Cognitive Neuropsychology

ISSN: 0264-3294 (Print) 1464-0627 (Online) Journal homepage: www.tandfonline.com/journals/pcgn20

Impact of imagery deficit on word-based object colour retrieval: Evidence from congenital aphantasia

Zhenjiang Cui, Xiangqi Luo, Yuxin Liu, Minhong Zhu, Zhiyun Dai, Xuliang Zhang & Zaizhu Han

To cite this article: Zhenjiang Cui, Xiangqi Luo, Yuxin Liu, Minhong Zhu, Zhiyun Dai, Xuliang Zhang & Zaizhu Han (25 Jul 2025): Impact of imagery deficit on word-based object colour retrieval: Evidence from congenital aphantasia, Cognitive Neuropsychology, DOI: 10.1080/02643294.2025.2536855

To link to this article: https://doi.org/10.1080/02643294.2025.2536855

Published online: 25 Jul 2025.

COGNITIVE NEUROPSYCHOLOGY
https://doi.org/10.1080/02643294.2025.2536855

Routledge
Taylor & Francis Group

Impact of imagery deficit on word-based object colour retrieval: Evidence from congenital aphantasia

Zhenjiang Cui, Xiangqi Luo, Yuxin Liu, Minhong Zhu, Zhiyun Dai, Xuliang Zhang and Zaizhu Han

National Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, People's Republic of China

ABSTRACT
Aphantasia is a form of neurodivergence characterized by an absence of voluntary mental imagery. This absence affects not only basic cognition but also the processing of complex semantic content such as object colour, which typically relies on both visual and verbal representations. According to the Dual Coding Theory (DCT), combining these representations enhances semantic processing. However, this advantage has not been fully investigated under cross-modal conditions. To address this, we tested 24 individuals with congenital aphantasia and 22 controls on object colour decision and retrieval tasks using picture and word stimuli. Unlike controls, individuals with aphantasia showed no benefit from picture-based learning when retrieving colour via words. While their accuracy was unimpaired, their response efficiency was reduced. These findings support DCT and demonstrate the importance of visual imagery in facilitating cross-modal retrieval of object colour under verbal conditions.

ARTICLE HISTORY
Received 20 December 2024
Revised 9 July 2025
Accepted 15 July 2025

KEYWORDS
Aphantasia; visual representation; verbal representation; dual coding theory; object colour

Introduction

Aphantasia is a form of neurodivergence, characterized by a reduced or absent capacity for voluntary mental imagery (Bainbridge et al., 2021; Bates & Farran, 2021; Dance et al., 2021; Keogh et al., 2021; Keogh & Pearson, 2018; Milton et al., 2021; Monzel et al., 2022, 2023; Pounder et al., 2022; Zeman et al., 2015). The absence of visual imagery in aphantasia influences various cognitive processes. Studies have shown that individuals with aphantasia exhibited poor performance in autobiographical memory (Milton et al., 2021), verbal and visual short-/long-term memory (Monzel et al., 2022), drawing recall (Bainbridge et al., 2021), face recognition (Dance et al., 2023) and demonstrate a lack of attentional guidance through visual imagery (Monzel et al., 2021). Moreover, in binocular rivalry tasks, individuals with aphantasia cannot induce a perceptual bias toward red or green Gabor patches through internal visual imagery (Keogh & Pearson, 2018) and fail to produce a pupillary light response when attempting to form internal visual representations (Kay et al., 2022). Besides, a study utilizing skin conductance measurements also revealed that individuals with aphantasia maintain a fear response to frightening images but not to written material (Wicken et al., 2021). Other studies have found that individuals with aphantasia perform similarly to controls in some low-difficulty visual working memory tasks, and in some cases even better (Jacobs et al., 2018; Keogh et al., 2021; Pounder et al., 2022). These findings support the idea that the cognitive processing procedures or strategies in individuals with aphantasia are altered due to the absence of visual imagery (Hurlburt & Akhter, 2006).

In addition to previously reported differences in basic cognitive functions such as memory, attention, and emotional experiences, individuals with aphantasia also exhibit differences from control participants in the processing of more complex and abstract semantic concepts. For instance, individuals with aphantasia are not affected by the verbal overshadowing effect (VOE) (Monzel et al., 2024; Kay et al., 2024). Additionally, compared to the control group, this population exhibited slower reaction time (RT) in voice-based imagery judgment tasks involving colour, shape, words, and faces, but their accuracy was similar (Liu et al., 2023; Liu & Bartolomeo, 2023). Among these properties, colour, as a visual property of objects, is often associated with specific object concepts (e.g., apples with red, bananas with yellow). Additionally, colour has uniqueness (Bi, 2021; Striem-Amit et al., 2018; Wang et al., 2020), unlike shape (which can be activated by somatosensory or tactile representations, such as the distinction

CONTACT Zaizhu Han zzhhan@bnu.edu.cn National Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing 100875, People's Republic of China

between sharp and smooth) and size (where descriptions of object size do not always correspond precisely to their objective measurements). The activation and retrieval of colour primarily rely on internal visual or verbal representations, and there is often a precise correspondence between internal visual representations and verbal representations of colour (Gleason et al., 2004; Kozhevnikov et al., 2005; Rich et al., 2006). However, existing research has not investigated the precise correspondence between the visual and verbal representations of object concepts and their colour properties. Furthermore, it has not explored how the absence of visual imagery affects the retrieval of object colour properties under cross-modal conditions when different types of learning or memory stimulus are experimentally manipulated.

Researchers who support Dual Coding Theory (DCT) argue that the simultaneous use of visual and verbal representations for cognitive processing is more advantageous than relying on a single strategy (Clark & Paivio, 1987; Linde & Paivio, 1979; Paivio, 1963; Paivio & Csapo, 1973). The discovery of aphantasia, combined with the unique nature of colour as an object property, provide a distinctive perspective to investigate the application of DCT in the semantic processing and memory retrieval of object colour. According to this theory, when retrieving the colour properties of objects, the concurrent use of visual and verbal representation strategies facilitates the retrieval of object colour (Hargis, 1978; Vandenberghe et al., 1996; Yui et al., 2017). However, it is unclear whether this hypothesis holds under cross-modal stimulus conditions. Several studies have indicated that under picture stimulus conditions, the frequency of using visual representation strategies is significantly reduced (D'Angiulli, 2002; Kay et al., 2022; Meng et al., 2023; Zeman et al., 2010; Monzel et al., 2021, p. 2022, 2024), therefore, there may be no significant behavioural differences between individuals with aphantasia and those with typical imagery. Furthermore, it remains unclear whether individuals with aphantasia differ from those with typical imagery when retrieving object colour properties based on word stimuli. Considering that processing word stimuli may involve both verbal and imagery representations, and that the content of these representations may depend on prior learning exposure, we assumed that when participants had previously learned to picture stimuli, individuals with aphantasia—who lack imagery representations—may experience certain difficulties in retrieval. The specific pattern of such difficulties deserves further investigation, as it could provide valuable insights for theories concerning semantic concept processing related to visual representation.

Based on the aforementioned gaps and theoretical assumptions, this study comprehensively explores the impact of imagery representation deficits on cross-modal object colour properties retrieval, particularly under word-based retrieval conditions. We recruited 24 congenital aphantasia and 22 control participants. Each participant completed two successive tasks, namely, object colour decision and retrieval (see Figure 1), along with several questionnaires aimed at measuring verbal and visual imagery ability (Figure 2). The stimuli for each task consisted of visually coloured object pictures or written Chinese names. In the colour decision task, participants were instructed to determine whether the colours of the pictures or words were typical or deviated from their mental representations, while also being instructed to remember the colours of the stimuli with atypical colours. In the colour retrieval task, participants were asked to ascertain whether the colours of the pictures or words were consistent with those identified in the colour decision task. The first task aimed to elucidate how deficits in imagery representation influence the retrieval of object colour knowledge based on long-term past life experience. The second task, in turn, aimed to examine how visual imagery deficits affect cross-modal object colour retrieval in short-term memory, while excluding the influence of long-term life experience. To our knowledge, this is the first study to specifically investigate how visual imagery deficits in individuals with aphantasia affect cross-modal retrieval of object colour properties, using both word and picture stimuli.

Materials and methods

Participants

This study determined the required sample size using G*Power 3.1.9.7 software. The Type I error probability was set at $\alpha = 0.05$, with a statistical power of $1 - \beta = 0.8$, and a medium effect size ($f = 0.25$). The correlation among repeated measures was set to the default value of 0.5. The calculation indicated that a total sample size of 34 participants was required.

In total, 46 participants were recruited, all of whom were native Chinese Mandarin speakers with normal or corrected-to-normal vision, no colourblindness, and no history of mental illness or other clinical conditions that could potentially affect the results. Both control and aphantasia participants were primarily recruited through active online social platforms commonly used by young people. Except for one 37-year-old male with a university degree in the aphantasia group, all other

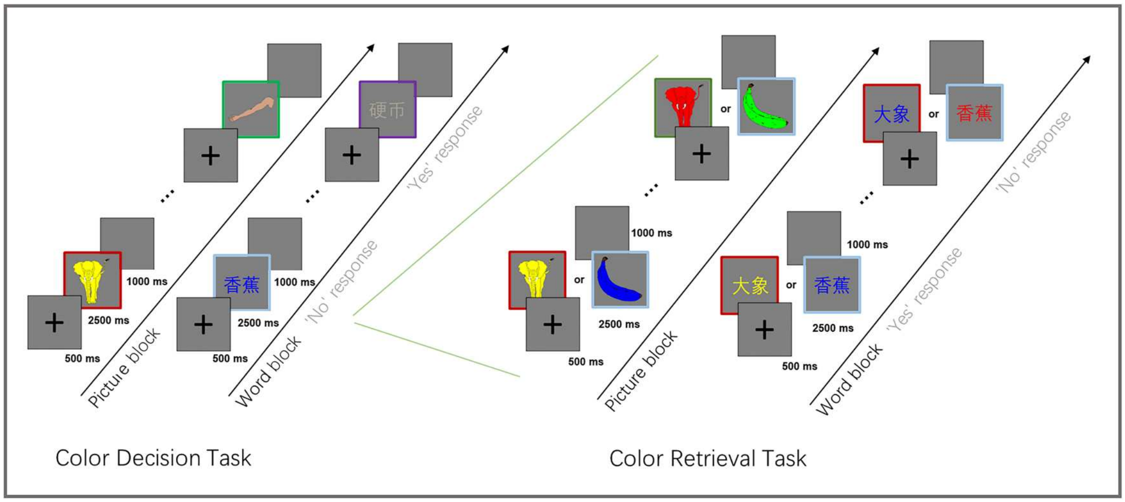

Figure 1. Experimental stimuli and procedure for the object colour decision and retrieval task. Different stimulus borders represent different categories of stimuli (red: animals, blue: vegetables/fruits, green: body parts, purple: other objects). [To view this figure in colour, please see the online version of this journal.]

