CORTEX 169 (2023) 220–234

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Viewpoint

Defining and 'diagnosing' aphantasia: Condition or individual difference?

Andrea Blomkvist [a,*] *and David F. Marks* [b]

[a] *Centre for Philosophy of Natural and Social Sciences, Department of Philosophy, Logic and Scientific Method, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK*
[b] *13200 Arles, Bouches-du-Rhône, Provence-Alpes-Côte d'Azur, France*

ARTICLE INFO

Article history:
Received 5 April 2023
Reviewed 05 June 2023
Revised 16 June 2023
Accepted 26 September 2023
Action editor Paolo Partolomeo
Published online 29 September 2023

Keywords:
Aphantasia
VVIQ
Visual imagery
Mental imagery
Vividness
Low-imagery
No-imagery
Methodology
Diagnosis
Individual differences

ABSTRACT

Research into the newly-coined 'condition' of 'aphantasia', an individual difference involving the self-reported absence of voluntary visual imagery, has taken off in recent years, and more and more people are 'self-diagnosing' as aphantasic. Yet, there is no consensus on whether aphantasia should really be described as a 'condition', and there is no battery of psychometric instruments to detect or 'diagnose' aphantasia. Instead, researchers currently rely on the *Vividness of Visual Imagery Questionnaire* (VVIQ) to 'diagnose' aphantasia. We review here fundamental and methodological problems affecting aphantasia research stemming from an inadequate focus on how we should define aphantasia, whether aphantasia is a pathological condition, and the extensive use of VVIQ as a 'diagnostic test' for aphantasia. Firstly, we draw attention to 'literature blindness' for visual imagery research from the 1960s–1990s concerning individual differences in visual imagery vividness. Secondly, despite aphantasia being defined as a 'condition' where voluntary visual imagery is *absent* as indicated by the lowest score on the VVIQ, aphantasia studies inconsistently employ samples comprised of a mixture of participants with no visual imagery and low visual imagery, and we argue that this hinders the uncovering of the underlying cause of aphantasia. Thirdly, the scores used to designate the boundary between aphantasia and non-aphantasia are arbitrary and differ between studies, compromising the possibility for cross-study comparison of results. Fourthly, the problems of 'diagnosing' aphantasia are not limited to the academic sphere, as one can 'self-diagnose' online, for example by using the variant-VVIQ on the Aphantasia Network website. However, the variant-VVIQ departs from the original in ways likely to impact validity and accuracy, which could lead people to falsely believe they have been 'diagnosed' with aphantasia by a scientifically-validated measure. Fifthly, we discuss the hypothesis that people who believe they have been 'diagnosed' with aphantasia might be vulnerable to health anxiety, distress, and stigma. We conclude with a discussion about some fundamental aspects of how to classify a disorder, and suggest the need for a new psychometric measure of aphantasia.

* *Corresponding author*. Centre for Philosophy of Natural and Social Sciences, Department of Philosophy, Logic and Scientific Method, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK.
E-mail address: a.blomkvist@lse.ac.uk (A. Blomkvist).

https://doi.org/10.1016/j.cortex.2023.09.004

Glossary of terms

Acquired aphantasia	aphantasia that occurs following brain injury, trauma, and, occasionally, other psychological conditions
AN	Aphantasia Network
AN-varVVIQ	a variant-VVIQ that appears on the AN's website.
Aphantasia	an individual difference commonly defined as the inability to generate conscious voluntary visual imagery
Aphantasic	a person with aphantasia
Binocular rivalry	a phenomenon of visual perception in which perception alternates between different images presented to each eye
Congenital aphantasia	a commonly defined individual difference that is assumed to present from birth
External validity	the quality of being sound and accurate as a measure of something
Self-report	a person's description of a mental experience
Vividness	in reference to a mental image, vividness is a combination of clarity and liveliness
VVIQ	acronym for the *Vividness of Visual Imagery Questionnaire*, a test of 16 items used for the measurement of individual differences in the vividness of individuals' visual mental imagery (Marks, 1973a)
VVIQ-2	acronym for the *Vividness of Visual Imagery Questionnaire-2*, an extended version of the VVIQ having 32 items and reversed rating scale (Marks, 1995)
Z-varVVIQ	the variant of the VVIQ used by Adam Zeman and associates (Zeman et al., 2020)

1. Introduction

Research into the newly-coined 'condition' of 'aphantasia' has taken off in recent years. Aphantasia is commonly defined as a 'condition' where people have no voluntary visual imagery (Dance et al., 2021a, b; Dawes et al., 2020; Greenberg & Knowlton, 2014; Keogh & Pearson, 2018; Milton et al., 2021; Pearson, 2019; Zeman et al., 2020), and it is commonly divided into two kinds: congenital aphantasia (Zeman et al., 2015), which is assumed to be present from birth,[1] and acquired aphantasia, which may occur following brain injury (Brain, 1954; Charcot & Bernard, 1883; Thorudottir et al., 2020), psychological or psychiatric conditions (de Vito & Bartolomeo, 2016), or as an unusual side-effect of procedures such as coronary angioplasty (Zeman et al., 2010).

