cognitive-science perspective, aphantasia serves as a natural "experimental window" into the neural basis of visual mental imagery, facilitating insights into the relationship between imagery and perception[9] as well as the diversity of human conscious experience. From the perspective of clinical practice, helping individuals with aphantasia compensate for the functional difficulties associated with absent or reduced voluntary visual imagery has become an important research goal in cognitive rehabilitation. However, aphantasia is not considered as a disease requiring treatment; rather, it is a variation in cognitive phenotype[10–12]. As such, supportive strategies should aim not to "cure" but to "compensate." Developing cognitive compensation strategies to address weaknesses in specific cognitive functions through external technological means or alternative cognitive pathways has emerged as a core direction for supportive approaches in aphantasia. This conceptual distinction helps clarify the aims of future support strategies.

Despite recent progress, three research gaps remain particularly relevant to cognitive compensation. First, current computational approaches are often discussed in isolation with limited comparison of their suitability for different imagery operations and sensory modalities. Second, the association between computational and neural representations remains poorly elucidated, and the generalizability of neural-reconstruction findings from healthy participants to individuals with aphantasia is uncertain. Third, an integrated and empirically validated framework that links computational methods, individual variability, multimodal support, and clinically relevant evaluation is still lacking.

Therefore, this review has three primary goals. First, it aims to classify and compare representative mental imagery generation methods, including geometric reasoning, deep learning–based reconstruction, neural decoding, and multimodal approaches. Second, it examines how computational representations may be functionally associated with neural representations and discusses the limitations of applying findings from healthy participants to individuals with aphantasia. Third, the review integrates fundamental imagery operations and compensation strategies into a conceptual framework while emphasizing evaluation requirements, current limitations, and future research directions.

The review begins with the theoretical background on mental imagery and aphantasia, followed by the review methodology and representative imagery generation approaches. The cognitive compensation discussion examines fundamental imagery operations and support strategies. The evaluation and discussion then compare the approaches and present a conceptual framework. The conclusions summarize the main findings and future research directions.

Theoretical background

This section discusses the theoretical foundations

underlying mental imagery generation, its neural substrates, the neurocognitive mechanisms of aphantasia, and the multimodal nature of mental imagery. These foundations strengthen the subsequent review of computational approaches and cognitive compensation strategies.

Mental imagery: definition and classic theories

Mental imagery constitutes an essential component of human cognitive function. It refers to the ability to actively generate, manipulate, and experience sensory information in the mind in the absence of corresponding external sensory input. In particular, visual mental imagery, commonly referred to as the "mind's eye," enables individuals to mentally "see" objects, scenes, and events that are not physically present. This capacity facilitates a wide range of higher-order cognitive functions, such as spatial navigation, episodic memory retrieval, emotion regulation, social cognition, and complex problem solving.

The computational theory, elaborated by Kosslyn[13] in the 1990s, has largely served as the basis for the theoretical understanding of mental imagery. This framework posits that visual imagery and visual perception share the same neural computational architecture; in other words, imagery generation operates as the "reverse" of the perceptual system in the absence of external input. Kosslyn distinguished between two types of internal representations: "image files," containing spatial and geometric information about objects, and "propositional files," containing semantic and categorical information. This dual-representation framework offers a computational basis for understanding how the brain generates, maintains, and manipulates visual mental images.

Pearson et al.[2] later proposed a functional mechanism model of mental imagery, highlighting its central role in perceptual prediction and cognitive reasoning. This predictive framework has important implications for understanding deficits in mental imagery and their cognitive consequences.

Neural substrates of visual mental imagery

Functional neuroimaging techniques have been used to extensively investigate the neural basis of visual mental imagery, and accumulating evidence has highlighted the critical role of the ventral visual pathway. Extending from the primary visual cortex (V1) through visual areas V2 and V4 to the inferotemporal cortex (IT), this pathway constitutes the core neural circuitry for object recognition and visual imagery generation[3].

Neuroimaging studies offer the key finding that visual imagery and visual perception engage overlapping neural substrates. Although imagery-related activation is generally weaker and more dependent on prefrontal and parietal control regions, when individuals perform visual imagery tasks, top-down activation patterns in the ventral pathway closely resemble those observed during actual visual perception[4]. Functional magnetic resonance imaging (fMRI) studies have consistently shown that category-selective regions in the ventral temporal cortex are activated by the act of imagining objects, including the fusiform face area for faces and the parahippocampal place area for scenes[14].

Neuroimaging findings should be distinguished from behavioral evidence concerning sensory imagery. Keogh and Pearson used a binocular rivalry task and questionnaires to assess imagery in individuals who reported having aphantasia[15]. Their findings provided behavioral evidence for reduced sensory visual imagery. These measures do not directly assess neural activation. This study therefore cannot establish reduced activation in the fusiform gyrus, lateral occipital complex, or lingual gyrus; such anatomical claims require independent neuroimaging evidence.

Beyond the ventral pathway, imagery generation recruits frontoparietal control networks that regulate the top-down reactivation of sensory representations. These control regions, including the dorsolateral prefrontal cortex and intraparietal sulcus, are thought to initiate and guide the imagery process, while the ventral pathway provides the content-specific sensory representations[3].

Neurocognitive mechanisms of aphantasia: competing accounts

Recent studies in neuroimaging have provided extensive evidence concerning aphantasia. Behavioral measures

assess sensory imagery, whereas neuroimaging examines neural systems. Together, they inform competing explanations of imagery differences. Behavioral effects alone cannot establish which brain regions show altered activity. Using lesion network mapping, Kutsche et al.[16] reported that brain lesions causing acquired aphantasia were all functionally associated with the imagery node in the fusiform gyrus, revealing a causal role of the ventral pathway in imagery generation. Furthermore, Noble et al.[17] found reduced temporal signal complexity in individuals with aphantasia, suggesting that their brains fail to effectively construct dynamic neural representations of imagery. Bartolomeo[18] proposed a taxonomic framework for aphantasia from a cognitive neuroanatomical perspective, distinguishing different subtypes and their neural underpinnings.

Despite increasing empirical attention, there is currently no consensus on the mechanisms underlying aphantasia. One study proposed that aphantasia is primarily due to an absence or substantial reduction of sensory visual representations[7]. Evidence from binocular-rivalry paradigms has indicated that there are reduced sensory imagery effects in individuals with aphantasia, suggesting that the phenomenon cannot be explained solely by inaccurate self-report or poor metacognitive awareness[15]. Conversely, other studies have indicated that at least some individuals may retain task-relevant representations without experiencing them as vivid conscious imagery. For example, preserved objective memory performance despite reduced subjective vividness has been interpreted as evidence that aphantasia may, in some individuals, involve differences in metacognitive access or conscious awareness rather than a complete absence of internal representations[12]. Providing a further perspective, a recent integration account proposed that rudimentary sensory information may be reactivated but cannot undergo the multistage integration required for conscious imagery experience[19]. These competing explanations—absent sensory representation, reduced conscious access, and incomplete integration—suggest that more than one neurocognitive pathway may be involved in the occurrence of aphantasia.

A related controversy concerns whether the subjective absence of imagery necessarily leads to impairments in objective cognitive performance. While individuals with aphantasia consistently report absent or markedly reduced voluntary imagery, objective findings vary substantially across tasks. Some studies have highlighted differences in object memory or imagery-related sensory measures, whereas others have reported preserved accuracy in spatial memory, visual working memory, and mental-rotation tasks. In mental rotation, for example, individuals with aphantasia may respond more slowly but more accurately than individuals with typical visual imagery, apparently because they rely more strongly on analytical strategies[20]. Correspondingly, recent evidence has indicated that individuals with aphantasia may be distributed across distinct spatial and verbal cognitive profiles rather than forming a single homogeneous group[21]. This heterogeneity may also characterize congenital and acquired aphantasia: acquired cases may be associated with identifiable neurological disruption, whereas congenital aphantasia may indicate more variable developmental and network-level mechanisms[16,18]. Therefore, distinctions between complete and reduced imagery, as well as between congenital and acquired aphantasia, should be taken into account, and it should not be assumed that aphantasia arises from a single underlying mechanism.

Multimodal mental imagery: beyond the visual domain

Although aphantasia is primarily characterized by deficits in visual imagery, mental imagery is not a unitary visual phenomenon. Through their multimodal theory of mental imagery, Nanay[22] argued that mental imagery encompasses multiple sensory modalities, including visual, auditory, tactile, and kinesthetic, which are integrated into a unified cognitive function. This theoretical perspective aids both in understanding aphantasia and in designing comprehensive compensation strategies.

Music listening provides a context for examining how visual imagery relates to emotional experience. Hashim et al.[23] conducted two online surveys comparing individuals with aphantasia and controls. The first compared 51 participants in each group during music listening; the second compared 29 participants in each group on emotion recognition and everyday music experience. The surveys revealed differences in emotional intensity and music-related recollection, but not in emotion discrim-

ination or musical reward sensitivity. These findings concern music experience; they do not establish an auditory imagery deficit or demonstrate effective compensation.

The multimodal theory further implies that individuals with aphantasia may compensate for absent mental imagery in daily life by relying more heavily on alternative cognitive strategies, including verbal coding, semantic organization, and analytical reasoning. For example, using qualitative methods, Hayes et al.[24] identified three possible cognitive compensation techniques used by individuals with aphantasia: semantic reliance, condensation of inner speech, and external recoding. These behavioral strategies align with the theoretical expectation that individuals without visual imagery may achieve comparable task outcomes by recruiting nonvisual cognitive systems.

With regard to cognitive compensation, the multimodal perspective suggests that support strategies should not be restricted to visual-channel compensation. When generation of visual imagery is challenging or absent, compensatory support could leverage auditory, tactile, or other sensory channels to aid in cognitive task completion.

Theoretical implications for cognitive compensation

The theoretical foundations reviewed above provide a conceptual basis for understanding mental imagery generation and its potential compensation in individuals with aphantasia. Four implications are particularly relevant for this review.

First, for understanding imagery operations—including object completion, spatial transformation, and mental rotation—Kosslyn's computational theory provides a framework that can be computationally modeled and potentially externalized through assistive systems. The distinction between image files (spatial information) and propositional files (semantic information) indicates that effective compensation strategies may need to address both the spatial and semantic aspects of imagery.