participants were undergraduate or graduate students. During the recruitment process, the aphantasia participants contacted us and self-reported their condition, which was later confirmed through the completion of other relevant questionnaires. Each participant completed the Chinese version of the Vividness of Visual Imagery Questionnaire (VVIQ; Marks, 1973; see details below), a widely used tool for distinguishing individuals with aphantasia. Previous large-sample studies have demonstrated that the Chinese version of the VVIQ has good reliability and validity (Sun et al., 2023; Zhou et al., 2020). According to established classification criteria (Bainbridge et al., 2021; Pounder et al., 2022), we classified 24 participants (14 females) as individuals with aphantasia (VVIQ score: 17.13 ± 2.58, range: 16-25; age: 24.00 ± 5.15, range: 18-37) and 22 participants (11 females) as controls (VVIQ score: 67.36 ± 8.18, range: 53-80; age: 23.82 ± 2.52, range 20-30). These two participant groups had significant differences in VVIQ scores ($t_{44} = 27.58$, *FDR-corrected* $q < 0.001$, Cohen's $d = 8.24$, $BF_{10} = 2.83 \times 10^{26}$) and were comparable in age ($t_{34.04} = -0.15$, *FDR-corrected* $q = 0.88$, Cohen's $d = -0.04$) and sex ($\chi^2 = 0.32$, *FDR-corrected* $q = 0.57$, Cramer's $V < 0.01$). All participants provided written informed consent and were compensated for their participation. Ethical approval of this study was obtained from the Institutional Review Board of the affiliated institution.

Questionnaire

All participants completed the questionnaire online. To ensure data quality, we included probe items in the questionnaire to assess whether participants were responding attentively. Specifically, certain items instructed participants to select a particular answer, and compliance with these instructions was used as an indicator of attentiveness. Only the data from participants who answered all of these items correctly were considered valid for analysis.

VVIQ

In previous studies, the VVIQ was used as a psychometric test of individual differences in visual imagery vividness (Bates & Farran, 2021; Keogh & Pearson, 2018; Milton et al., 2021; Zeman et al., 2015). The VVIQ involves participants imagining four scenes (e.g., a friend's physical features, a rising sun, a shop or a country scene), each encompassing four specific aspects related to them (e.g., the overall appearance of the shop, a specific scenario of exchanging money with the counter assistant). Thus, the VVIQ comprises a total of 16 items. For each item, participants are asked to subjectively rate each aspect on a scale of 1–5 based on the vividness of the image in their mind (1 = no image at all, 3 = medium, 5 = perfectly clear). The scores from the VVIQ completed by participants were summed, with the total possible score ranging between 16 and 80 points. A higher score indicates greater vividness of the image (Dance et al., 2023). Although VVIQ scores have been extensively used to screen for aphantasia, the range of scores used to define aphantasia varies across studies. For instance, some psychologists set the criteria for aphantasia within the range of 16–32 points (Dawes et al., 2020; Milton et al., 2021), while another group defined it as 16–23 points (Milton et al., 2021; Zeman et al., 2020).

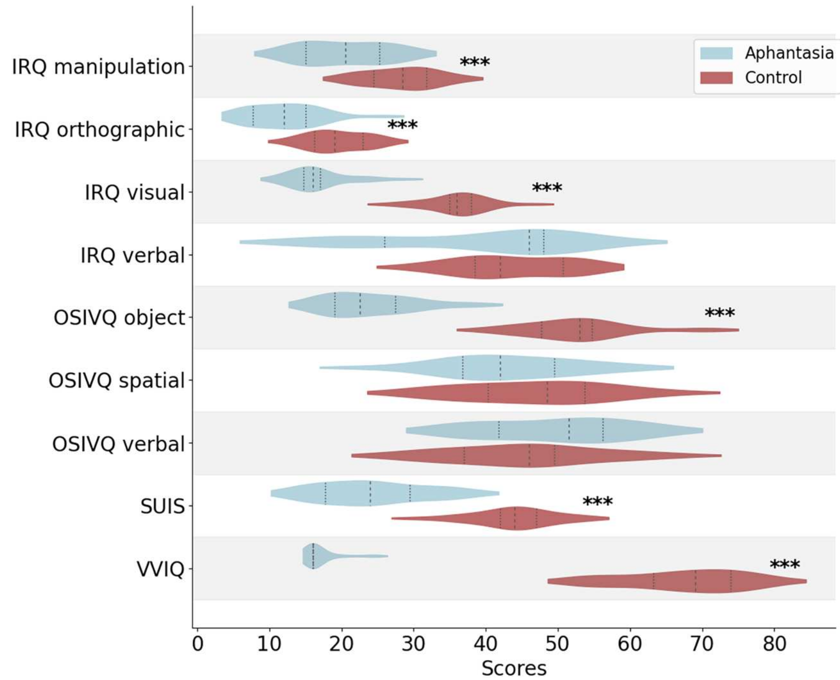

Figure 2. Violin plots comparing scores on various cognitive and imagery-related questionnaires between the aphantasia and control groups. *** BF_{10}: >30. [To view this figure in colour, please see the online version of this journal.]

Some recent researchers utilized a range of 16–25 points (Bainbridge et al., 2021; Pounder et al., 2022), which was also adopted in this study.

Other relevant questionnaires

In addition to the differences observed in VVIQ scores between individuals with aphantasia and controls, group differences have also been observed in other questionnaires designed to assess various dimensions of mental representational abilities (Blazhenkova & Pechenkova, 2019; Keogh & Pearson, 2018; Takahashi et al., 2023). The first questionnaire is the Object-Spatial Imagery and Verbal Questionnaire (OSIVQ), which measures three processing systems in participants (Huang et al., 2025). The first is the verbal system, where scores reflect participants' ability to express, comprehend, and produce spoken and written language. Second, for the object system, the scores indicate the participant's ability to represent objects or scenes in colour, detail, and shape and to recognize objects. Finally, the spatial system, with scores indicating the participant's ability to understand location, movement, spatial relationships, manipulation and transformation of objects (Blazhenkova & Kozhevnikov, 2009).

The second questionnaire is the Internal Representations Questionnaire (IRQ), which has been demonstrated to possess high internal validity and good test-retest reliability. Initially, the IRQ contained only items measuring the extent to which people experience their thoughts in the form of language and use language to guide their thinking (verbal dimension). Subsequently, it was expanded to include measures for the vividness of visual imagery (visual dimension), items pertaining to orthographic imagery (orthographic dimension), and items that measure people's subjective ease of manipulating mental representations across modalities (manipulation dimension) (Roebuck & Lupyan, 2020).

The last questionnaire used was the Spontaneous Use of Imagery Scale (SUIS). Participants were asked to rate 12 items on a scale from 1 to 5, indicating how well each item fit them (5 = perfectly suitable, and 1 = not suitable). The questionnaire measures the extent of participants' spontaneous use of mental imagery, for example, "when I am looking for new furniture in a store, I always visualize what the furniture would look like in particular places in my home", and "when I hear a friend's voice, a visual image of him or her almost always springs to mind" (Reisberg et al., 2003).

To ensure that participants responded attentively throughout the questionnaires, we embedded attention probe items at predetermined points following several standard questions. These items featured the same response options as regular questions but included explicit instructions directing participants to select a specific response (e.g., "Please choose the "medium" option," corresponding to the third choice, "medium," in the VVIQ questionnaire). Participants were not informed in advance about the presence or placement of these probes. A total of six attention probe items were included, and only those who responded correctly to all of them were included in the subsequent analyses.

The above questionnaires were administered in their Chinese versions following standard procedures. Scores from these questionnaires provide supplementary data for exploring the multidimensional disparities in internal visual representation between individuals with and without aphantasia.

Reliability and validity of the Chinese questionnaires

Given the lack of large-sample studies examining the reliability and validity of the Chinese versions of the aforementioned questionnaires, and the limited number of participants in the behavioural task, we collected an additional dataset from a larger sample of native Chinese speakers to better evaluate their psychometric properties. These participants were recruited using the same procedures as those described earlier for the control and aphantasia groups—that is, through online social platforms—and completed the questionnaires online. Only participants who correctly answered all attention probe items were included in the final sample. We collected a total of 321 questionnaire responses. After data screening, a total of 278 valid questionnaires were obtained (including the 22 controls and 24 individuals with aphantasia mentioned earlier). The final sample consisted of 115 males and 163 females (age: 23.09 ± 3.71, range: 18-50), most of whom were undergraduate or graduate students.

Based on this larger sample size, we assessed the reliability and structural validity of the questionnaires (see Table 1). Reliability was evaluated using Cronbach's alpha coefficients, while structural validity was examined through the Kaiser-Meyer-Olkin (KMO) measure and Bartlett's test of sphericity. The results indicated that all questionnaires demonstrated acceptable reliability and structural validity: Cronbach's alpha coefficients were all above 0.75, KMO values exceeded 0.75, and Bartlett's tests were statistically significant ($p < 0.001$, marked as *** in Table 1). In addition, exploratory factor analyses (EFA) revealed total variance explained ranging from 46.62% to 80.16%, further supporting the structural validity of the instruments.

Table 1. Psychometric Properties of the Chinese Questionnaires.