Various deficiencies and differences in visual imagery were known before the term 'aphantasia' was coined in 2015 (Bartolomeo et al., 2002; Moro et al., 2008; and for an overview see Liu et al., 2022; Zeman, 2015). Studying these loss of function cases are important for establishing the function of visual imagery, as well as its relation to vision and its neurological underpinnings. For example, Basso et al. (1980) describe a patient with a visual imagery deficits who on a CT scan showed an area of reduced density in the left occipital lobe and left temporal lobe, and Riddoch et al. (1990) describe a patient with a deficit in imagery generation despite no apparent occipital damage. Certain patients have also been reported to have more specific imagery deficits, for example in generating imagery of object form and colour (Farah et al., 1988) (this contrast with aphantasia which is normally taken to be domain-general).[2] In contrast to these cases, the vast majority of cases involves apparent congenital aphantasia where there is no known neurological damage or previous trauma, and hence provides a way to study the function of visual imagery in healthy individuals. The form of aphantasia labelled 'congenital aphantasia' is the focus of this viewpoint paper.

The most common way to establish whether a person has aphantasia is by administering the *Vividness of Visual Imagery Questionnaire* (VVIQ) (Marks, 1973a), and variations of this questionnaires can be found online for people to 'self-diagnose' (e.g., see the questionnaire on the Aphantasia Network, https://aphantasia.com/vviq/). This questionnaire aims to establish the vividness of someone's voluntary visual imagery by asking them to visualise different scenes and rate the vividness; a minimum or very low score on the VVIQ is commonly taken to indicate that a person has aphantasia.

In this viewpoint, we will argue that this practice is problematic, as the VVIQ is currently being used to evaluate cases of aphantasia in inconsistent ways across studies. This could potentially lead to false positives where individuals are 'diagnosed' as aphantasic, which could have negative consequences for a person's well-being (Monzel, 2022). To move forward in aphantasia research, we suggest that further resources need to be put towards investigating the questions of: i) whether there is a cluster of symptoms exhibited by aphantasic individuals that could signify any kind of disorder, and ii) how the well-being of individuals given an aphantasia label are affected by this label. Further, we will suggest that a new psychometric tool targeting aphantasia is necessary, as the VVIQ only measures the reduction in the vividness of voluntary visual imagery and not in any other sensory modalities, nor any other cognitive functions.

Firstly, we draw attention to a large body of research from the 1960s to the 1990s document multiple cases of participants with low or no imagery ability using the VVIQ (Marks, 1972, 1973a, 1973b, 1986; Isaac & Marks, 1994). But this body of

[1] It should be noted that although one form of aphantasia is defined as being present at birth, no infants or children have been tested, and no longitudinal studies have yet been conducted to distinguish between the acquired and congenital forms of aphantasia.

[2] For an overview of neurological cases of reduced general visual imagery, or reduced visual imagery pertaining to one modality, see Bartolomeo (2002).

research is not appealed to by modern day researchers of aphantasia despite its relevance to the field (§1). Secondly, despite aphantasia being defined as a 'condition' where people have *no* voluntary visual imagery, aphantasia samples in studies comprise a mixture of participants with no visual imagery *or* low visual imagery. This practice could hinder the possibility of finding an underlying mechanism responsible for causing aphantasia (§2). Thirdly, cross-study comparison is problematic because the scores used to designate the boundary between aphantasia and non-aphantasia are arbitrary and differ from study to study (§3). Fourthly, the problem of 'diagnosing' aphantasia is not limited to what cut-off point on the VVIQ to use, but also which version of the VVIQ to use. As aphantasia has received a lot of media attention, various ways of 'diagnosing' aphantasia are available online (§4). In particular, the popular Aphantasia Network claims that their version of VVIQ can tell a person whether they have aphantasia. However, their unvalidated version of the VVIQ departs from the original in multiple ways which likely impacts the validity of the measure. The practice of modifying the VVIQ is also present in academia, and version of the VVIQ are used in multiple studies with unknown impacts on the measure's validity (Zeman et al., 2015, 2020). Hence, the use of the VVIQ in aphantasia research is problematic with both respect to cut-off points and unvalidated versions. Fifthly, we believe that this should be a concern to the research community as it could lead people to falsely believe that they have aphantasia, which could have negative effects on one's wellbeing (Monzel et al., 2022) (§5). Finally, going forward, we suggest that we consider whether there are any grounds for classifying aphantasia as a 'disorder', and the potential impact this would have on people's lives. We also suggest that in order to solve many of the problems raised in this article, a specialized psychometric measure ought to be developed for the assessment of potential cases of aphantasia (§6).