Second, the ongoing debates concerning the neurocognitive mechanisms of aphantasia imply that no single compensation strategy is likely to suit all individuals. The heterogeneity of aphantasia, whether considered in terms of sensory representation, conscious access, integration processes, or cognitive profile, highlights the need for personalized approaches that consider individual characteristics and preserved cognitive strategies.

Third, the multimodal nature of mental imagery suggests that compensation strategies should extend beyond the visual domain to include auditory, tactile, and other sensory modalities. This is particularly important in light of evidence that imagery deficits in aphantasia may extend across several sensory modalities.

Fourth, the dissociation between theoretical accounts of aphantasia and the design of existing computational techniques highlights a central gap: most existing technologies have been developed for general-purpose imagery generation or neural decoding rather than specifically for cognitive compensation in individuals with aphantasia. This gap provides the impetus for the methodology-focused review that follows and the conceptual framework proposed in later sections.

Review methodology

This study serves as a methodology-oriented narrative review rather than a systematic review or meta-analysis. Recent advances in cognitive neuroscience, artificial intelligence, and multimodal information processing have fostered the development of intelligent cognitive compensation technologies for individuals with aphantasia. Instead of reviewing these studies chronologically, the present review adopted a methodology-oriented perspective to provide a structured conceptual synthesis of representative studies. Based on their underlying technical principles and functional characteristics, the methods of existing studies were classified into four groups: visual imagery generation methods, auditory imagery generation methods, multimodal mental imagery methods, and cognitive compensation strategy methods. This classification established a coherent foundation for the subsequent comparative analysis of representative methodologies, their underlying cognitive mechanisms, technical advantages, existing limitations, and future research directions. The reviewed literature was selected

from representative studies in cognitive neuroscience, artificial intelligence, and assistive technologies, with particular emphasis on research addressing mental imagery generation, neural representation, and cognitive compensation mechanisms. Studies were considered relevant when they addressed the neurocognitive or behavioral features of aphantasia, representative strategies for mental imagery generation or neural decoding, alternative cognitive strategies, or technologies with potential relevance to cognitive compensation. Given that this narrative review did not employ a preregistered systematic-review protocol, exhaustive database screening, or formal risk-of-bias assessment, it should not be interpreted as a comprehensive systematic review of all available evidence. Its aim is to organize representative interdisciplinary evidence while identifying conceptual relationships and priorities for future empirical investigation.

Visual imagery generation methods

Visual imagery generation is one of the most extensively studied areas in mental imagery. The related methods are an important technological foundation for intelligent cognitive compensation in individuals with aphantasia. These methods are primarily aimed at reconstructing or generating complete visual mental representations from incomplete sensory information, prior knowledge, or learned representations. In accordance with their underlying computational principles, existing visual imagery generation approaches can be classified into several representative categories, each exhibiting distinct characteristics with respect to interpretability, reconstruction accuracy, computational complexity, and applicability.

Geometric topology–based reasoning methods

Among the existing approaches, geometric topology–based reasoning methods are regarded as one of the well-established and cognitively interpretable paradigms. Drawing on the sensitivity of the human visual system to geometric relationships, these methods extract structural information of objects (e.g., contours, line segments, corner points, intersections) to construct relational matrices or topological representations, and then employ geometric constraint solving to achieve occlusion completion and morphological transformations. Marr's[25] computational theory of vision serves as the classic theoretical foundation for this direction, positing that the reconstruction of three-dimensional structure from two-dimensional images fundamentally relies on geometric reasoning. In Kosslyn's computational theory of imagery[13], he further pointed out that the generation and manipulation of visual imagery depend on "image files" maintained in the brain—structured representations containing spatial-relational information about objects. In recent years, methods based on line-segment relational matrices have demonstrated strong performance in occlusion completion and mental rotation tasks. These methods extract visible contour segments from objects, compute geometric features such as relative positions, distances, angles, intersections, and adjacency relationships among segments, construct a complete geometric relational matrix, and then iteratively generate missing parts through constraint solving.

The following formulation is included as a conceptual illustration of how geometric relationships may be represented and does not constitute a newly implemented or experimentally validated model.

The geometric relational matrix can be formally defined as follows. This matrix is the core data structure of geometric topology methods, encoding the contour segments of an object and their mutual spatial constraints into a computable mathematical object.

Let the object contour consist of N line segments. Define the geometric relational matrix G as an $N \times N$ matrix whose element g_{ij} describes the geometric relationship between the i-th segment L_i and the j-th segment L_j. Each g_{ij} is a quadruple (Equation 1):

$$g_{ij} = (r_{ij}, \theta_{ij}, d_{ij}, t_{ij}) \quad (1)$$

where r_{ij} denotes the relative positional relationship (e.g., parallel, perpendicular, intersecting, or collinear), θ_{ij} is the included angle ($0 \leq \theta_{ij} \leq \pi$), d_{ij} is the minimum Euclidean distance, and t_{ij} is the adjacency type (endpoint connection, crossing connection, or no connection).

The geometric relational matrix G provides a topology-oriented representation of object structures, where the representation is mainly determined by spatial relationships among constituent elements rather than absolute

coordinates or object dimensions. This representational property allows subsequent object completion and mental rotation operations to be performed while preserving geometric constraints, thereby ensuring the physical plausibility of the generated or transformed outcomes. For cognitive compensation applications in aphantasia, such interpretable geometric representations provide a transparent mechanism for spatial reasoning and offer complementary advantages compared with black-box deep learning approaches, particularly in terms of interpretability and traceability. The main strengths of this approach include: (1) high interpretability, as every completion step can be explicitly tracked and verified; (2) independence from large-scale training data, reducing the black-box limitation commonly associated with some deep learning–based approaches; and (3) high structural consistency of the generated imagery. However, these methods still face challenges when dealing with highly complex or unstructured scenes.

Deep learning–based reconstruction methods

The rise of deep learning has opened entirely new technical pathways for visual imagery generation. Generative adversarial networks (GANs), variational autoencoders (VAEs), and diffusion models have achieved substantial advances in tasks such as image inpainting, super-resolution, and image generation[26–28]. In the domain of visual imagery generation, deep learning–based image reconstruction methods can recover complete visual content from partial observations. For instance, Shen et al.[29] proposed an image reconstruction method based on deep generative models that can reconstruct images viewed by participants from fMRI signals. Pan et al.[30], targeting electroencephalography (EEG) signals, introduced the Deep Visual Representation Model (DVRM), which employs residual dense blocks in the encoder to extract distributional features between EEG and images, and uses deep neural network decoders to reconstruct visual stimulus images from learned deep visual representations, achieving end-to-end image generation. The advantage of these methods lies in their generation results, which typically surpass traditional methods in visual realism and detail richness. However, deep learning approaches have significant limitations: (1) limited interpretability, which makes it difficult to explicitly understand the relationship between learned representations and underlying cognitive mechanisms; (2) heavy dependence on the quality and scale of training data, with limited generalization to scenarios outside the training distribution; and (3) difficulty in establishing correspondences with cognitive theories, which constrains their theoretical persuasiveness in aphantasia-assistive interventions.

Neural decoding–based methods

Neural decoding–based methods aim to infer or reconstruct visual imagery-related information from neural signals. These methods typically acquire fMRI, EEG, or magnetoencephalography (MEG) signals from participants during visual perception or imagery tasks, and build mapping relationships between neural activity patterns and perceptual/imagery content using machine learning models[31,32]. Horikawa and Kamitani[33] proposed a method that decodes visual content from fMRI signals using deep neural network–based image classifiers, enabling prediction of object categories participants are viewing or imagining. Naselaris et al.[34] employed voxel-wise encoding models to decode mental imagery by modeling visual perceptual activity, confirming that low-level visual features are encoded in imagery and demonstrating the feasibility of imagery-guided image search. Lee et al.[35], from a brain-computer interface (BCI) perspective, compared the decoding of visual imagery and imagined speech using EEG signals, finding that high-frequency neural activity and language-related brain regions contributed significantly to both tasks, validating the effectiveness of visual imagery as a control signal for multi-class BCI. The outstanding advantage of neural decoding methods is that they establish a direct mapping between neural activity and mental imagery, providing a neural-level standard for objective assessment of imagery generation quality. For individuals with aphantasia, neural decoding techniques may provide objective biomarkers for characterizing residual imagery abilities and identifying individual differences, which could support personalized cognitive compensation strategies. However, these methods are technically demanding, rely on expensive neuroimaging equipment, have poor real-time performance, and are currently difficult to translate into portable daily assistive tools.

However, most of the neural reconstruction studies

included in this review only included healthy participants with typical imagery. Although these studies highlight the technical feasibility of associating neural activity with visual or imagery-associated content, their findings cannot be directly generalized to individuals with aphantasia, who may have different imagery-associated neural activity, subjective experiences, and cognitive strategies. Future studies should therefore directly evaluate individuals with aphantasia while combining reconstruction accuracy with behavioral and subjective assessments. Until such validation is available, neural reconstruction should be considered a potential technical basis for cognitive compensation rather than an established intervention for aphantasia.

Auditory imagery generation methods

Auditory mental imagery—that is, the ability to "hear" sounds or music in the mind—is an important component of mental imagery, yet it has long received insufficient attention. Auditory imagery should be distinguished from emotional responses to externally presented music. Hashim et al.[23] found differences in selected aspects of music experience in individuals with aphantasia, while emotion discrimination and musical reward sensitivity did not differ from controls. These findings indicate that music experience is not uniformly affected by reduced visual imagery. They provide context for considering individual differences, rather than direct evidence of impaired auditory imagery or successful compensation. Any proposed auditory support therefore requires direct evaluation of imagery abilities and functional outcomes. At the computational modeling level, auditory imagery generation techniques are mainly divided into three categories: timbre parameter–based synthesis, deep learning–based audio generation, and neural signal–based decoding[36,37]. Compared with visual imagery, auditory imagery exhibits stronger temporal variability and is more difficult to model, yet it also holds unique value in cognitive compensation. For example, for individuals with aphantasia who cannot imagine a loved one's voice or a familiar melody, auditory imagery generation systems can provide external playback to compensate for the lack of internal auditory imagery. However, auditory imagery compensation still faces several challenges, including the complexity of modeling temporal auditory information, individual variability in auditory imagery ability, and the difficulty of achieving natural and personalized auditory experiences.