Name		Validity		Reliability
		KMO (Bartlett's *p*-value)	Total Variance Explained	Cronbach's α
VVIQ		0.97 ***	80.16%	0.98
IRQ	Total	0.92 ***	61.02%	0.92
	manipulation	0.78 ***	60.87%	0.80
	orthographic	0.79 ***	46.62%	0.76
	Visual	0.95 ***	62.66%	0.90
	Verbal	0.89 ***	52.19%	0.87
OSIVQ	Total	0.92 ***	63.36%	0.82
	object	0.96 ***	66.27%	0.94
	spatial	0.84 ***	60.85%	0.76
	verbal	0.83 ***	57.60%	0.77
SUIS		0.94 ***	58.64%	0.90

*** indicates *p*-value < 0.001

Behavioural tasks

All behavioural tasks in this study were coded in Psychopy (version 2021.2.3, https://www.psychopy.org/) and presented on the NaoDao online platform (https://www.naodao.com/). During the online behavioural tasks, participants were required to be in a quiet environment, share their screens, and keep their microphones on throughout the session. The entire procedure was supervised by the experimenter via a virtual meeting platform (e.g., Zoom or Tencent Meeting). In addition, the experimenter was available throughout the session to address any questions or technical issues the participants might have encountered.

Each participant completed two tasks consecutively: the object colour decision task followed by the object colour retrieval task. Each participant completed all trials in both tasks. Each task could be considered as a two-way or three-way mixed design, with the participant group as a between-subject factor and the other factors as within-subject factors (e.g., learning stimulus type or retrieval stimulus type).

Object color decision task

In this task, we initially selected 72 common objects (e.g., elephant, banana, arm, coin) from the Multilingual Picture databank (Duñabeitia et al., 2018), with 36 objects presented as picture input stimuli and the remaining 36 objects presented as their written Chinese names (word stimuli). Both types of stimuli were evenly distributed across semantic categories: animals, vegetables/fruits, body parts, and other objects. Subsequently, 24 picture stimuli were modified to display atypical colour using Adobe Photoshop 2023 software. The remaining 12 object pictures retained their original, typical colours from the Picture databank. For the word stimuli, 24 object names were selected for atypical colour processing, while the remaining 12 object names underwent typical colour processing based on the original and typical colours of objects in the Picture databank. All colour manipulations were performed using Adobe Photoshop 2023. There were 8 atypical colours, defined by their RGB values: (255, 0, 0), (0, 255, 0), (0, 0, 255), (255, 255, 0), (0, 128, 128), (102, 102, 51), (88, 50, 50), and (200, 200, 255). Each atypical colour was paired with objects for each type of stimulus.

Consequently, the 72 distinct object stimuli were further categorized into four types: 24 pictures with atypical colour (e.g., a yellow elephant), 24 words with atypical colour (e.g., a blue written "香蕉" meaning banana), 12 pictures with typical colour (e.g., a flesh-coloured arm, RGB: 211, 161, 134), and 12 words with typical colour (e.g., a silver "硬币" meaning a Chinese coin, typically silver RGB: 168, 166, 151) (examples in Figure 1). Additionally, to avoid colour ambiguity, we deliberately avoided using colours close to the object's typical colour (e.g., yellow or green were not used as atypical colours for bananas). In total, 48 items with atypical colours constituted the "no" response trials, while the 24 items with typical colours constituted the "yes" response trials.

The stimuli were then combined into 2 word blocks and 2 picture blocks, each containing 12 objects with atypical colours and 6 objects with typical colours. The presentation order of the stimuli within each block was randomized for each participant. In addition to deciding whether the colour of the objects was their typical colour, participants were also required to try to remember the colours of the stimuli with atypical colours. Prior to each stimulus presentation, a 500 ms fixation point was displayed, followed by the stimulus for 2500 ms, during which participants were required to determine whether the object's colour in the picture or word matched its typical colour in real life. Responses were recorded by pressing the "1" key for "yes" and the "0" key for "no" on the main area of the computer keyboard. Following stimulus presentation, there was a 1000 ms blank screen interval before proceeding to the fixation point for the next stimulus (see Figure 1). To exclude the effects of the presentation order between the word and picture stimuli, we employed an ABBA experimental sequence.

Therefore, the data collected from colour decision task followed a two-way mixed design: 2 participant groups (aphantasia, control) * 2 stimulus types (picture, word). The first factor included 24 participants in the aphantasia and 22 in the control groups. For the second factor, each stimulus type comprised 36 objects.

Object color retrieval task

After completing a colour decision task for one block (whether word or picture), participants immediately completed one block of the colour retrieval task based on word stimuli and one block of the colour retrieval task based on picture stimuli. In this task (i.e., the second task), the 48 objects with atypical colours in the above colour decision task (i.e., the first task) (comprising 24 pictures and 24 words) were each expanded into four stimuli: one picture and one word with the same colour as in the first task and one picture and one word with another atypical colour derived from the eight options previously selected. Consequently, the 48 original objects were expanded into 196 items with atypical colours: 48 pictures and 48 words matching the colour in the first task and an additional 48 pictures and 48 words with colours different from those in the first task.

The 192 items were totally divided into 8 blocks: 4 picture blocks and 4 word blocks. Each block contained 24 objects, including 12 with the same colour as in the first task and 12 with colours different from those in the first task. For each trial, the experimental procedure in this task was identical to that of the first task, except that participants were required to judge whether the colours of the objects matched those presented in the first task (see Figure 1). Half of the stimuli ($n = 96$) elicited "yes" responses, and the other half ($n = 96$) elicited "no" responses.

This task also followed a two-way mixed design: 2 participant groups (aphantasia, control) * 4 stimulus consistency types (picture-picture, picture-word, word-picture, word-word). The first factor included 24 participants in the aphantasia group and 22 in the control group. The second factor indicated the type of stimulus used during both the learning and retrieval phases. Among the 192 total items, 48 involved pictures at both stages (picture-picture), 48 involved words at both stages (word-word), 48 used pictures during learning and words during retrieval (picture-word), and 48 used words during learning and pictures during retrieval (word-picture). Each condition included 24 "yes" and 24 "no" response trials. The goal of this task was to investigate whether differences between stimulus types across the learning and retrieval stages affected memory retrieval processes. To minimize the influence of participants' prior life experiences, all stimuli in this task represented objects with atypical colours.

Given that the stimulus consistency type in the object colour retrieval task includes four levels (picture-picture, picture-word, word-picture, word-word), directly analyzing these four levels may obscure whether it is the type of the learning stimulus or the type of the retrieval stimulus that influences behavioural performance. Therefore, in the current analysis, we firstly reorganized these four conditions into two factors. For the learning stimulus type, we grouped conditions based on the stimulus presented during learning: the picture-picture and picture-word conditions were combined to form the picture learning condition, while the word-picture and word-word conditions were combined to form the word learning condition. A similar approach was applied to define the retrieval stimulus type, with conditions regrouped based on the type of stimulus presented during retrieval.

Post-experiment interview

Following the completion of the two behavioural tasks, brief post-experiment interviews were conducted with participants in the aphantasia group to explore the strategies they employed during the tasks. Some individuals with aphantasia reported that, although they were unable to generate internal visual representations of objects, they were still able to access or infer visual and colour properties through verbal or linguistic means—a process they described as akin to having a phonograph in the mind.

Data analysis

The scores of four questionnaires were computed and analyzed using Python. Independent-samples t-tests were used to examine between-group differences, with false discovery rate (FDR) correction for multiple comparisons.

For the object colour decision task, a two-way mixed-design analysis was performed using linear mixed-effects (LME) models. The statistical analysis for the object colour retrieval task followed the same procedure as the decision task, investigating the interactions between learning or retrieval stimulus types and group. The LME models tested the fixed effects of subject group (aphantasia vs. control) and stimulus type (picture vs. word), with random effects including variation in intercepts across subjects. The model syntax is as follows:

$$accuracy\ (correctRT) \sim group{*}stimulus\_type + (1|subject)$$

To further investigate the Impact of long-term and short-term memory on object colour retrieval in aphantasia and control groups, the colour decision task was compared with the object colour retrieval task. A three-way mixed-design analysis based on LME models was also conducted. The model tested the fixed effect of participant group (aphantasia vs. control), task type (colour decision vs. object colour retrieval), and retrieval stimulus type (picture vs. word), with random effects including variation in intercepts across subjects. The model syntax is as follows:

$$accuracy\ (correct\ RT) \sim group{*}task\_type{*}retrieval\_stimulus\_type + (1|subject)$$

Finally, in the object colour retrieval task, a three-way mixed-design analysis based on linear mixed-effects (LME) models was conducted to examine the interaction among group (aphantasia vs. control), learning stimulus type (picture vs. word) and retrieval stimulus type (picture vs. word). The model syntax is as follows:

$$accuracy\ (correct\ RT) \sim group{*}learning\_type{*}retrieval\_type + (1|subject)$$

All analyses related to the LME models were performed in MATLAB R2022a. A FDR correction was applied for multiple comparisons. Additionally, Bayesian ANOVAs and t-tests (Keysers et al., 2020; Vincent, 2015) were conducted in JASP software (version 0.18.3.0, https://jasp-stats.org/) to complement the frequentist analyses.

Results

Demographics and questionnaire scores

Figure 2 illustrates the differences between the two participant groups on the 4 questionnaire measures. Besides scoring significantly lower on the VVIQ compared to the control group, we also found that the aphantasia group had lower on object OSIVQ ($t_{44} = -14.24$, *FDR-corrected* $q < 0.001$, Cohen's $d = -4.20$, $BF_{10} > 1000$), SUIS ($t_{44} = -10.16$, *FDR-corrected* $q < 0.001$, Cohen's $d = -3.00$, $BF_{10} > 1000$), visual IRQ ($t_{44} = -15.12$, *FDR-corrected* $q < 0.001$, Cohen's $d = -4.46$, $BF_{10} > 1000$), orthographic IRQ ($t_{44} = -5.39$, *FDR-corrected* $q < 0.001$, Cohen's $d = -1.59$, $BF_{10} > 1000$) and manipulation IRQ ($t_{44} = -5.01$, *FDR-corrected* $q < 0.001$, Cohen's $d = -1.48$, $BF_{10} > 1000$) scores than control group. These findings further confirm that aphantasia had a difficulty in visual imagery. In contrast, the two groups had comparable scores on the verbal OSIVQ ($t_{44} = 1.49$, *FDR-corrected* $q = 0.17$, Cohen's $d = 0.44$, $BF_{10} = 0.72$), spatial OSIVQ ($t_{44} = -1.31$, *FDR-corrected* $q = 0.20$, Cohen's $d = -0.39$, $BF_{10} = 0.59$) and verbal IRQ ($t_{37.11} = -1.49$, *FDR-corrected* $q = 0.17$, Cohen's $d = -0.44$, $BF_{10} = 0.72$).

These results indicate that individuals with aphantasia in this study exhibited normal internal verbal representation ability at the group level. This finding aligns with participants' self-reports during post-task interviews.

Performance in the object color decision task

Figure 3 presents the results of the LME models for mean accuracy and correct RT.

For accuracy, neither main effects was significant (Group: $F_{(1,\ 44)} = 1.35$, $p = 0.25$, $\eta_p^2 = 0.03$, $BF_{incl} = 0.44$; Stimulus Type: $F_{(1,\ 44)} = 1.67$, $p = 0.20$, $\eta_p^2 = 0.04$, $BF_{incl} = 0.33$). Their interaction was also non-significant ($F_{(1,\ 44)} = 0.46$, $p = 0.50$, $\eta_p^2 = 0.01$, $BF_{incl} = 0.16$). No significant differences were found between groups under either stimulus condition (Picture: $t_{44} = -0.66$, *FDR-corrected* $q = 0.69$, Cohen's $d = -0.19$, $BF_{10} = 0.35$; Word: $t_{44} = -1.30$, *FDR-corrected* $q = 0.40$, Cohen's $d = -0.382$, $BF_{10} = 0.58$). Similarly, neither group showed significant differences between picture and word conditions (Aphantasia: $t_{23} = 1.40$, *FDR-corrected* $q = 0.35$, Cohen's $d = 0.29$, $BF_{10} = 0.51$; Control: $t_{21} = 0.43$, *FDR-corrected* $q = 0.79$, Cohen's $d = 0.09$, $BF_{10} = 0.24$).

For correct RT, the main effect of group was not significant ($F_{(1,\ 44)} = 1.03$, $p = 0.32$, $\eta_p^2 = 0.23$, $BF_{incl} = 1.06$), whereas the main effect of stimulus type was significant ($F_{(1,\ 44)} = 6.70$, $p = 0.01$, $\eta_p^2 = 0.13$, $BF_{incl} = 4.03$). A significant interaction between them was also observed ($F_{(1,\ 44)} = 4.96$, $p = 0.03$, $\eta_p^2 = 0.10$). However, Bayesian ANOVA provided only anecdotal evidence for this interaction ($BF_{incl} = 2.52$). No significant between-group differences were found for either stimulus type (Picture: $t_{44} = 0.14$, *FDR-corrected* $q = 0.89$, Cohen's $d = 0.04$, $BF_{10} = 0.30$; Word: $t_{44} = 1.72$, *FDR-corrected* $q = 0.37$, Cohen's $d = 0.51$, $BF_{10} = 0.96$). However, a simple effects analysis revealed that the aphantasia group responded significantly slower to word stimuli than to pictures ($t_{23} = -3.30$, *FDR-corrected* $q = 0.01$, Cohen's $d = -0.67$, $BF_{10} = 13.11$). whereas no such difference was found in the control group ($t_{21} = -0.27$, *FDR-corrected* $q = 0.79$, Cohen's $d = -0.06$, $BF_{10} = 0.23$).

In summary, these findings suggest that the visual imagery deficit in individuals with aphantasia selectively impairs the efficiency of retrieving object colour knowledge from word-based stimuli, while leaving the accessibility of that knowledge intact.

Performance in the object color retrieval task

The impact of learning stimulus type on object color retrieval

Similar LME models, along with t-tests, were conducted to investigate the main effects and interaction between learning stimulus type and participant group on accuracy and correct RT (see Figure 4).

For accuracy, the main effect of participant group was not significant ($F_{(1,\ 44)} = 0.04$, $p = 0.85$, $\eta_p^2 < 0.01$, $BF_{incl} = 0.40$), whereas the main effect of learning stimulus type was significant ($F_{(1,\ 44)} = 46.58$, $p < 0.001$, $\eta_p^2 = 0.12$, $BF_{incl} > 1000$). The interaction between them was not significant ($F_{(1,\ 44)} = 0.04$, $p = 0.84$, $\eta_p^2 < 0.01$, $BF_{incl} = 0.38$). No significant group differences were found under either the picture or word learning condition (Picture: $t_{44} = 0.21$, *FDR-corrected* $q = 0.92$, Cohen's $d = 0.06$, $BF_{10} = 0.30$; Word: $t_{44} = -0.25$, *FDR-corrected* $q = 0.92$, Cohen's $d = -0.08$, $BF_{10} = 0.30$). Paired-sample t tests revealed that both aphantasia and control groups performed significantly more accurately in the picture learning condition than in the word condition (Aphantasia: $t_{23} = 5.43$, *FDR*-corrected $q < 0.001$, Cohen's $d = 1.11$, $BF_{10} = 1416.99$; Control: $t_{21} = 4.30$, *FDR*-corrected $q < 0.001$, Cohen's $d = 0.92$, $BF_{10} = 97.81$).

For correct RT, while the main effect of participant group was again not significant ($F_{(1,\ 44)} = 0.43$, $p = 0.52$, $\eta_p^2 = 0.01$, $BF_{incl} = 0.95$), while the main effect of stimulus type was significant ($F_{(1,\ 44)} = 62.69$, $p < 0.001$, $\eta_p^2 = 0.59$, $BF_{incl} > 1000$). The interaction between them was marginally significant ($F_{(1,\ 44)} = 3.79$, $p = 0.06$, $\eta_p^2 = 0.08$). with Bayesian analysis providing anecdotal support for the interaction effect ($BF_{incl} = 1.99$). No significant between-group differences were observed under either learning condition (Picture: $t_{44} = 1.12$, *FDR-corrected* $q = 0.92$, Cohen's $d = 0.33$, $BF_{10} = 0.49$; Word: $t_{44} = -0.11$, *FDR-corrected* $q = 0.92$, Cohen's $d = -0.03$, $BF_{10} = 0.29$). Paired-sample t-tests indicated that both groups responded significantly faster in the picture learning condition than in the word condition (Aphantasia: $t_{23} = -3.92$, *FDR*-corrected $q < 0.001$, Cohen's $d = -0.80$, $BF_{10} = 48.69$; Control: $t_{21} = -7.80$, *FDR*-corrected $q < 0.001$, Cohen's $d = -1.66$, $BF_{10} > 1000$).

In summary, these results suggest that the absence of visual imagery in individuals with aphantasia does not alter sensitivity to the type of learning stimulus, as both groups benefited similarly from picture-based learning in terms of accuracy and speed.

The impact of retrieval stimulus type on object color retrieval

Similar LME models, along with t-tests, were conducted to investigate the main effects and interaction between retrieval stimulus type and participant group on accuracy and correct RT (see Figure 5).

For accuracy, neither the main effect of participant group ($F_{(1,\ 44)} = 0.04$, $p = 0.85$, $\eta_p^2 < 0.01$, $BF_{incl} = 0.40$) nor the main effect of stimulus type ($F_{(1,\ 44)} = 1.43$, $p = 0.24$, $\eta_p^2 = 0.03$, $BF_{incl} = 0.37$) reached significance. Their interaction was also not significant ($F_{(1,\ 44)} = 1.86$, $p = 0.18$, $\eta_p^2 = 0.04$, $BF_{incl} = 0.25$). Given the absence of significant effects, no additional t-tests on accuracy were reported.

For correct RT, the main effect of participant group remained non-significant ($F_{(1,\ 44)} = 0.52$, $p = 0.47$, $\eta_p^2 = 0.01$, $BF_{incl} = 0.58$), while a significant main effect of stimulus type was observed ($F_{(1,\ 44)} = 33.49$, $p < 0.001$, $\eta_p^2 = 0.43$, $BF_{incl} > 1000$). The interaction between them was not significant ($F_{(1,\ 44)} = 0.34$, $p = 0.56$, $\eta_p^2 < 0.01$, $BF_{incl} = 0.54$). No significant group differences were found under either the picture or word condition (Picture: $t_{44} = 0.59$, *FDR*-corrected $q = 0.76$, Cohen's $d = 0.17$, $BF_{10} = 0.34$; Word: $t_{44} = 0.80$, *FDR*-corrected $q = 0.76$, Cohen's $d = 0.24$, $BF_{10} = 0.38$). However, paired-sample t-tests revealed that both the aphantasia and control groups responded significantly faster when retrieving colour knowledge from pictures compared to words (Aphantasia: $t_{23} = -4.36$, *FDR*-corrected $q <$

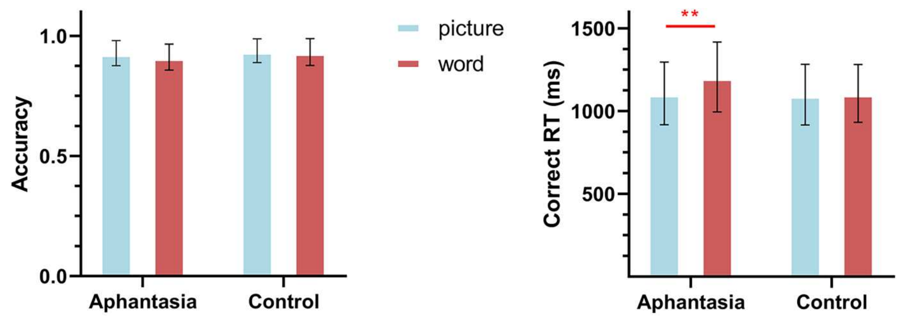

Figure 3. Behavioural performance of participants in the colour decision task under two retrieval stimulus types. ** BF_{10}: 10-30. [To view this figure in colour, please see the online version of this journal.]

0.001, Cohen's $d = -0.89$, $BF_{10} = 128.51$; Control: $t_{21} = -3.86$, *FDR*-corrected $q < 0.001$, Cohen's $d = -0.82$, $BF_{10} = 38.71$).