Multimodal mental imagery methods

Multimodal mental imagery methods aim to address the generation of mental imagery across different sensory channels in a unified manner. The core view of these methods is that mental imagery across modalities shares certain common cognitive-computational principles, yet differs in specific implementation mechanisms[22]. In 2011, Briscoe[38] examined how hidden object features are represented in amodal perception. His analysis challenged a unitary account of this capacity, instead arguing that amodal perception is not a single, unified cognitive ability but is jointly constituted by three distinct types of representations: mental imagery, genuine visual representations, and beliefs. This multi-level analysis provides a more refined theoretical framework for understanding the cognitive architecture of multimodal imagery. In computational modeling, the main challenge faced by multimodal mental imagery approaches lies in how to establish effective mappings between different modalities. Some researchers have attempted to use cross-modal learning techniques to unify visual and auditory imagery into a shared representational space[39]. For example, by aligning shape features of visual objects with corresponding sound features during training, models can perform cross-modal reasoning and generation. Such approaches offer unique advantages in aphantasia cognitive compensation: when visual imagery generation is limited, the compensatory function of auditory or tactile channels can be leveraged to assist in completing cognitive tasks. However, multimodal mental imagery approaches still face challenges, including the difficulty of aligning heterogeneous sensory representations, the requirement for large-scale multimodal datasets, and uncertainty regarding their effectiveness for individuals with different aphantasia subtypes.

Cognitive compensation strategy methods

Cognitive compensation strategies, as a core pathway for aphantasia-assistive interventions, have received increasing attention in recent years. Hayes et al.[24],

using qualitative research methods and thematic analysis of individuals with aphantasia, identified three possible cognitive compensation strategies—semantic reliance, condensation of inner speech, and external recoding—providing systematic qualitative evidence for understanding the behavioral patterns with which individuals with aphantasia cope with the absence of visual imagery. In the domain of spatial cognition, Kay et al.[20] found that individuals with aphantasia are slower but more accurate in mental rotation tasks; the key lies in their preference for analytical strategies (e.g., comparing local features one by one) rather than the holistic imagery-rotation strategies commonly used by controls, indicating that alternative strategies can effectively compensate for the absence of visual imagery. At the theoretical integration level, Scholz et al.[19] proposed an integrative model of aphantasia, arguing that reactivated sensory information must undergo multistage integration to produce imagery experience; this model explains aphantasia as the failure of sensory precursors to complete integration into perceptual representations, rather than as a product of unconscious imagery, thereby providing a unified theoretical framework for understanding the cognitive mechanisms of aphantasia. Regarding short-term memory binding, Delhaye et al.[40] found no difference between individuals with aphantasia and typical imagers in color-shape conjunction binding tasks; even under conditions where shapes were difficult to label verbally, aphantasic individuals' performance remained significantly above chance, suggesting that neither visual imagery nor verbal coding is necessary for conjunction binding. The authors speculated that individuals with aphantasia may rely on implicit representations or non-visual spatial/sensorimotor strategies to accomplish the task. Building on this, Delem et al.[21], using unsupervised clustering, revealed heterogeneous cognitive profiles within the aphantasic population—some favoring spatial strategies, others favoring verbal strategies—rather than a single cognitive mode. The classic study by Wicken et al.[41] demonstrated that individuals with aphantasia show significantly reduced physiological responses (e.g., skin conductance) during the imagination of frightening narratives compared to typically developing controls, suggesting that visual imagery can amplify emotion elicited by imagined events. Collectively, these studies reveal the complexity and diversity of cognitive compensation in aphantasia, providing behavioral and neuroscientific evidence for the development of technology-assisted compensation strategies.

Cognitive compensation

This section presents a review of the fundamental imagery operations that may support cognitive compensation, such as internal representation, object completion, displacement and scaling, mental rotation, and their integration into different compensation strategies.

Internal representation

To link these computational techniques to a neurocognitive basis of imagery, a critical distinction must be made between computational and neural representations. Computational representations are deliberately constructed numerical descriptions designed to support specific algorithmic operations and include geometric relational matrices, latent feature vectors, or cross-modal embeddings. Neural representations, by contrast, comprise distributed, dynamic, and context-dependent patterns of activity across interacting brain regions. As a result, computational representations should not be considered direct replicas of the representations used by the human brain.

The relationship between computational and neural representations is functional and probabilistic, as opposed to structurally identical. Nevertheless, operational links may be established using encoding, decoding, and representational alignment. Encoding models link neural activity to computational features; decoding models estimate perceptual or imagery-related information from neural signals, and cross-modal alignment connects information across visual, verbal, auditory, tactile, or neural domains.

For cognitive compensation, available sensory, linguistic, or neural inputs can be encoded into manipulable computational representations that support operations such as completion, transformation, rotation, or cross-modal integration. Subsequently, the resulting information may be externally presented through visual displays, verbal descriptions, auditory cues, or tactile feedback. Although this process may be conducive to imagery-related task performance, it should not be interpreted as directly

restoring subjective imagery or reproducing neural representations.

Internal representation is the foundation of mental imagery generation, referring to the manner in which an individual or system internally encodes and stores objects, scenes, or concepts. From a cognitive science perspective, the human brain's representation of objects is multi-layered and multimodal: it includes visual feature–based representations (shape, color, texture), spatial relation–based representations (position, orientation, scale), semantic concept–based representations (category, function, attributes), and movement dynamics–based representations. Kosslyn[13], in his computational theory, divided visual imagery representations into "image files" and "propositional files," where the former stores spatial information and the latter stores semantic information. In computational models, internal representations can be implemented in various ways. Geometric topology–based methods construct structured relational matrices using relative positions, distances, angles, and adjacency relationships among line segments as representations; deep learning–based methods utilize continuous vectors in latent space as implicit representations; neural decoding–based methods encode measured neural activity patterns into computational features for decoding or reconstruction.

In the context of aphantasia cognitive compensation, the quality of internal representation construction directly affects the efficacy of all subsequent mental imagery operations[42]. A well-designed internal representation should possess the following properties: (1) completeness, i.e., the ability to recover the complete structure of an object from partial information; (2) invariance, i.e., maintaining representational stability under transformations such as rotation, translation, and scaling; and (3) operability, i.e., supporting subsequent operations such as completion, transformation, and reasoning. It is worth noting that different object types may require different internal representation strategies. For structured objects (e.g., buildings, mechanical parts), representations based on geometric-topological relationships are generally more effective; for unstructured objects (e.g., natural scenes, organic forms), representations based on deep learning distributions may be more appropriate. In the domain of auditory mental imagery, internal representations may involve multi-dimensional encoding approaches including the temporal dynamic characteristics of audio spectra, timbre parameter spaces, or musical structural hierarchies. Therefore, the internal representation module in cognitive compensation strategies needs to flexibly select appropriate representation methods according to the characteristics of the target object.

Object completion

Object completion is one of the core operations of mental imagery, referring to the ability to infer and generate the complete structure of an object when only partial information is available. The human visual system possesses remarkable completion capabilities—even when most of an object is occluded, we can still accurately judge its complete shape. This ability contributes to everyday object recognition and spatial cognition. In individuals with aphantasia, object completion ability may differ from performance in individuals with typical imagery, which is a key factor hindering their daily cognitive functioning[15].

At the computational level, object completion can be formulated as a constrained energy minimization problem. Given partial observations, the completion algorithm needs to infer the most plausible complete structure while satisfying the observational constraints. We define the following energy function to quantify the quality of any completion candidate in Equation 2:

$$E(S_{comp}) = \lambda_1 \cdot E_{data}(S_{comp}, S_{obs}) + \lambda_2 \cdot E_{smooth}(S_{comp}) + \lambda_3 \cdot E_{prior}(S_{comp}) \quad (2)$$

where S_{obs} is the visible portion, S_{comp} is the completion result; E_{data} measures the consistency between the completion result and the visible portion, E_{smooth} encourages smoothness of the completed surface, E_{prior} incorporates prior knowledge about object shapes (e.g., symmetry, continuity); and λ_1, λ_2, λ_3 are weighting coefficients. In this conceptual formulation, the weighting coefficients λ_1, λ_2 and λ_3 represent the relative importance assigned to observational consistency, structural smoothness, and prior knowledge, respectively. Their optimal values may vary depending on object characteristics, task requirements,

and individual compensation scenarios, and should be determined through future empirical validation.

Equation 2 provides a general optimization framework, within which different implementation methods (e.g., geometric constraint solving, deep learning–based inpainting) essentially choose different specific forms of E_{data}, E_{smooth}, and E_{prior} as well as optimization strategies. However, Equation 2 represents a general conceptual form rather than a fully specified optimization model. The precise definitions of E_{data}, E_{smooth}, and E_{prior} depend on the selected implementation method, object type, dataset, and task requirements. No fixed numerical values are assigned to λ_1, λ_2 or λ_3 in this review, and these coefficients have not been optimized for individuals with aphantasia.

In a future implemented system, these coefficients would need to be determined empirically using appropriate training, validation, and test data. Ablation analyses could evaluate the contribution of each energy term by varying or removing individual coefficients and comparing structural accuracy, perceptual quality, task performance, cognitive load, and user experience. Therefore, the appropriate weighting scheme remains an open empirical question rather than an established parameter setting.

At the computational level, object completion can be implemented through various methods. Geometric constraint–based methods infer the structure of missing portions by analyzing geometric features of the visible parts (e.g., continuity, parallelism, symmetry, topological relationships of intersections). The advantage of these methods lies in maintaining geometric consistency of the completion results, making them suitable for rigid and structured objects. Deep learning–based methods leverage prior knowledge of objects learned from large-scale training data to infer missing content; image inpainting is a typical representative of such approaches[26,28]. In recent years, diffusion model–based inpainting methods have achieved breakthrough progress in visual quality, capable of generating object completion results with high consistency in both texture and semantics.