In summary, both the aphantasia and control groups showed a robust performance advantage when retrieving object colour information from picture-based stimuli, reflected in significantly faster correct RT.

The impact of long-term and short-term memory on object color retrieval in aphantasia and control groups

In the object colour decision task, participants retrieved object colour knowledge based on their long-term real-world experience, whereas in the object colour retrieval task, the colour knowledge was drawn from short-term memory. The results of the LME model analyses comparing the two tasks are as follows:

For accuracy, only the main effect of task type was significant ($F_{(1,\ 44)} = 176.36$, $p < 0.001$, $\eta_p^2 = 0.80$, $BF_{incl} > 1000$), while no significant interactions were found between task type and stimulus type or participant group ($F_{(1,\ 44)} < 0.2$, $p = 0.664 \sim 0.989$, $\eta_p^2 < 0.01$, $BF_{incl} = 0.06 \sim 0.23$). Post hoc comparisons further revealed that both the aphantasia and control groups performed significantly better in the object colour decision task than in the object colour retrieval task (Aphantasia: $t_{47} = 14.04$, *FDR*-corrected $q < 0.001$, Cohen's $d = 2.03$, $BF_{10} > 1000$; Control: $t_{43} = 10.92$, *FDR*-corrected $q < 0.001$, Cohen's $d = 1.65$, $BF_{10} > 1000$).

For correct RT, again only the main effect of task type reached significance ($F_{(1,\ 44)} = 22.10$, $p < 0.001$, $\eta_p^2 = 0.33$, $BF_{incl} = 558.53$). However, no significant interactions were found between task type and either stimulus type or group ($F_{(1,\ 44)} = 0.10 \sim 2.67$, $p = 0.11 \sim 0.76$, $\eta_p^2 = 0.002 \sim 0.06$, $BF_{incl} = 0.40 \sim 0.74$). Post hoc comparisons showed that both the aphantasia and control groups responded significantly faster in the object colour decision task than in the object colour retrieval task (Aphantasia: $t_{47} = -4.65$, *FDR*-corrected $q < 0.001$, Cohen's $d = -0.67$, $BF_{10} = 770.43$; Control: $t_{43} = -4.02$, *FDR*-corrected $q < 0.001$, Cohen's $d = -0.61$, $BF_{10} = 110.68$).

In summary, the findings support that tasks relying on long-term memory (object colour decision) yield better performance—characterized by shorter correct RT and higher accuracy—than those relying on short-term memory (object colour retrieval). However, no interaction was observed between the reliance on memory type (long-term vs. short-term) and the absence of visual imagery.

Impact of stimulus modality consistency between learning and retrieval on object color retrieval

In the colour retrieval task, the results of the LME model analysis investigating the interaction among group (aphantasia vs. control), learning stimulus type (picture- vs. Word-), and retrieval stimulus type (-picture vs. -word) are as follows:

For accuracy, the results revealed a significant main effect of learning stimulus type ($F_{(1,\ 44)} = 65.68$, $p < 0.001$, $\eta_p^2 = 0.60$, $BF_{incl} > 1000$) and a significant interaction between learning and retrieval stimulus types ($F_{(1,\ 44)} = 25.43$, $p < 0.001$, $\eta_p^2 = 0.37$, $BF_{incl} > 1000$). Interestingly, although the frequentist ANOVA yielded a non-significant result for the main effect of retrieval stimulus type ($F_{(1,\ 44)} = 1.08$, $p = 0.28$, $\eta_p^2 = 0.02$), the Bayesian analysis provided strong evidence in favour of including this effect ($BF_{incl} > 1000$). No other significant main effect of group was found ($F_{(1,\ 44)} = 0.51$, $p = 0.48$, $\eta_p^2 = 0.01$, $BF_{incl} = 0.29$), nor were any significant interactions involving group observed ($F_{(1,\ 44)} = 0.13 \sim 1.47$, $p = 0.26 \sim 0.71$, $\eta_p^2 = 0.002 \sim 0.03$, $BF_{incl} = 0.27 \sim 0.37$) were observed.

For correct RT, there were significant main effects of both learning stimulus type and retrieval stimulus type ($F_{(1,\ 44)} = 21.13 \sim 64.02$, $p < 0.001$, $\eta_p^2 = 0.32 \sim 0.59$, $BF_{incl} > 1000$), as well as a significant interaction between them ($F_{(1,\ 44)} = 240.46$, $p < 0.001$, $\eta_p^2 = 0.85$, $BF_{incl} > 1000$). There was a significant three-way interaction among group, learning stimulus type, and retrieval stimulus type ($F_{(1,\ 44)} = 4.39$, $p = 0.04$, $\eta_p^2 = 0.09$, $BF_{incl} =$

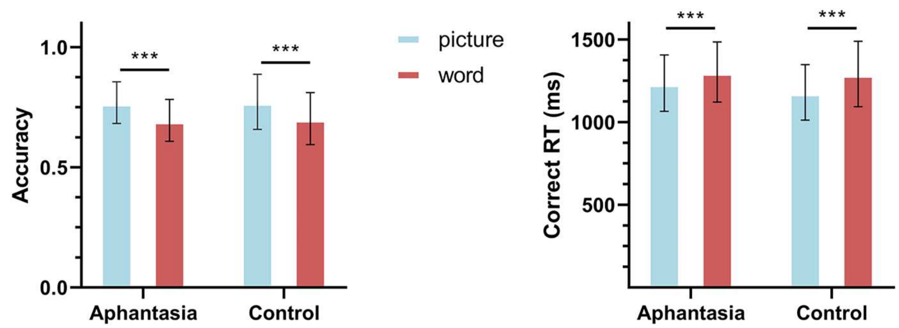

Figure 4. Behavioural performance of participants in the colour retrieval task under two learning stimulus conditions. *** BF_{10}: >30. [To view this figure in colour, please see the online version of this journal.]

4.05). In addition, there was a weak interaction between group and learning stimulus type ($F_{(1,\ 44)} = 3.38$, $p = 0.07$, $\eta_p^2 = 0.01{\sim}0.07$, $BF_{incl} = 1.35$). No significant main effect of group ($F_{(1,\ 44)} = 0.66$, $p = 0.42$, $\eta_p^2 = 0.01$, $BF_{incl} = 0.76$) or interaction between group and retrieval stimulus type was observed ($F_{(1,\ 44)} = 0.82$, $p = 0.37$, $\eta_p^2 = 0.02$, $BF_{incl} = 0.89$).

In summary, although no significant group-level effects were observed, the consistency of the effect patterns of learning and retrieval stimulus types within each group further investigation.

Impact of stimulus modality consistency between learning and retrieval on object color retrieval within groups

Given the absence of significant group-related effects on both accuracy and correct RT, we only examined the interaction between stimulus type and retrieval type separately within the aphantasia and control groups. Considering the complexity of the simple effects and post hoc analyses, we selectively report only the parts where distinct patterns emerged between the aphantasia and control groups (see Figure 6).

For accuracy, LME models revealed a significant main effect of learning stimulus type ($F_{(1,\ 44)} = 29.35{\sim}31.90$, $p < 0.001$, $\eta_p^2 = 0.56{\sim}0.60$, $BF_{incl} > 1000$) and a significant interaction between learning and retrieval stimulus types in both groups ($F_{(1,\ 44)} = 8.34{\sim}20.19$, $p < 0.001$, $\eta_p^2 = 0.28{\sim}0.47$, $BF_{incl} > 1000$). Although the frequentist ANOVA did not detect a significant main effect of retrieval stimulus type ($F_{(1,\ 44)} = 0.28{\sim}0.88$, $p = 0.61{\sim}0.36$, $\eta_p^2 = 0.01{\sim}0.04$), Bayesian analysis provided strong evidence for a pronounced main effect of retrieval stimulus type in the control group ($BF_{incl} = 657.38$) but not in the aphantasia group ($BF_{incl} = 2.27$). The following analysis revealed that in the control group, when the retrieval stimuli were words, participants showed significantly higher accuracy when the learning stimuli were pictures compared to words ($t_{21} = 2.78$, *FDR*-corrected $q = 0.01$, Cohen's $d = 0.59$, $BF_{10} = 4.52$). However, this advantage for picture-based learning under word retrieval conditions was not observed in the aphantasia group ($t_{23} = 0.70$, *FDR*-corrected $q = 0.51$, Cohen's $d = 0.14$, $BF_{10} = 0.27$).

For correct RT, significant main effects of both learning and retrieval stimulus types, as well as their interaction, were observed in both the aphantasia and control groups ($F_{(1,\ 44)} = 7.36{\sim}156.69$, $p < 0.05$, $\eta_p^2 = 0.26{\sim}0.87$, $BF_{incl} > 1000$). Further analysis showed that in the control group, when the retrieval stimuli were words, correct RT was marginally longer following picture-based learning compared to word-based

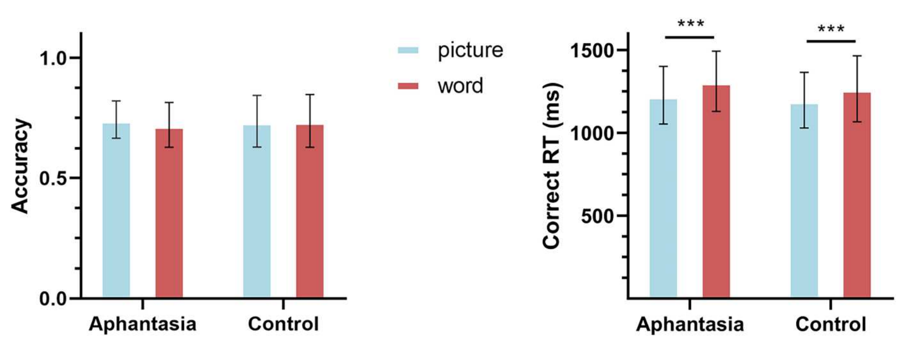

Figure 5. Behavioural performance of participants in the colour retrieval task under two retrieval stimulus conditions. *** BF_{10}: >30. [To view this figure in colour, please see the online version of this journal.]