Regarding the evaluation of object completion, traditional visual imagery methods (e.g., geometric reasoning methods) typically focus on geometric fidelity metrics (e.g., contour overlap, angle preservation rate, topological consistency), whereas deep learning methods place greater emphasis on perceptual quality metrics (e.g., Fréchet Inception Distance [FID], Learned Perceptual Image Patch Similarity [LPIPS]). In the context of aphantasia cognitive compensation, the correctness and safety of object completion are more important than visual aesthetics—incorrect completion results may convey erroneous spatial information to the individual, paradoxically increasing cognitive load. Therefore, the object completion module in cognitive compensation strategies should prioritize ensuring the correctness of completion results in terms of geometric structure and functional semantics. In addition, cross-modal object completion is also worth exploring—for example, when visual imagery generation is difficult, tactile or auditory information completion can assist in object recognition tasks.

Displacement and scaling

Displacement and scaling are fundamental spatial transformation capabilities in mental imagery operations, referring to the ability to translate positions and adjust scales of objects in mental space. In daily life, displacement and scaling are ubiquitous: we imagine an object moving from one place to another, imagine the appearance of a building changing when viewed from different distances, or imagine an object being magnified to examine its details. For individuals with aphantasia, these seemingly simple mental operations become extremely difficult or even impossible[7]. At the computational implementation level, displacement and scaling can be realized through multiple technical approaches. Geometric representation–based methods directly apply coordinate transformations to the object's geometric representation—adding or subtracting coordinates for translation, multiplying by scaling factors for scale changes. These methods are computationally simple and highly interpretable, but require high-precision geometric representations of objects. Deep feature–based methods perform displacement and scaling operations in latent space, learning transformed object representations through trained encoder-decoder networks. These methods can handle more complex objects, but offer less control precision and interpretability compared to geometric methods. Of note, displacement and scaling may differ across modalities of mental imagery. For

visual imagery, displacement primarily involves positional changes of objects in two- or three-dimensional (2D or 3D) space; for auditory imagery, "displacement" may correspond to changes in the spatial position of sounds (i.e., mental simulation of sound source localization), while "scaling" may correspond to changes in volume or perceived distance of sounds. In cognitive compensation strategies, the displacement and scaling module should support fine-grained parametric control, allowing individuals or systems to precisely specify the magnitude and direction of transformations while ensuring that the object's structure and features remain unchanged during the transformation process.

Displacement and scaling operations can be precisely controlled computationally through a spatial transformation matrix. For any point $p_i = (x_i,\ y_i)^T$ in the representation, its translation and uniform scaling in the 2D plane can be uniformly expressed as the following homogeneous coordinate transformation (Equation 3):

$$\begin{bmatrix} x_i' \\ y_i' \\ 1 \end{bmatrix} = \begin{bmatrix} s & 0 & t_x \\ 0 & s & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \tag{3}$$

where s is the scaling factor, and t_x and t_y are the translation amounts along the X-axis and Y-axis, respectively.

The transformation matrix $T_{(t_x, t_y, s)}$ described in Equation 3 offers high interpretability and parametric controllability. Individuals or therapists can precisely specify transformation magnitudes by adjusting the three parameters s, t_x, t_y, while the system ensures that all points transform synchronously, preserving the object's relative structure. This fine-grained quantitative control is difficult for end-to-end deep learning models to directly provide, yet it is precisely in spatial training for aphantasia that individuals need to explicitly perceive the geometric meaning of each operation step to establish correct spatial cognitive mappings. Therefore, the design of the displacement and scaling module should prioritize such explicit parameterization, providing clear interfaces for subsequent personalized adjustment and feedback.

Mental rotation

Mental rotation is one of the most classic visual imagery operations in humans, first systematically investigated by Shepard and Metzler (1971) in their seminal experiments. This operation requires rotating an object in mental space and evaluating its appearance and structure from different angles. Mental rotation ability is closely related to spatial intelligence and has important application value in professional domains such as engineering drafting, surgical planning, and navigation[43]. Individuals with aphantasia may respond more slowly but more accurately than typically developing controls in mental rotation tasks, suggesting differences in cognitive strategies[20]. At the computational level, mental rotation methods vary according to representational forms[44]. Geometric topology–based methods achieve imagery rotation by rotating line segment endpoints and relational matrices: given the object's geometric relational matrix, the system computes new coordinates after rotation while maintaining the relational constraints among segments unchanged. 3D model–based methods maintain 3D models of objects and generate 2D views from arbitrary angles through projection transformations. Generative model–based methods learn mappings from source viewpoints to target viewpoints through trained conditional generation networks. In recent years, 3D-aware generative model–based methods have been able to infer the 3D structure of objects directly from a single 2D image, thereby generating rotated views from arbitrary angles.

The evaluation of mental rotation typically focuses on angle preservation rate and structural consistency. For aphantasia cognitive compensation, the precision of rotation generation is crucial: if the rotated imagery is geometrically inconsistent with true physical rotation, it may mislead the individual's spatial judgment. Therefore, the mental rotation module needs to establish rigorous geometric verification mechanisms to ensure that the topological relationships among object parts remain stable before and after rotation. In addition, factors such as the choice of rotation axes, continuity of rotation angles, and psychological realism of rotation speed all affect compensation effectiveness. From the perspective of the user experience of individuals with aphantasia, rotation operations should be smooth, continuous, and reversible, simulating the cognitive experience of genuine human mental rotation.

For mental rotation of 3D objects, its mathematical

essence is applying a 3D rotation matrix to each point. Let the rotation angles around the X, Y, and Z axes be α, β, and γ, respectively. The corresponding elementary rotation matrices are (Equation 4):

$$R_x(\alpha)=\begin{bmatrix}1 & 0 & 0\\ 0 & \cos\alpha & -\sin\alpha\\ 0 & \sin\alpha & \cos\alpha\end{bmatrix},$$
$$R_y(\beta)=\begin{bmatrix}\cos\beta & 0 & \sin\beta\\ 0 & 1 & 0\\ -\sin\beta & 0 & \cos\beta\end{bmatrix}, \quad (4)$$
$$R_z(\gamma)=\begin{bmatrix}\cos\gamma & -\sin\gamma & 0\\ \sin\gamma & \cos\gamma & 0\\ 0 & 0 & 1\end{bmatrix}$$

The composite rotation matrix is the product of the three: $R_{(\alpha,\beta,\gamma)}=R_x(\alpha)R_y(\beta)R_z(\gamma)$. For any point $p=(x, y, z)^T$, the rotated coordinates are $p'=R_{(\alpha,\beta,\gamma)}\cdot p$. The multiplication order shown here represents one rotation convention because 3D rotations are noncommutative. A future implementation should therefore specify the coordinate convention, angle units, rotation order, and direction of positive rotation to ensure reproducibility.

The rotation matrices given in Equation 4 precisely describe rigid rotational transformations of objects in 3D space, with the core property of preserving the relative distances and topological relationships among all internal points of the object. This property is critical for aphantasia compensation—because cognitive compensation strategies may rely on system-generated rotated representations to support spatial judgments, any non-rigid deformation or topological errors during rotation would severely mislead the individual's cognition. Therefore, the mental rotation module must be based on precise rotation matrix computations, supplemented by verification mechanisms ensuring geometric consistency of objects before and after rotation. Furthermore, considering that individuals with aphantasia may exhibit differential sensitivity to various rotation axes and angles, the system should also support independent adjustment of α, β, and γ as well as continuous smooth demonstrations, mimicking the dynamic process of genuine human mental rotation.

For nonvisual spatial transformations, tactile object rotation and changes in imagined sound location should be distinguished. Neither operation can be assumed to be equivalent to visual mental rotation. Their task definitions, computational models, and evaluation methods require separate validation. Their potential use in supporting individuals with aphantasia remains a direction for future research.

Cognitive compensation strategies

Cognitive compensation strategies refer to the process of integrating the above four core modules—internal representation construction, object completion, displacement and scaling, and mental rotation—into a unified assistive system. Depending on the specific needs of individuals with aphantasia, different types of compensation strategies may employ different combinations of modules. According to the objectives and approaches of compensation, cognitive compensation strategies can be classified into three types: substitutive compensation, assistive compensation, and augmentative compensation. Substitutive compensation refers to the external system completely replacing the individual's deficient mental imagery function—for example, when the individual cannot imagine what an object would look like after rotation, the system directly generates and displays the rotated image. Assistive compensation refers to the system providing partial support to help the individual better accomplish mental imagery tasks—for example, the system displays the contour skeleton of an object, guiding the individual to complete details independently. Augmentative compensation refers to the system enhancing the individual's cognitive capabilities by providing additional information channels—for example, transforming visual imagery information into tactile vibration signals, leveraging cross-modal alternative perception.

In practical applications, cognitive compensation strategy design must account for individual differences among individuals, as those with different imagery profiles may require different forms or combinations of assistance[45]. Therefore, the appropriate strategy should be determined based on individual cognitive characteristics, task requirements, preserved abilities, and user preferences rather than on the aphantasia subtype. Furthermore,

the interaction modality of compensation strategies is crucial. Intuitive, natural, and low-cognitive-load interaction designs can lighten the psychological burden on individuals using assistive systems. Future research may deepen the theory and practice of cognitive compensation strategies in the following directions: (1) developing personalized and adaptive compensation plans that dynamically adjust compensation strategies according to individuals' specific deficit characteristics and support outcomes; (2) exploring multimodal integration compensation pathways and comprehensively exploiting the compensatory potential of multiple sensory channels including vision, audition, and touch; and (3) developing objective evaluation systems for cognitive compensation effects, encompassing multidimensional quantitative assessments, such as behavioral, neurophysiological, and subjective experience measures.

Evaluation and discussion

Evaluation framework

The following evaluation framework is proposed as a structure for future research rather than an evaluation conducted in the present review. The metrics summarized below were not applied to an implemented cognitive compensation system, and no original behavioral, neurophysiological, or clinical data were collected in this study.

The evaluation environments for mental imagery generation technologies vary considerably depending on the methodological approach. Geometric topology–based reasoning methods typically run on standard computing platforms, primarily relying on central processing units (CPUs) or general-purpose graphics processing units (GPUs) for geometric computation and constraint solving, with relatively low hardware requirements suitable for deployment on portable devices. In many existing studies, deep learning–based methods require high-performance GPUs (e.g., NVIDIA V100/A100) for model training, although hardware demands during inference can be reduced through model compression and quantization techniques[46]. Neural decoding–based methods impose the most stringent requirements on experimental environments, necessitating dedicated neuroimaging equipment such as fMRI, high-density EEG, or MEG, with experimental procedures requiring electromagnetic interference shielding and careful control of participant head movements[47]. Regarding software platforms, commonly used environments include the PyTorch and TensorFlow deep learning frameworks, the OpenCV computer vision library, and the Computational Geometry Algorithms Library (CGAL) and Eigen for geometric computation. Psychophysical experiments typically employ specialized tools such as Psychtoolbox or PsychoPy for precise visual stimulus presentation and behavioral data acquisition. At the data level, commonly used datasets in visual imagery generation experiments include ImageNet (natural image classification), ShapeNet (3D models), CelebA (facial images), and synthetic datasets specifically designed for mental rotation and occlusion completion tasks[48].