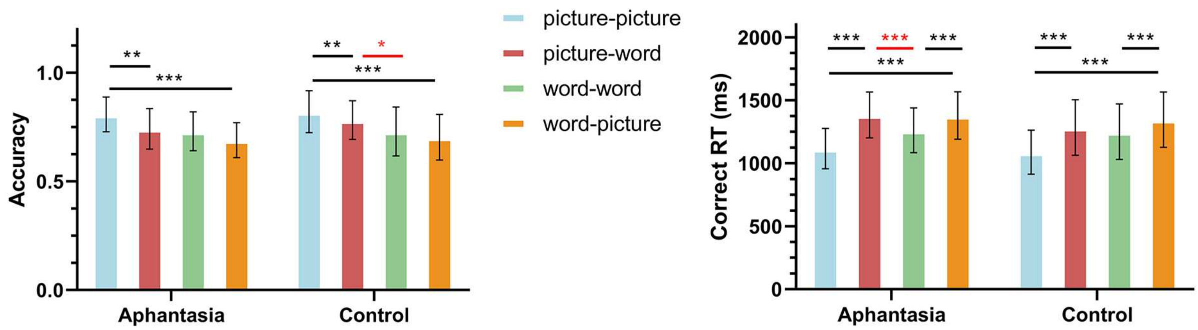

Figure 6. Behavioural performance of participants in the colour retrieval task under different learning and retrieval stimulus conditions. * BF_{10}: 3-10, ** BF_{10}: 10-30, *** BF_{10}: >30. [To view this figure in colour, please see the online version of this journal.]

learning according to the frequentist analysis ($t_{21} = 2.22$, *FDR*-corrected $q = 0.04$, Cohen's $d = 0.47$). However, the Bayesian analysis did not provide substantial support for this difference ($BF_{10} = 1.70$). In contrast, in the aphantasia group, when the retrieval stimuli were words, correct RT was significantly longer following picture-based learning compared to word-based learning ($t_{23} = 5.08$, *FDR*-corrected $q < 0.001$, Cohen's $d = 1.04$, $BF_{10} = 636.88$).

Moreover, under the word retrieval condition, an ANOVA with group and learning stimulus type as fixed factors revealed a significant interaction between group and learning stimulus type in correct RT ($F_{(1,\ 44)} = 9.08$, $p < 0.01$, $\eta_p^2 = 0.17$, $BF_{incl} = 14.01$) but not in accuracy ($F_{(1,\ 44)} = 2.75$, $p = 0.11$, $\eta_p^2 = 0.01$, $BF_{incl} = 0.79$).

In summary, the results of this section support the notion that visual imagery ability not only enhances the accuracy of object colour retrieval in controls, but also suggest that it improves the efficiency of object colour retrieval even under word-based retrieval conditions.

Other impact indicators

To further investigate potential differences in sensitivity and response criterion between the aphantasia and control groups, we applied signal detection theory (SDT) measures to assess interactions between group and various task-related factors across the two tasks. In SDT, *d′* (sensitivity) and *β* (response criterion) are commonly used indices. These measures were calculated as follows:

(1) *d′*: For each participant, the accuracy in "yes" trials was treated as the hit rate (P_{hit}), and accuracy in "no" trials was used to estimate the correct rejection rate ($P_{correct\ rejection}$). The false alarm rate ($P_{false\ alarm}$) was calculated as (1 – $P_{correct\ rejection}$). Both P_{hit} and $P_{false\ alarm}$ were converted into *z*-scores (Z_{hit} and $Z_{false\ alarm}$), and *d′* was calculated as Z_{hit} - $Z_{false\ alarm}$.

(2) β (response criterion): The response criterion β, which reflects the participant's decision bias, was calculated using a simplified formula from standard SDT:

$$log(\beta) = Z^2_{falsealarm} - Z^2_{hit}$$

For the object colour decision task, the analysis revealed no significant main effects of group or stimulus type, nor a significant interaction between them, for either sensitivity ($F_{(1,\ 44)} < 1.062$, $p > 0.31$, $\eta_p^2 < 0.02$, $BF_{incl} < 0.34$) or response criterion ($F_{(1,\ 44)} < 0.91$, $p > 0.34$, $\eta_p^2 < 0.02$, $BF_{incl} < 0.24$).

In the object colour retrieval task, based on previously reported significant findings, we conducted two follow-up analyses: first, we examined the interaction between aphantasia and task type (colour retrieval vs. colour decision) under different retrieval conditions; second, we tested the interaction between group and learning material type when the retrieval cue was a word. However, ANOVA results based on LME models showed that all interactions were not significant (*d′*: $F_{(1,\ 44)} < 3.345$, $p > 0.74$, $\eta_p^2 < 0.02$, $BF_{incl} < 0.29$; log(β): $F_{(1,\ 44)} < 0.57$, $p > 0.46$, $\eta_p^2 < 0.01$, $BF_{incl} < 0.05$).

These findings suggest that the differences in performance patterns between the aphantasia and control groups in object colour decision or retrieval tasks can not be attributable to differences in perceptual sensitivity or response criterion.

Discussion

This study employed both word and picture stimuli to investigate how deficits in visual or imagery representation affect object colour retrieval under cross-modal conditions. To begin with, we administered a battery of questionnaires to comprehensively assess the cognitive abilities of individuals with congenital aphantasia. The results revealed that, compared to controls, the aphantasia group not only exhibited a reduced capacity for visual imagery of objects but also showed poorer performance in visual orthographic representation and mental manipulation. Moreover, individuals with aphantasia reported rarely using visual representation strategies spontaneously in their daily lives. Importantly, no significant differences were observed between the aphantasia and control groups in terms of spatial or verbal representational abilities (Figure 2).

While prior studies involving tasks such as mental rotation have provided evidence for reduced mental manipulation abilities in individuals with aphantasia, the potential decline in orthographic processing within this population has yet to be systematically examined. This raises an intriguing question for future research: although individuals with aphantasia report being unable to generate vivid mental images, some of them describe being able to form simple lines in their minds—could these lines potentially be further organized into meaningful letters or even sequences of words? This possibility remains an open and worthwhile avenue for future investigation.

The results of the object colour decision task showed that individuals with aphantasia exhibited significantly longer correct RT when retrieving object colour properties from long-term semantic memory based on word stimuli, compared to picture-based stimuli. Furthermore, in the object colour retrieval task assessing short-term memory, both the aphantasia and control groups exhibited significantly longer RT when retrieving colour properties under word-based retrieval conditions compared to picture-based conditions. Under word-based retrieval conditions, a further investigation into the interaction between stimulus type and learning stimulus type revealed that, in the control group, learning with pictures resulted in comparable retrieval speed to word learning, but led to higher accuracy. However, in the aphantasia group, this advantage of picture-based learning was not significant, and these individuals even exhibited slower retrieval speeds for picture-based materials. Given the unique nature of colour as a property (Bi, 2021; Striem-Amit et al., 2018; Wang et al., 2020), individuals with aphantasia seem unlikely to employ strategies beyond verbal representation to support this retrieval. Follow-up interviews with participants in the aphantasia group further supported this view, suggesting that some individuals may have relied primarily on verbal representational strategies to complete the behavioural tasks. According to embodied cognition theories, understanding or retrieving an object's visual properties such as colour typically involves the reactivation of brain regions associated with visual functions, along with the subjective experience of visual mental imagery (Marre et al., 2021; Meteyard et al., 2012; Muraki et al., 2023). However, these theories are largely grounded in findings from individuals who are capable of generating such imagery. In contrast, individuals with aphantasia report little to no subjective experience of visual mental imagery (Zeman et al., 2015, 2020; Takahashi et al., 2023; Speed et al., 2024). Our findings indicate that while individuals with aphantasia do not exhibit impairments in accessing object colour knowledge, their retrieval is characterized by significantly slower speed. This raises important questions about the validity of strong embodied cognition theories in explaining conceptual processing in aphantasia. Specifically, it remains unclear whether the reactivation of visual brain regions is essential for successful retrieval of visual properties, or whether it merely plays a facilitative role. Future neuroimaging studies in individuals with aphantasia may offer novel insights into this issue.

In the object colour retrieval task, we did not find a significant effect of learning stimulus on either the accuracy or the correct RT in the aphantasia or control group. One possible explanation is that this effect may have been influenced by the retrieval stimulus. Specifically, the picture-picture condition yielded the best behavioural performance in both groups, while the word-picture condition resulted in the poorest performance. The picture-word condition produced performance comparable to the word-word condition in the aphantasia group and even better performance than the word-word condition in the control group. This pattern may have amplified the overall advantage of picture-based learning stimuli over word-based learning stimuli across both groups, thereby masking any potential interaction effect between learning stimulus and group. A possible explanation for this picture learning advantage effect is as follows: when pictures served as the learning stimuli, participants were required to recognize and name objects—a largely automatic process that placed relatively low demands on cognitive resources. In this case, participants mainly needed to associate the visual features of the object picture with its corresponding name (Curran & Doyle, 2011; McBride & Anne Dosher, 2002; Suzuki & Takahashi, 1997). In contrast, when words were used as the learning stimuli, participants may have spontaneously attempted to generate mental images in preparation for the subsequent colour retrieval task. This often involved imagining an object's image with an atypical colour, a process likely to place a higher cognitive load than imagining a typically coloured object. Although individuals with aphantasia are unable to generate such mental imagery, studies on pupillary responses suggest that merely attempting to form visual imagery still consumes cognitive resources in these individuals.

This increased cognitive load during the learning phase may have negatively affected participants' ability to store and learn word-based stimuli effectively, thus resulting in poorer learning performance for word stimuli across both groups.

The current analysis did not reveal any significant main effects or interactions of retrieval stimulus or group on accuracy in the object colour retrieval task, consistent with the findings from the colour decision task. However, for correct RT, both the aphantasia and control groups demonstrated a picture-based advantage. The object colour retrieval task primarily assessed the relationship between retrieval stimulus type and short-term memory, whereas the object colour decision task targeted long-term or semantic memory. In the latter, a picture retrieval advantage was observed only in the aphantasia group—an effect that appeared to stem more from a tendency toward slower RT for word stimuli, rather than faster responses to picture stimuli. This pattern of results reflects, on one hand, differences in the retrieval of object colour properties from conceptual representations depending on whether short-term or long-term memory is involved. On the other hand, it also suggests that the facilitative role of visual imagery in retrieving object colour information may depend on repeated use and practice of such a strategy. In other words, even individuals with visual imagery may not retain a vivid mental image of an object after a single exposure—despite the potential of visual imagery to support rapid recall. Only through repeated memory encoding does the advantage of visual imagery in facilitating object colour retrieval become evident. At the same time, the absence of visual imagery did not impair participants'accuracy in retrieving object colour information. In other words, the lack of visual imagery may affect the efficiency—but not the accuracy—of object colour property retrieval.

When learning stimulus type, retrieval stimulus type, and group were all included in the statistical model, the results revealed a significant interaction: in the condition where the retrieval stimulus was a word, the control group—who possessed visual imagery—showed a clear advantage when the learning stimulus was a picture (Arterberry et al., 2001; Cherry et al., 2003; Keitz, 1976). Specifically, compared to word-based learning, picture-based learning led to higher accuracy within the control group. This pattern aligns with the DCT, which proposes that individuals with intact visual imagery can simultaneously engage both verbal and imagery strategies during task performance. Such dual engagement could account for the observed differences in performance patterns between the control and aphantasia groups (Linde & Paivio, 1979; Clark & Paivio, 1987, 1991; Monzel et al., 2022, 2024). Moreover, our findings further support the applicability of DCT in cross-modal object colour property retrieval.

These findings highlight a fundamental difference in how object colour properties are processed between individuals with and without visual imagery. They also call into question existing theories of object colour processing developed based solely on data from neurotypical populations. More broadly, the absence of visual imagery may influence how humans process semantic concepts. Beyond visual aphantasia, future research into other forms of multisensory aphantasia could help clarify how imagery in different sensory modalities contributes to semantic processing (Dance et al., 2021; Dawes et al., 2024; Palermo et al., 2022; Takahashi et al., 2023; Zeman et al., 2016).

While the present study offers preliminary evidence for the role of visual imagery in retrieving object colour information under cross-modal conditions, it is limited by a relatively small sample size and focuses solely on individuals with visual aphantasia. Future research could design experiments targeting different subtypes of aphantasia (Keogh & Pearson, 2024; Palermo et al., 2022; Takahashi et al., 2023; Zeman et al., 2016). Moreover, traditional psychological theories—developed for individuals with intact imagery—may not fully capture the cognitive behaviours of individuals with congenital aphantasia. Regardless, studying aphantasia provides a unique opportunity to refine and expand these theories. In the future, combining larger sample sizes with neuroimaging techniques such as electroencephalography (EEG) and magnetic resonance imaging (MRI) will provide more objective and compelling evidence regarding the role of visual imagery in cognition.

Conclusions

In conclusion, our findings indicate that while the absence of visual imagery as a cognitive strategy does not impair the accuracy of object colour properties retrieval, it can reduce the efficiency of this process. Furthermore, this study provides further support for DCT and demonstrates that visual representations can facilitate cross-modal retrieval of object colour properties under word retrieval conditions.

Acknowledgments

We thank all the research participants for their patience and time. We would also like to express our sincere appreciation to Professor Yanchao Bi and her lab members for their generous provision of the Chinese-translated version of the measurement questionnaire employed in this study.

CRediT authorship contribution statement

Zhenjiang Cui: Conceptualization, Methodology, Writing - original draft, Writing - review & editing,

Investigation, Visualization, Software, Data curation, Formal analysis. **Xiangqi Luo:** Investigation, Software, Conceptualization, Data curation. **Yuxin Liu:** Investigation, Data curation. **Minhong Zhu:** Investigation, Data curation. **Zhiyun Dai:** Conceptualization, Investigation. **Xuliang Zhang:** Conceptualization, Investigation. **Zaizhu Han:** Conceptualization, Methodology, Resources, Investigation, Formal analysis, Writing - original draft, Writing - review and editing, Supervision, Funding acquisition.

Data availability statement

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Major Projects of National Social Science Foundation [grant number 24&ZD252] and National Natural Science Foundation of China [grant numbers 32271091 and 82372555].

Ethical approval

This study was approved by the Institutional Review Board of the National Key Laboratory of Cognitive Neuroscience and Learning at Beijing Normal University.

References

Arterberry, M. E., Milburn, M. M., Loza, H. L., & Willert, A. S. (2001). Retrieval of episodic information from memory: Comparisons among 3- and 4-year-olds, 7- and 8-year-olds, and adults. *Journal of Cognition and Development*, *2*(3), 283–305. https://doi.org/10.1207/S15327647JCD0203_3

Bainbridge, W. A., Pounder, Z., Eardley, A. F., & Baker, C. I. (2021). Quantifying aphantasia through drawing: Those without visual imagery show deficits in object but not spatial memory. *Cortex*, *135*, 159–172. https://doi.org/10.1016/j.cortex.2020.11.014

Bates, K. E., & Farran, E. K. (2021). Mental imagery and visual working memory abilities appear to be unrelated in childhood: Evidence for individual differences in strategy use. *Cognitive Development*, *60*, 101120. https://doi.org/10.1016/j.cogdev.2021.101120

Bi, Y. (2021). Dual coding of knowledge in the human brain. *Trends in Cognitive Sciences*, *25*(10), 883–895. https://doi.org/10.1016/j.tics.2021.07.006

Blazhenkova, O., & Kozhevnikov, M. (2009). The new object-spatial-verbal cognitive style model: Theory and measurement. *Applied Cognitive Psychology*, *23*(5), 638–663. https://doi.org/10.1002/acp.1473

Blazhenkova, O., & Pechenkova, E. (2019). The two eyes of the blind mind: Object vs. Spatial aphantasia? *The Russian Journal of Cognitive Science*, *6*(4), 51–62. https://doi.org/10.47010/19.4.5

Cherry, K. E., Dokey, D. K., Reese, C. M., & Brigman, S. (2003). Pictorial illustrations enhance memory for sentences in younger and older adults. *Experimental Aging Research*, *29*(3), 353–370. https://doi.org/10.1080/03610730303720

Clark, J. M., & Paivio, A. (1987). A dual coding perspective on encoding processes. In *Imagery and related mnemonic processes* (pp. 5–33). Springer New York. https://doi.org/10.1007/978-1-4612-4676-3_1.

Curran, T., & Doyle, J. (2011). Picture superiority doubly dissociates the ERP correlates of recollection and familiarity. *Journal of Cognitive Neuroscience*, *23*(5), 1247–1262. https://doi.org/10.1162/jocn.2010.21464

Dance, C. J., Hole, G., & Simner, J. (2023). The role of visual imagery in face recognition and the construction of facial composites. Evidence from aphantasia. *Cortex*, *167*, 318–334. https://doi.org/10.1016/j.cortex.2023.06.015

Dance, C. J., Ward, J., & Simner, J. (2021). What is the link between mental imagery and sensory sensitivity? Insights from aphantasia. *Perception*, *50*(9), 757–782. https://doi.org/10.1177/03010066211042186

D'Angiulli, A. (2002). Mental image generation and the contrast sensitivity function. *Cognition*, *85*(1), B11–B19. https://doi.org/10.1016/S0010-0277(02)00075-6

Dawes, A. J., Keogh, R., Andrillon, T., & Pearson, J. (2020). A cognitive profile of multi-sensory imagery, memory and dreaming in aphantasia. *Scientific Reports*, *10*(1), 10022. https://doi.org/10.1038/s41598-020-65705-7

Dawes, A. J., Keogh, R., & Pearson, J. (2024). Multisensory subtypes of aphantasia: Mental imagery as supramodal perception in reverse. *Neuroscience Research*, *201*, 50–59. https://doi.org/10.1016/j.neures.2023.11.009

Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M. (2018). Multipic: A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology*, *71*(4), 808–816. https://doi.org/10.1080/17470218.2017.1310261

Gleason, T. R., Fiske, K. E., & Chan, R. K. (2004). The verbal nature of representations of the canonical colors of objects. *Cognitive Development*, *19*(1), 1–14. https://doi.org/10.1016/S0885-2014(03)00044-3

Hargis, G. (1978). The function of imagery in word recognition development. *The Reading Teacher*, *31*, 870–874. https://doi.org/10.7287/peerj.preprints.2719v1

Huang, L., Xu, Q., Zhang, J., Li, K., Wang, J., Jia, H., & Ji, M. (2025). *Development and validation of a contextual-interactive spatial visualization Test (Cisvt)*. https://doi.org/10.2139/ssrn.5096889

Hurlburt, R. T., & Akhter, S. A. (2006). The descriptive experience sampling method. *Phenomenology and the Cognitive Sciences*, *5*(3–4), 271–301. https://doi.org/10.1007/s11097-006-9024-0

Jacobs, C., Schwarzkopf, D. S., & Silvanto, J. (2018). Visual working memory performance in aphantasia. *Cortex*, *105*, 61–73. https://doi.org/10.1016/j.cortex.2017.10.014

Kay, L., Keogh, R., Andrillon, T., & Pearson, J. (2022). The pupillary light response as a physiological index of aphantasia, sensory and phenomenological imagery strength. *eLife*, *11*, e72484. https://doi.org/10.7554/eLife.72484

Kay, L., Keogh, R., & Pearson, J. (2024). Slower but more accurate mental rotation performance in aphantasia linked to differences in cognitive strategies. *Consciousness and Cognition*, *121*, 103694. https://doi.org/10.1016/j.concog.2024.103694

Keitz, S. M. (1976). Age differences in adults' free recall of pictorial and word stimuli. *Educational Gerontology*, *1*(3), 237–241. https://doi.org/10.1080/0360127760010303

Keogh, R., & Pearson, J. (2018). The blind mind: No sensory visual imagery in aphantasia. *Cortex*, *105*, 53–60. https://doi.org/10.1016/j.cortex.2017.10.012

Keogh, R., & Pearson, J. (2024). Revisiting the blind mind: Still no evidence for sensory visual imagery in individuals with aphantasia. *Neuroscience Research*, *S0168010224000129*), https://doi.org/10.1016/j.neures.2024.01.008