Evaluation methods

Neural decoding–based methods

The evaluation metrics for mental imagery generation technologies are diverse, varying according to application scenarios and technical approaches. This review categorizes existing evaluation metrics into the following five categories:

(1) **Geometric fidelity metrics**: these include Intersection over Union (IoU), Structural Similarity Index Measure (SSIM), Root Mean Square Error (RMSE), and Hausdorff Distance, among others, used to measure the geometric similarity between generated imagery and real objects. These metrics are particularly important in geometry-based reasoning methods, as they can precisely reflect the geometric correctness of completion or rotation results[49].

Specifically, IoU is the most intuitive region overlap measure. Let the predicted contour pixel set be P and the ground-truth contour pixel set be G. IoU is defined as the ratio of the intersection to the union of the two sets (Equation 5):

$$\mathrm{IoU}(P,G)=\frac{|P\cap G|}{|P\cup G|} \tag{5}$$

IoU ranges between 0 and 1, with values closer to 1

indicating higher overlap between the generated and ground-truth contours. This metric is computationally simple and intuitive, particularly suitable for evaluating binarized contour completion results, although it is sensitive to minor misalignments.

The SSIM evaluates image quality from a global structural perspective. For two image patches x and y to be compared, SSIM integrates three components: luminance, contrast, and structure (Equation 6):

$$\mathrm{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (6)$$

where μ_x, μ_y are the pixel means of the two image patches, σ_x, σ_y are the standard deviations, σ_{xy} is the covariance, and c_1, c_2 are small constants to prevent division by zero. SSIM better aligns with human visual system perception of image quality, but it reflects global similarity rather than precise geometric deviations.

RMSE is the most fundamental pixel-wise error metric. For N corresponding pixels, it is defined as the square root of the mean squared difference between predicted values $\hat{y}_i$ and ground-truth values y_i (Equation 7):

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \quad (7)$$

RMSE assigns higher weight to larger errors, thus sensitively capturing significant geometric deviations, although its ability to distinguish between global shifts and local deformations is limited.

The Hausdorff Distance is a powerful tool for evaluating worst-case boundary errors. For two point sets P and G, the bidirectional Hausdorff distance is defined as the larger of the maximum minimum distances from either point set to the other (Equation 8):

$$d_H(P, G) = \max\left\{\sup_{p\in P}\inf_{g\in G}\|p-g\|, \sup_{g\in G}\inf_{p\in P}\|g-p\|\right\} \quad (8)$$

where $\sup_{p\in P}\inf_{g\in G}\|p\text{-}g\|$ means: for each point p in set P, first compute its minimum distance to all points in set G (i.e., find the closest point in G to p), and then take the maximum of these minimum distances. This term intuitively characterizes "how far the point in P farthest from G is from G"—a large value indicates that some regions of the generated contour deviate severely from the ground truth. The second term, $\sup_{g\in G}\inf_{p\in P}\|g\text{-}p\|$, is symmetric and measures "how far the point in G farthest from P is from P." The final bidirectional Hausdorff distance takes the larger of these two terms, reflecting the maximum degree of mismatch between two contours—i.e., the distance between the farthest corresponding points. This metric is particularly sensitive to local distortions or discontinuities and offers unique advantages in detecting local geometric distortions in generated imagery. In aphantasia cognitive compensation applications, geometric fidelity represents an important criterion for ensuring that generated spatial information remains structurally reliable. The above four formulas characterize the geometric quality of generated imagery from four complementary mathematical perspectives—overlap ratio, structural similarity, pointwise error, and maximum deviation—and should be used in combination rather than in isolation.

(2) **Perceptual quality metrics**: these include FID, LPIPS, and Natural Image Quality Evaluator (NIQE), used to assess the visual naturalness and perceptual realism of generated imagery. These metrics are widely applied in deep learning–based methods[50].

FID is one of the most commonly used metrics for evaluating generated image quality. It extracts high-level semantic features using the Inception network, models the distributions of generated and real image sets as multivariate Gaussian distributions, and computes the Fréchet distance between the two distributions (Equation 9):

$$\mathrm{FID} = \|\mu_r - \mu_g\|^2 + \mathrm{Tr}\left(\Sigma_r + \Sigma_g - 2(\Sigma_r\Sigma_g)^{1/2}\right) \quad (9)$$

where μ_r, Σ_r and μ_g, Σ_g are the means and covariance matrices of the real and generated image sets in the Inception feature space, respectively. Lower FID values indicate greater similarity between the generated and real image distributions. Compared to earlier metrics (e.g., Inception Score), FID more sensitively captures both diversity and fidelity of generated images, although its computation depends on a pre-trained Inception network,

which may introduce bias when evaluating images outside the training distribution.

LPIPS more closely approximates human perceptual similarity judgments. It extracts multi-level feature maps through a deep network and computes weighted distances between paired features (Equation 10):

$$\mathrm{LPIPS}(x, y) = \sum_{l} \frac{1}{H_l W_l} \sum_{h, w} \left\| w_l \odot \left(\hat{y}^l_{hw} - y^l_{0hw} \right) \right\|_2^2 \quad (10)$$

where l is the network layer index, $\hat{y}^l$ are normalized feature vectors, and w_l are learnable weight vectors. LPIPS has become a standard evaluation tool in generation tasks such as inpainting and super-resolution. In deep learning–dominated visual imagery generation research, perceptual quality metrics (especially FID and LPIPS) have become the de facto standard for visual quality evaluation. However, they emphasize visual naturalness rather than spatial accuracy and should be used as supplementary metrics in aphantasia cognitive compensation scenarios.

(3) **Cognitive relevance metrics**: these include behavioral metrics such as Vividness of Visual Imagery Questionnaire (VVIQ) scores, mental rotation reaction times, and task completion accuracy, as well as neurophysiological metrics such as fMRI brain region activation strength and EEG event-related potentials (e.g., rotation-related negativity [RRN]). These metrics directly assess the compensatory effects of generated imagery on individuals' cognitive functions[51].

Mental rotation reaction time is the most classic cognitive relevance metric at the behavioral level. In mental rotation experiments, reaction time RT typically exhibits a linear relationship with rotation angle θ, as described by a linear mental-rotation model (Equation 11):

$$RT(\theta) = RT_0 + k \cdot \theta \quad (11)$$

where RT_0 is the baseline reaction time (corresponding to 0° rotation), k is the time-per-angle coefficient, and θ is the rotation angle. In a future evaluation study, changes in k could be compared across baseline, system-assisted, and appropriate control conditions. However, changes in reaction time alone would not establish compensation effectiveness and should be interpreted together with accuracy, strategy use, cognitive load, and subjective experience. Task completion accuracy is a more direct measure of compensation effectiveness, defined as the ratio of correct responses $N_{correct}$ to total trials N_{total} (Equation 12):

$$\mathrm{Accuracy} = \frac{N_{correct}}{N_{total}} \times 100\% \quad (12)$$

Neurophysiological metrics within the cognitive relevance category (e.g., fMRI brain region activation strength, EEG event-related potentials) also possess well-defined quantitative characteristics. Cognitive relevance metrics are included because technical performance alone cannot establish behavioral or clinical benefit. Future studies should examine whether improvements in technical metrics are accompanied by measurable changes in task performance, cognitive load, daily functioning, or user experience.

(4) **Topological consistency metrics**: these include adjacency preservation rate, relative position error, and topological structural stability, used to evaluate the correctness of generated imagery in terms of topological relationships. For imagery tasks involving spatial relation reasoning, topological consistency is more critical than pixel-level geometric fidelity[52].

The adjacency preservation rate is a fundamental metric for topological consistency. Let the set of all adjacent element pairs (e.g., between line segments, between regions) in the object be E_{true}, and the set of element pairs that maintain the same adjacency relationship in the generated imagery be E_{pred}. The adjacency preservation rate is defined as (Equation 13):

$$\mathrm{Adjacency\ Preservation} = \frac{|E_{pred} \cap E_{true}|}{|E_{true}|} \quad (13)$$

This metric characterizes the extent to which the generated imagery preserves the object's connection structure and is crucial for evaluating topological correctness in mental rotation and occlusion completion tasks. Relative position error focuses on whether spatial order among elements is correctly maintained, typically measured by Kendall's rank correlation coefficient or similar

methods. Topological structural stability further requires that the topological features of generated imagery do not change drastically under small perturbations (e.g., minor changes in input information), mathematically defined through Lipschitz continuity constraints in the generation process. In the daily cognition of individuals with aphantasia, the usefulness of spatial information depends first on the correctness of structural relationships and only secondarily on the completeness of visual details—for example, in mental rotation tasks, if connections among object parts are incorrectly altered, even high overall visual quality would convey distorted spatial information to the individual. Therefore, topological consistency metrics should be regarded as the baseline standard in the aphantasia cognitive compensation evaluation framework.

(5) **Task performance metrics**: these include completion time, error rate, and user satisfaction in downstream tasks such as occlusion completion, mental rotation, and scene reconstruction. These metrics could provide evidence concerning the usability and task-level performance of future cognitive compensation systems. In actual research, multiple categories of metrics are typically used in combination to comprehensively evaluate the performance of mental imagery generation technologies.

Task completion time reflects the system's real-time performance and cognitive efficiency, fundamentally expressed as the difference between the end time t_{end} and start time t_{start} (Equation 14):

$$T_{task} = t_{end} - t_{start} \tag{14}$$

For tasks requiring repeated execution, the average $\overline{T}_{task} = \frac{1}{n}\sum_{i=1}^{n} T_i$ is typically taken to eliminate random fluctuations. Error rate is the complement of accuracy, i.e., Error Rate = 100% - Accuracy. User satisfaction is a subjective evaluation metric, typically collected through standardized questionnaires (e.g., Likert scales) and quantitatively converted through statistical processing. The greatest value of task performance metrics lies in their comprehensiveness—they ultimately aggregate the strengths and weaknesses of the previous four metric categories into the definitive question of "whether the system is genuinely useful in real-world use," and may help determine whether a future system is sufficiently usable and reliable for further real-world evaluation.

Summary of experimental results

Regarding the main experimental results of mental imagery generation technologies, this review summarizes from two dimensions: methodological approaches and cognitive compensation tasks. For geometric topology–based reasoning methods, experimental results indicate that these approaches maintain high structural accuracy in occlusion completion tasks, with performance remaining robust at lower occlusion levels. They also preserve geometric relationships effectively during mental rotation tasks. Deep learning–based reconstruction methods have demonstrated strong perceptual quality in standard image inpainting tasks, with generative models producing visually realistic completions[26,28]. However, in tasks requiring precise spatial reasoning, such as mental rotation and pose estimation, deep learning methods generally underperform specialized geometric reasoning approaches in terms of geometric fidelity. For neural decoding methods, fMRI-based image reconstruction has achieved promising results on natural images, although reconstruction resolution and real-time performance remain limited[29,31]. EEG-based methods offer high temporal resolution but face constraints in spatial resolution and reconstruction quality[47].

These findings suggest that geometrically constrained approaches may be better suited for applications requiring precise structural manipulation, whereas deep learning methods are more appropriate when perceptual naturalness is the primary objective.

Comparative analysis

A comprehensive comparison of various approaches reveals that different technical pathways have respective strengths and weaknesses in different aspects of mental imagery generation. In terms of internal representation construction capability, geometric topology methods generally perform well with structured objects, deep learning methods demonstrate advantages in handling unstructured scenes, and neural decoding methods

provide closer links to biological neural representations. In object completion capability, geometric methods dominate in geometric fidelity, while deep learning methods dominate in visual quality. In displacement and scaling capability, geometric methods offer finer control precision, whereas deep learning methods provide greater flexibility. In mental rotation capability, 3D model–based methods can directly generate views from arbitrary angles, while 2D image–based methods require additional learning of viewpoint transformations[53]. In the specific application scenario of aphantasia cognitive compensation, the key performance metrics are not visual realism but geometric correctness and interpretability. Individuals with aphantasia need to intuitively understand how the system-generated imagery supports spatial reasoning, which is crucial for building trust and enabling effective use of assistive systems. Therefore, the technical pathway suitable for aphantasia cognitive compensation should be a hybrid approach that integrates the precision of geometric reasoning with the perceptual quality of deep learning.

Table 1 compares the principles, main advantages, limitations, and potential relevance of representative approaches to cognitive compensation. The proposed relevance to aphantasia should be interpreted as conceptual unless direct evidence from individuals with aphantasia is available.

The comparison indicates that these approaches offer

Table 1: Comparative analysis of representative mental imagery generation approaches for cognitive compensation in aphantasia.

Approach	Core representation/ principle	Main advantages	Main limitations	Potential relevance to cognitive compensation
Geometric topology–based reasoning	Explicit geometric representations based on contours, line segments, relational matrices, and topological constraints	High interpretability; precise structural control; strong geometric consistency; does not rely heavily on large-scale training datasets	Limited capacity for complex, naturalistic, or highly unstructured scenes; performance depends on reliable geometric feature extraction	May support transparent external spatial reasoning, object completion, and mental-rotation tasks in structured environments
Deep learning–based reconstruction	Implicit feature representations learned from large-scale datasets using generative models such as GANs, VAEs, and diffusion models	Strong image generation capability; high perceptual realism; effective handling of complex visual information	Requires substantial training data and computational resources; limited interpretability; possible hallucination and structural inconsistency	May generate external visual representations from verbal, partial visual, or multimodal inputs, but its effectiveness for individuals with aphantasia requires direct evaluation
Neural decoding–based methods	Mapping neural activity patterns from fMRI, EEG, or MEG signals to visual imagery content through machine learning models	Provides direct links between neural activity and imagery representation; enables objective investigation of neural mechanisms	Expensive or restrictive acquisition equipment; limited portability; participant-specific variability; reconstruction accuracy remains limited	May help characterize individual neural profiles and inform future personalized assistance, but currently lacks evidence as a practical compensation intervention
Multimodal mental imagery methods	Construction of shared representations across multiple sensory modalities, including visual, auditory, and tactile information	Enables cross-modal compensation; integrates complementary sensory channels; supports personalized strategies	Limited evidence in aphantasia; difficult multimodal alignment; increased system and evaluation complexity	May support alternative nonvisual strategies when voluntary visual imagery is absent or reduced
External/verbal cognitive strategies	Use of verbal coding, semantic organization, external diagrams, physical manipulation, or stepwise reasoning	Low cost; accessible; interpretable; consistent with some reported compensatory strategies used by individuals with aphantasia	May increase cognitive load; effectiveness varies substantially across individuals and tasks; does not restore visual imagery	May provide immediately usable substitutive or assistive support without assuming restoration of internal imagery

Abbreviations: GANs, generative adversarial networks; VAEs, variational autoencoders; fMRI, functional magnetic resonance imaging; EEG, electroencephalography; MEG, magnetoencephalography.

complementary capabilities rather than a single optimal solution. Geometric reasoning provides transparent structural representations, whereas deep generative methods support flexible processing of complex information. Neural decoding methods may contribute to the investigation of individual neural mechanisms, while multimodal and external cognitive strategies may provide alternative channels when voluntary visual imagery is absent or reduced. However, most of these potential applications have not been directly evaluated as cognitive compensation interventions for individuals with aphantasia. Therefore, future studies should test whether the technical capabilities summarized in Table 1 translate into measurable behavioral or functional benefits.

Discussion and proposed conceptual framework

Building on the preceding review, the core purpose of this conceptual framework (Figure 1) is to organize existing computational approaches to mental imagery generation and illustrate how external computational systems may support individuals with aphantasia in performing imagery-related cognitive tasks. The framework is

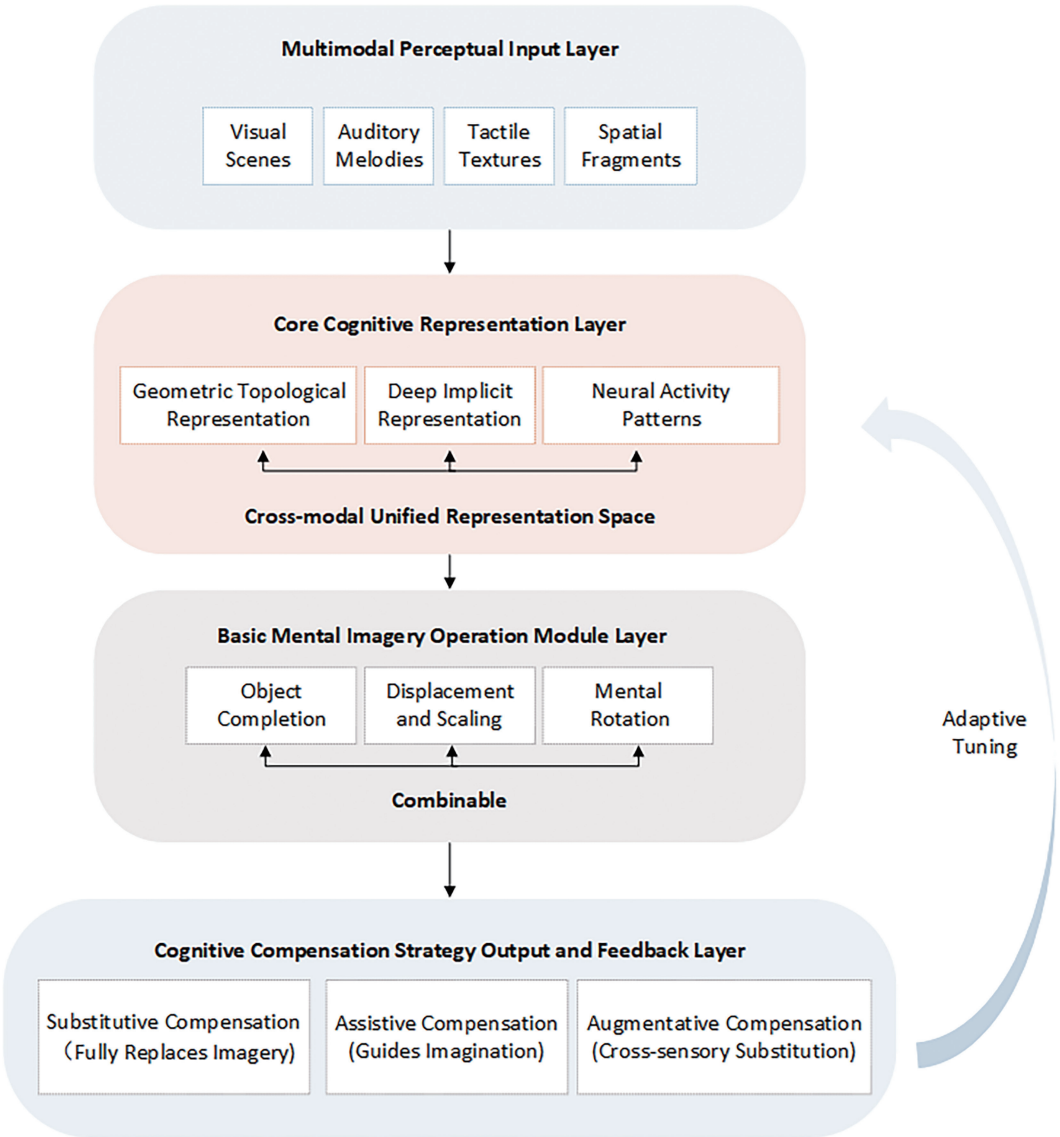


Figure 1: Conceptual framework of a multimodal cognitive compensation system for aphantasia.

intended as a conceptual synthesis rather than an experimentally or clinically validated intervention system.

The framework consists of four layers: the multimodal perception input layer receives available visual, auditory, linguistic, and other contextual information; the core representation layer converts these inputs into geometric, latent, semantic, or neural-signal-derived computational representations; the imagery operation layer supports fundamental functions including internal representation construction, object completion, displacement and scaling, and mental rotation; and the cognitive compensation strategy layer provides substitutive, assistive, or augmentative forms of support according to individual imagery characteristics, preserved cognitive strategies, task requirements, and user preferences.