Keogh, R., Wicken, M., & Pearson, J. (2021). Visual working memory in aphantasia: Retained accuracy and capacity with a different strategy. *Cortex*, *143*, 237–253. https://doi.org/10.1016/j.cortex.2021.07.012

Keysers, C., Gazzola, V., & Wagenmakers, E.-J. (2020). Using Bayes factor hypothesis testing in neuroscience to establish evidence of absence. *Nature Neuroscience*, *23*(7), 788–799. https://doi.org/10.1038/s41593-020-0660-4

Kozhevnikov, M., Kosslyn, S., & Shephard, J. (2005). Spatial versus object visualizers: A new characterization of visual cognitive style. *Memory & Cognition*, *33*(4), 710–726. https://doi.org/10.3758/BF03195337

Linde, J. T., & Paivio, A. (1979). Symbolic comparison of color similarity. *Memory & Cognition*, *7*(2), 141–148. https://doi.org/10.3758/BF03197594

Liu, J., & Bartolomeo, P. (2023). *Probing the unimaginable: The impact of aphantasia on distinct domains of visual mental imagery and visual perception* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/wzy5b

Liu, J., Zhan, M., Hajhajate, D., Spagna, A., Dehaene, S., Cohen, L., & Bartolomeo, P. (2023). *Visual mental imagery in typical imagers and in aphantasia: A millimeter-scale 7-T fMRI study* [Preprint]. Neuroscience. https://doi.org/10.1101/2023.06.14.544909

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, *64*(1), 17–24. https://doi.org/10.1111/j.2044-8295.1973.tb01322.x

Marre, Q., Huet, N., & Labeye, E. (2021). Embodied mental imagery improves memory. *Quarterly Journal of Experimental Psychology*, *74*(8), 1396–1405. https://doi.org/10.1177/17470218211009227

McBride, D. M., & Anne Dosher, B. (2002). A comparison of conscious and automatic memory processes for picture and word stimuli: A process dissociation analysis. *Consciousness and Cognition*, *11*(3), 423–460. https://doi.org/10.1016/S1053-8100(02)00007-7

Meng, M., Chang, S., Zhang, X., & Pearson, J. (2023). *Imageless imagery in aphantasia: Decoding non-sensory imagery in aphantasia* [Preprint]. In Review. https://doi.org/10.21203/rs.3.rs-3162223/v1

Meteyard, L., Cuadrado, S. R., Bahrami, B., & Vigliocco, G. (2012). Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex*, *48*(7), 788–804. https://doi.org/10.1016/j.cortex.2010.11.002

Milton, F., Fulford, J., Dance, C., Gaddum, J., Heuerman-Williamson, B., Jones, K., Knight, K. F., MacKisack, M., Winlove, C., & Zeman, A. (2021). Behavioral and neural signatures of visual imagery vividness extremes: Aphantasia versus hyperphantasia. *Cerebral Cortex Communications*, *2*(2), tgab035. https://doi.org/10.1093/texcom/tgab035

Monzel, M., Dance, C., Azañón, E., & Simner, J. (2023). Aphantasia within the framework of neurodivergence: Some preliminary data and the curse of the confidence gap. *Consciousness and Cognition*, *115*, 103567. https://doi.org/10.1016/j.concog.2023.103567

Monzel, M., Handlogten, J., & Reuter, M. (2024). No verbal overshadowing in aphantasia: The role of visual imagery for the verbal overshadowing effect. *Cognition*, *245*, 105732. http://dx.doi.org/10.1016/j.cognition.2024.105732

Monzel, M., Keidel, K., & Reuter, M. (2021). Imagine, and you will find – lack of attentional guidance through visual imagery in aphantasics. *Attention, Perception, & Psychophysics*, *83*(6), 2486–2497. https://doi.org/10.3758/s13414-021-02307-z

Monzel, M., Vetterlein, A., & Reuter, M. (2022). Memory deficits in aphantasics are not restricted to autobiographical memory – Perspectives from the Dual Coding Approach. *Journal of Neuropsychology*, *16*(*2*), 444–461. http://dx.doi.org/10.1111/jnp.v16.2

Muraki, E. J., Speed, L. J., & Pexman, P. M. (2023). Insights into embodied cognition and mental imagery from aphantasia. *Nature Reviews Psychology*, *2*(10), 591–605. https://doi.org/10.1038/s44159-023-00221-9

Paivio, A. (1963). Learning of adjective-noun paired associates as a function of adjective-noun word order and noun abstractness. *Canadian Journal of Psychology / Revue Canadienne de Psychologie*, *17*(4), 370–379. https://doi.org/10.1037/h0083277

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology / Revue Canadienne de Psychologie*, *45*(3), 255–287. https://doi.org/10.1037/h0084295

Paivio, A., & Csapo, K. (1973). Picture superiority in free recall: Imagery or dual coding? *Cognitive Psychology*, *5*(2), 176–206. https://doi.org/10.1016/0010-0285(73)90032-7

Palermo, L., Boccia, M., Piccardi, L., & Nori, R. (2022). Congenital lack and extraordinary ability in object and spatial imagery: An investigation on sub-types of aphantasia and hyperphantasia. *Consciousness and Cognition*, *103*, 103360. https://doi.org/10.1016/j.concog.2022.103360

Pounder, Z., Jacob, J., Evans, S., Loveday, C., Eardley, A. F., & Silvanto, J. (2022). Only minimal differences between individuals with congenital aphantasia and those with typical imagery on neuropsychological tasks that involve imagery. *Cortex*, *148*, 180–192. https://doi.org/10.1016/j.cortex.2021.12.010

Reisberg, D., Pearson, D. G., & Kosslyn, S. M. (2003). Intuitions and introspections about imagery: The role of imagery experience in shaping an investigator's theoretical views. *Applied Cognitive Psychology*, *17*(2), 147–160. https://doi.org/10.1002/acp.858

Rich, A. N., Williams, M. A., Puce, A., Syngeniotis, A., Howard, M. A., McGlone, F., & Mattingley, J. B. (2006). Neural correlates of imagined and synaesthetic colours. *Neuropsychologia*, *44*(14), 2918–2925. https://doi.org/10.1016/j.neuropsychologia.2006.06.024

Roebuck, H., & Lupyan, G. (2020). The internal representations questionnaire: Measuring modes of thinking. *Behavior*

Research Methods, *52*(5), 2053–2070. https://doi.org/10.3758/s13428-020-01354-y

Speed, L. J., Eekhof, L. S., & Mak, M. (2024). The role of visual imagery in story reading: Evidence from aphantasia. *Consciousness and Cognition*, *118*, 103645. http://dx.doi.org/10.1016/j.concog.2024.103645

Striem-Amit, E., Wang, X., Bi, Y., & Caramazza, A. (2018). Neural representation of visual concepts in people born blind. *Nature Communications*, *9*(1), 5250. https://doi.org/10.1038/s41467-018-07574-3

Sun, Y., Xue, X., Li, Z., Ma, H., & Zhang, D. (2023). Association of visual conscious experience vividness with human cardiopulmonary function. *Stress and Brain*, *3*(2), 80–95. https://doi.org/10.26599/SAB.2022.9060033

Suzuki, K., & Takahashi, R. (1997). Effectiveness of color in picture recognition memory. *Japanese Psychological Research*, *39*(1), 25–32. https://doi.org/10.1111/1468-5884.00033

Takahashi, J., Saito, G., Omura, K., Yasunaga, D., Sugimura, S., Sakamoto, S., Horikawa, T., & Gyoba, J. (2023). Diversity of aphantasia revealed by multiple assessments of visual imagery, multisensory imagery, and cognitive style. *Frontiers in Psychology*, *14*, 1174873. https://doi.org/10.3389/fpsyg.2023.1174873

Vandenberghe, R., Price, C., Wise, R., Josephs, O., & Frackowiak, R. S. J. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature*, *383*(6597), 254–256. https://doi.org/10.1038/383254a0

Vincent, B. T. (2015). A tutorial on Bayesian models of perception. *Journal of Mathematical Psychology*, *66*, 103–114. https://doi.org/10.1016/j.jmp.2015.02.001

Wang, X., Men, W., Gao, J., Caramazza, A., & Bi, Y. (2020). Two forms of knowledge representations in the human brain. *Neuron*, *107*(2), 383–393.e5. https://doi.org/10.1016/j.neuron.2020.04.010

Wicken, M., Keogh, R., & Pearson, J. (2021). The critical role of mental imagery in human emotion: Insights from fear-based imagery and aphantasia. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1946), 20210267. https://doi.org/10.1098/rspb.2021.0267

Yui, L., Ng, R., & Perera, W. A. H. (2017). *Concrete vs abstract words – what do you recall better? A study on dual coding theory*. https://doi.org/10.7287/peerj.preprints.2719v1

Zeman, A. Z. J., Della Sala, S., Torrens, L. A., Gountouna, V.-E., McGonigle, D. J., & Logie, R. H. (2010). Loss of imagery phenomenology with intact visuo-spatial task performance: A case of 'blind imagination'. *Neuropsychologia*, *48*(1), 145–155. https://doi.org/10.1016/j.neuropsychologia.2009.08.024

Zeman, A., Dewar, M., & Della Sala, S. (2015). Lives without imagery – congenital aphantasia. *Cortex*, *73*, 378–380. https://doi.org/10.1016/j.cortex.2015.05.019

Zeman, A., Dewar, M., & Della Sala, S. (2016). Reflections on aphantasia. *Cortex*, *74*, 336–337. https://doi.org/10.1016/j.cortex.2015.08.015

Zeman, A., Milton, F., Della Sala, S., Dewar, M., Frayling, T., Gaddum, J., Hattersley, A., Heuerman-Williamson, B., Jones, K., MacKisack, M., & Winlove, C. (2020). Phantasia–The psychological significance of lifelong visual imagery vividness extremes. *Cortex*, *130*, 426–440. https://doi.org/10.1016/j.cortex.2020.04.003

Zhou, W., Li, P., Lei, X., & Yh, Y. (2020). *Positive mental imagery in the intervention of negative mood: A randomized controlled trial*. https://doi.org/10.21203/rs.3.rs-60071