Within this framework, the core representation layer should be understood as a functional interface rather than a digital replica of neural representation. Sensory, linguistic, or neural inputs may be encoded into manipulable computational representations, upon which completion, transformation, and cross-modal integration can be performed. The resulting information is then presented through external visual, auditory, verbal, or tactile outputs, while behavioral, subjective, or neural feedback may be used to refine subsequent outputs. Accordingly, the arrows between the layers represent functional information flow and iterative adjustment rather than direct equivalence between computational representations and patterns of brain activity.

In this framework, the feedback loop represents a proposed iterative process associated with the cognitive compensation strategy layer. After an imagery-related task, task accuracy, response time, interaction behavior, self-reported cognitive load, perceived usefulness, and modality preference may be collected as feedback signals. Based on these signals, a future system could adjust the amount of guidance, the output modality, the presentation format, or the type of compensation strategy. For example, when partial visual cues do not adequately support task performance, the system could provide a complete external representation or introduce verbal, auditory, or tactile information. This feedback process remains a conceptual design proposal requiring future empirical evaluation.

Substitutive, assistive, and augmentative forms of compensation represent different functional forms of support rather than strategies assigned to fixed diagnostic subtypes. Their selection should depend on individual imagery characteristics, preserved cognitive abilities, habitual strategies, task requirements, cognitive load, and user preferences. At present, there is insufficient evidence to establish a clinically validated one-to-one mapping between particular aphantasia profiles and these compensation strategies. The three forms of compensation are described in greater detail in Section “Cognitive compensation strategies.” Within the imagery operation layer, four core functional modules are considered: internal representation construction, object completion, displacement and scaling, and mental rotation.

These four modules constitute the fundamental functional units of cognitive compensation for mental imagery. They can operate either independently or in combination through serial or parallel configurations to collaboratively accomplish more complex cognitive tasks. Unlike existing literature that focuses on specific technical implementations (e.g., geometric relational matrices or deep generative models), the present framework describes the methodology of mental imagery operations at a higher level of abstraction. That is, any mental imagery generation technique—whether based on geometric reasoning, deep learning, or neural decoding—can be mapped onto the four core modules of this framework. For example, the “geometric relational matrix” is essentially just one specific encoding approach for implementing “internal representation” of objects; “line-segment contour completion” is merely one particular algorithm for achieving “object completion.” The generality of this framework enables the integration of different technical pathways into a systematic methodological perspective for studying cognitive compensation.

Although the proposed framework integrates the reviewed approaches at a conceptual level, its practical implementation remains constrained by limitations in the existing evidence and technologies.

Mental imagery generation technologies may have potential relevance to cognitive compensation for aphantasia across multiple domains, although direct evidence

of their effectiveness remains limited. In the field of clinical rehabilitation, assistive systems based on multimodal mental imagery generation technologies could potentially support individuals with aphantasia in performing imagery-related cognitive tasks such as scene construction, spatial navigation, and object recognition. However, whether such systems produce measurable improvements in daily functioning, quality of life, or social participation has not yet been established. In education, such technologies may provide external representations or alternative information formats for students with aphantasia in learning activities that involve visual imagery, such as geometric reasoning, understanding molecular structures, and map comprehension. Their educational effectiveness requires direct evaluation rather than being assumed from their technical capabilities. In professional support, imagery compensation systems could potentially provide external spatial information for tasks such as architectural design, engineering drafting, and surgical planning. Nevertheless, there is currently insufficient evidence to conclude that these systems improve professional performance or expand career opportunities for individuals with aphantasia. At the technological development level, advances in brain–computer interfaces, augmented reality (AR), and explainable artificial intelligence (XAI) may contribute to the future development of portable, real-time, and personalized assistive systems. These possibilities should be interpreted as future research directions rather than as established clinical or practical applications.

The main limitations of current cognitive compensation strategies include the following aspects: First, most existing technologies are confined to the single modality of visual imagery, with limited research on compensation for multimodal mental imagery including auditory and tactile imagery. However, the impact of aphantasia extends well beyond the visual domain—individuals may also face difficulties in musical imagination, speech imagination, tactile imagination, and other areas, necessitating multimodal compensation solutions. Second, current systems are mostly validated in controlled experimental environments, and their robustness and effectiveness in real-world, dynamically changing scenarios have not been fully verified. Third, insufficient personalization is another prominent issue. Individuals with aphantasia exhibit significant differences in deficit severity, compensation needs, and adaptive capacity, yet most existing systems adopt a "one-size-fits-all" compensation strategy, lacking the ability to dynamically adjust according to individual characteristics. Fourth, human-computer interaction design has not received adequate attention. If compensation systems are complex to operate or provide unintuitive feedback, individuals may find it difficult to use them effectively in daily life scenarios. Fifth, current evaluation frameworks are dominated by technical metrics, with insufficient attention to clinically relevant indicators such as individual subjective experience, improvement in daily functioning, and quality of life enhancement.

Practical translation may also be constrained by medical-device regulatory requirements and by the cost, portability, battery life, comfort, and maintenance requirements of wearable hardware. Before clinical deployment, future systems would require evidence of safety, effectiveness, usability, and clearly defined responsibilities for system monitoring and decision-making. Because individuals with aphantasia differ in imagery characteristics, preserved cognitive strategies, and support needs, future algorithms should allow individualized calibration of assistance level, output modality, and feedback format.

Furthermore, from an ethical perspective, cognitive compensation technologies raise a series of thought-provoking questions: Would compensation technologies weaken individuals' own cognitive development? What impact would long-term reliance on external assistive tools have on individuals' brain plasticity? How can a balance be achieved between compensation effectiveness and ethical boundaries? These questions warrant greater attention and discussion in future research. Additional ethical concerns arise when assistive systems acquire or process neural signals, because such data may contain sensitive information concerning users' cognitive states or health-related characteristics. Future development should therefore consider informed consent, data minimization, secure storage and transmission, access control, transparency regarding secondary data use, and users' right to withdraw their data. Potential risks associated with inaccurate system output, algorithmic bias, and unequal access resulting from hardware cost should also be evaluated. Overall, cognitive compensation strategies show great potential in supporting individuals with aphantasia, yet a significant gap remains between

basic research and clinical translation. Future research should focus on key directions including multimodal integration, personalized adaptation, ecosystem integration, and long-term effectiveness evaluation. Future studies should establish standardized evaluation protocols combining behavioral measurements, neurophysiological indicators, and user-centered assessments to systematically evaluate the effectiveness of cognitive compensation strategies. Ethical oversight, data-governance requirements, safety monitoring, and longitudinal assessment should also be incorporated before clinical deployment is considered.

Conclusion

Based on the perspective that aphantasia represents a form of cognitive diversity rather than a pathological condition, this overview provides a structured narrative review of cognitive compensation strategies grounded in mental imagery generation technologies. Drawing on a comprehensive review and analysis of the neurocognitive basis of aphantasia, the theoretical frameworks of mental imagery, and existing technical approaches, this study offers the following main conclusions. First, aphantasia should be understood as a form of cognitive phenotypic diversity rather than a disease, with its hallmark being the inability to voluntarily generate and manipulate visual and other modal mental imagery. This conceptual shift has important implications for the development of supportive strategies. Notably, future approaches should not aim to "cure" aphantasia but to explore compensatory solutions that accommodate individual cognitive characteristics. The coordinated functioning of four fundamental operational modules—internal representation construction, object completion, displacement and scaling, and mental rotation—constitutes the core of cognitive compensation strategies, and any mental imagery generation technique can be mapped onto this general framework.

Methodologically, this review broadens the scope of previous studies, which have typically centered on specific technical implementations, and proposes a methodological system for cognitive compensation at a higher level of abstraction. Instead of depending on any particular algorithm or model, such as geometric relational matrices or deep neural networks, this system focuses on the functional essence and general methodology of mental imagery operations. The proposed framework integrates the benefits from different technical pathways while guiding the systematic development of future cognitive assistive systems for aphantasia. From a technical perspective, this review summarizes research progress along three major technical approaches: geometric topology–based reasoning, deep learning–based reconstruction, and neural decoding. It analyzes the strengths and limitations of each from the perspective of cognitive compensation, and suggests that hybrid approaches integrating the precision of geometric reasoning with the perceptual quality of deep learning may offer a promising technical direction for aphantasia cognitive compensation.

It is important not to regard computational representations as direct replicas of neural representations. The functional connections that may be established through encoding, decoding, cross-modal alignment, and feedback highlight their relevance. Because most neural reconstruction evidence is obtained from healthy participants, its direct relevance to aphantasia remains to be validated.

At the same time, this review acknowledges the challenges and shortcomings that existing cognitive compensation strategies must still address. To transition from providing a description of aphantasia as a phenomenon to technological interventions, sustained efforts in multiple directions are warranted, including multimodal compensation, personalized adaptation, ecological validity validation, and interaction design optimization. Future cognitive compensation systems should: develop multimodal, cross-sensory compensation solutions that transcend the restrictions imposed by a single visual modality; introduce personalized and adaptive compensation mechanisms to meet the diverse needs of individuals; construct closed-loop systems that integrate perception, reasoning, and feedback, thus achieving real-time and accurate cognitive assistance; and improve objective, multidimensional evaluation systems to comprehensively measure compensation effectiveness. In addition, future research should prioritize investigating the multimodal nature of aphantasia. That is, it should focus on deficits not only in visual imagery but also in other modalities such as auditory imagery. More broadly, research on

mental imagery generation technologies and their cognitive compensation strategies not only benefits individuals with aphantasia but also provides a unique lens for understanding the flexibility and diversity of human cognition. As the concept of cognitive variation suggests, the diversity of human cognitive abilities is itself a valuable ecological asset, and the power of technology lies in enabling individuals to fully realize their unique cognitive potential.

Author contributions

Luping Wang and Jiayi Liu contributed equally to this work. Luping Wang and Jiayi Liu conceived the study, conducted the literature review, and drafted the original manuscript. Lijuan Chen supervised the entire project, provided critical revisions to the manuscript, and is responsible for the overall scientific integrity of the work. All authors reviewed, edited, and approved the final version of the manuscript and agree to its publication.

Acknowledgments

This work was supported by the NSFC Project (Grant No. 62003212).

We thank LetPub for its linguistic assistance during the preparation of this manuscript.

Data availability statement

No datasets were generated or analyzed for this work.

Ethics statement

Not applicable.

AI statement

The authors used GPT-5.6 (OpenAI, online version) as a generative artificial intelligence (AI)-assisted language-editing tool to improve grammar, clarity, and readability. The AI tool was not used for data analysis, data generation, image processing, or scientific content creation. All outputs were carefully reviewed and verified by the authors, who take full responsibility for the final content.

Conflicts of interest

Luping Wang serves as a member of the Editorial Board of this journal. He was not involved in the editorial review or decision-making process for this manuscript. All editorial decisions were made independently by other members of the Editorial Board who have no conflicts of interest.

References

1. Zeman AZJ, Della Sala S, Torrens LA, Gountouna VE, McGonigle DJ, Logie RH. Loss of imagery phenomenology with intact visuo-spatial task performance: a case of "blind imagination". *Neuropsychologia*. 2010;48(1):145-155.
2. Pearson J, Naselaris T, Holmes EA, Kosslyn SM. Mental imagery: functional mechanisms and clinical applications. *Trends Cogn Sci*. 2015;19(10):590-602.
3. Kosslyn SM, Thompson WL, Ganis G. *The Case for Mental Imagery*. New York, NY: Oxford University Press; 2006.
4. Ganis G, Thompson WL, Kosslyn SM. Brain areas underlying visual mental imagery and visual perception: an fMRI study. *Brain Res Cogn Brain Res*. 2004;20(2):226-241.
5. Zeman A, Dewar M, Della Sala S. Lives without imagery—Congenital aphantasia. *Cortex*. 2015;73:378-380.
6. Keogh R, Pearson J. Revisiting the blind mind: Still no evidence for sensory visual imagery in individuals with aphantasia. *Neurosci Res*. 2024;201:27-30.
7. Zeman A, Milton F, Della Sala S, et al. Phantasia–The psychological significance of lifelong visual imagery vividness extremes. *Cortex*. 2020;130:426-440.
8. Monzel M, Pickering JW, Condon DM, Beran MJ, Ebeyer T. Reimagining the vividness of visual imagery questionnaire as a single item screener for aphantasia. *Conscious Cogn*. 2026;142:104061.
9. Bartolomeo P. The relationship between visual perception and visual mental imagery: a reappraisal of the neuropsychological evidence. *Cortex*. 2002;38(3):357-378.
10. Knight KF, Milton F, Zeman A. Aphantasia and visual working memory: No direct evidence of impaired visual working memory in aphantasics, either in behavioral performance or the accuracy of a multivoxel pattern classifier. *Neuropsychologia*. 2026;226:109430.
11. Zeman A. Aphantasia and hyperphantasia: Exploring imagery vividness extremes. *Trends Cogn Sci*. 2024;28(5):467-480.
12. Siena MJ, Simons JS. Metacognitive awareness and the subjective experience of remembering in aphantasia. *J Cogn Neurosci*. 2024;36(8):1578-1598.
13. Kosslyn SM. *Image and brain: The resolution of the imagery debate*. Cambridge, MA: MIT Press; 1994.
14. O'Craven KM, Kanwisher N. Mental imagery of faces

and places activates corresponding stimulus-specific brain regions. *J Cogn Neurosci*. 2000;12(6):1013-1023.
15. Keogh R, Pearson J. The blind mind: No sensory visual imagery in aphantasia. *Cortex*. 2018;105:53-60.
16. Kutsche J, Howard C, Castro Palacin A, et al. Lesions causing aphantasia are connected to the fusiform imagery node. *Cortex*. 2026;197:1-12.
17. Noble C, Taylor NL, Milton F, et al. Seeing through the static: Reduced imagery vividness in aphantasia is associated with impaired temporal lobe signal complexity. *Neuropsychologia*. 2026;221:109322.
18. Bartolomeo P. Mapping the imageless mind: towards a taxonomy of aphantasia. *Neuropsychologia*. 2025;219: 109276.
19. Scholz CO, Monzel M, Kvamme TL, Liu J, Silvanto J. An integration model of mental imagery and aphantasia: Conceptual framework, neuromechanistic pathways, and clinical implications. *Neuropsychologia*. 2026;225: 109401.
20. Kay L, Keogh R, Pearson J. Slower but more accurate mental rotation performance in aphantasia linked to differences in cognitive strategies. *Conscious Cogn*. 2024;121:103694.
21. Delem M, Turkben S, Cavalli E, Cousineau D, Plancher G. Unsupervised clustering reveals spatial and verbal cognitive profiles in aphantasia and typical imagery. *Neuropsychologia*. 2025;219:109279.
22. Nanay B. Multimodal mental imagery. *Cortex*. 2018;105: 125-134.
23. Hashim S, Pulcini C, Jansari A, Küssner MB, Omigie D. The experience of music in aphantasia: Emotion, reward, and everyday functions. *Music Sci*. 2024;7:1-18.
24. Hayes SJ, Miles GE, Evans SA. "Unseen strategies": What can the experience of aphantasia teach us about cognitive strategies in memory? *New Ideas Psychol*. 2026;80:101215.
25. Marr D. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco, CA: Freeman; 1982.
26. Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks. *Commun ACM*. 2020;63(11): 139-144.
27. Kingma DP, Welling M. Auto-encoding variational Bayes. arXiv. Preprint posted online December 20, 2013. arXiv: 1312.6114.
28. Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models. In: *Proceedings of the 34th International Conference on Neural Information Processing Systems (NeurIPS)*; 2020;33:6840-6851.
29. Shen G, Horikawa T, Majima K, Kamitani Y. Deep image reconstruction from human brain activity. *PLoS Comput Biol*. 2019;15(1):e1006633.
30. Pan H, Li Z, Fu Y, et al. Reconstructing visual stimulus images from EEG signals based on deep visual representation model. *IEEE Trans Human Mach Syst*. 2024; 54(6):711-722.
31. Goense J, Nothnagel N. fMRI at ultra-high field: from acquisition to interpretation. *Neuron*. 2026;114(17):3133-3155.
32. Qiao K, Chen J, Wang L, Zhang C, Tong L, Yan B. BigGAN-based Bayesian reconstruction of natural images from human brain activity. *Neuroscience*. 2020;444:92-105.
33. Horikawa T, Kamitani Y. Generic decoding of seen and imagined objects using hierarchical visual features. *Nat Commun*. 2017;8:15037.
34. Naselaris T, Olman CA, Stansbury DE, Ugurbil K, Gallant JL. A voxel-wise encoding model for early visual areas decodes mental images of remembered scenes. *Neuroimage*. 2015;105:215-228.
35. Lee SH, Lee M, Lee SW. Neural decoding of imagined speech and visual imagery as intuitive paradigms for BCI communication. *IEEE Trans Neural Syst Rehabil Eng*. 2020;28(12):2647-2659.
36. Lambert AJ, Schelp ZM, Quigley-Tump G, Tan V, Purdy S, Sibley C. Anauralia and aphantasia: Prevalence and distinct associations with personality, well-being and self-regulation in a large, representative sample. *Neuropsychologia*. 2026; 230:109527.
37. Halpern AR. Differences in auditory imagery self-report predict neural and behavioral outcomes. *Psychomusicology: Music, Mind, and Brain*. 2015;25(1):37-47.
38. Briscoe RE. Mental imagery and the varieties of amodal perception. *Pacific Philosophical Quarterly*. 2011;92(2): 153-173.
39. Vetter P, Smith FW, Muckli L. Decoding sound and imagery content in early visual cortex. *Curr Biol*. 2014;24(11): 1256-1262.
40. Delhaye E, Fritz P, Martial C, Bastin C. The role of visual imagery and verbal coding in short-term memory conjunctive binding: evidence from aphantasia. *Conscious Cogn*. 2026;141:104036.
41. Wicken M, Keogh R, Pearson J. The critical role of mental imagery in human emotion: insights from fear-based imagery and aphantasia. *Proc Biol Sci*. 2021; 288(1946):20210267.
42. Tooming U, Jakapi R. Aphantasia as a challenge for Humean abstraction. *Neuropsychologia*. 2026;227:109465.
43. Shepard RN, Metzler J. Mental rotation of three-dimensional objects. *Science*. 1971;171(3972):701-703.
44. Searle JA, Hamm JP. Asymmetric response time functions during left-/right-facing discriminations of rotated objects: the short and the long of it. *Mem Cognit*. 2016;44(1):124-142.
45. Borst G, Kosslyn SM. Individual differences in spatial mental imagery. *Q J Exp Psychol (Hove)*. 2010; 63(10):2031-2050.
46. Howard J, Ruder S. Universal language model fine-tuning for text classification. In: *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*

(ACL); 2018;1:328-339.
47. Yadav H, Maini S. Decoding brain signals: a comprehensive review of EEG-Based BCI paradigms, signal processing and applications. *Comput Biol Med*. 2025; 196(Pt C):110937.
48. Chang AX, Funkhouser TA, Guibas LJ, et al. ShapeNet: An information-rich 3D model repository. arXiv. Preprint posted online December 9, 2015;arXiv:1512.03012.
49. Anderson AJ, Bruni E, Lopopolo A, Poesio M, Baroni M. Reading visually embodied meaning from the brain: visually grounded computational models decode visual-object mental imagery induced by written text. *Neuroimage*. 2015;120:309-322.
50. Heusel M, Ramsauer H, Unterthiner T, Nessler B, Hochreiter S. GANs trained by a two time-scale update rule converge to a local Nash equilibrium. In: *Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS)*; 2017;30:6626-6637.
51. Milton F, Fulford J, Dance C, et al. Behavioral and neural signatures of visual imagery vividness extremes: aphantasia versus hyperphantasia. *Cereb Cortex Commun*. 2021; 2(2):tgab035.
52. Kvamme TL, Lumaca M, Bajada CJ, et al. Neural network topologies supporting individual variations in vividness of visual imagery. *Neuroimage*. 2025;321:121520.
53. Tiwari A, Pachori RB, Sanjram PK. Dimensionality and angular disparity influence mental rotation in computer gaming. *Comput Mater Contin*. 2022;72(1):887-905.