As this article centres around the use of the *Vividness of Visual Imagery Questionnaire* (VVIQ) in aphantasia research, we will first briefly review the VVIQ (Marks 1973) and the VVIQ-2 (Marks, 1995). This will facilitate the comparison of these measures to the Z-varVVIQ and the AN-varVVIQ, which we discuss later in the article (§4). The VVIQ is a procedure for the measurement of individual differences in the vividness of individuals' voluntary visual mental imagery. The VVIQ, with 16 items, and the VVIQ-2, with 32 items, are available in multiple languages and have been made freely available for researchers since their first publication.

The five-point rating scale of the VVIQ is presented below. Some researchers reverse the numerical scale to make 5 = perfectly clear and as vivid as normal vision, and 1 = no image at all, you only "know" that you are thinking of an object, as was the case for the VVIQ-2 (Table 1).

The VVIQ consists of 16 items that a person should attempt to visualise. The 16 items are arranged in blocks of four, in which each has a theme (i.e., a person, natural scenery, a shop). Each theme is provided with a narrative to guide a progression of visual imagery. At least one item in each cluster, eight items in total, describes an activity or movement, indexing liveliness. The aim of the VVIQ is to assess visual imagery vividness under conditions which allow a progressive development of scenes, situations, or events as naturally as possible. The items are intended to evoke sufficient interest, meaning, and affect conducive to image generation. Participants rate the vividness of their images separately with eyes open and eyes closed. There are four sets of items with four items per set. The first set is described as follows (see Supplementary Material for the full original VVIQ):

Table 1 – VVIQ ratings (Marks 1973).

Rating	The Image Aroused by an Item Might Be
1	Perfectly clear and as vivid as normal vision
2	Clear and reasonably vivid
3	Moderately clear and vivid
4	Vague and dim
5	No image at all, you only "know" that you are thinking of an object

> "For items 1 to 4, think of some relative or friend whom you frequently see (but who is not with you at present) and consider carefully the picture that comes before your mind's eye.
>
> 1. The exact contour of face, head, shoulders, and body.
>
> 2. Characteristic poses of head, attitudes of body, etc.
>
> 3. The precise carriage, length of step, etc. in walking.
>
> 4. The different colours worn in some familiar clothes."

The VVIQ continues with 12 more items. The participants' responses to the questionnaire are scored by calculating the numerical total of the 32 obtained vividness ratings (16 ratings with eyes closed and 16 ratings with eyes open). Research has shown that the VVIQ has excellent psychometric reliability (test-retest and split-half reliability) and it has been validated in hundreds of independent studies (e.g., for reviews, see McKelvie, 1995; Campos & Pérez-Fabello, 2009; Marks, 2023).

2. Imagery research in the late 20th century

Aphantasia is described as a 'condition' in which people are unable to generate voluntary visual imagery (Dance et al., 2021a, b; Dawes et al., 2020; Greenberg & Knowlton, 2014; Keogh & Pearson, 2018; Milton et al., 2021; Pearson, 2019; Zeman et al., 2020) (but see Blomkvist, 2022). The authors prefer the term 'individual difference' as a descriptor of aphantasia rather than the term 'condition', which is loaded with the connotations of a medical diagnosis. People who report having no voluntary visual imagery are people who rate every single VVIQ item as "no image at all, you only 'know' that you are thinking of the object" while imaging with their eyes closed and, once again, with their eyes open.

The discovery of people who are unable to voluntarily summon visual imagery is normally attributed to Galton (1880). The second instance of aphantasia commonly mentioned is patient MX studied by Zeman et al. (2010). We

here aim to draw attention to a significant body of research from the 1950s and forwards that discusses individuals with low visual imagery, sometimes seemingly congenital and sometimes due to neurological impairments.

For example, in the period 1957–1975, a productive group of imagery researchers made substantial contributions to our understanding of individual differences in voluntary mental imagery which had been described by Galton (e.g., Barber, 1959; Gur & Hilgard, 1975; Haber, 1969; Richardson, 1969; Hebb, 1968; Segal & Gordon, 1969; Horowitz, 1970; Shepard and Metzler, 1971; Sheehan, 1972; Hatakeyama, 1974; McKellar, 1957; Paivio, 1969; Sarbin & Juhasz, 1967; West, 1962; Singer & Antrobus, 1963; Holt, 1964).

Marks (1972, 1973a, 1973b) investigated individual differences in voluntary visual imagery vividness with the VVIQ in multiple studies and, in so doing, identified within the large participant pool a few individuals (<.5% of the total) who adamantly stated that *they did not know what the terms 'mental imagery' or 'visual imagery' meant*. These individuals consistently used the rating '5' for 'no image at all' on the VVIQ. In the studies, they were considered to be 'VVIQ non-imagers', extreme responders in an approximately normal VVIQ distribution. As would be expected, there was an equal number of people at the opposite end of the distribution who consistently rated their visual imagery vividness as '1' (*Perfectly clear and as vivid as normal vision*), another set of extreme responders which in current day research are referred to as 'hyperphantasics' (Zeman et al., 2020). Despite its relevancy to contemporary aphantasia research, Marks's (1972, 1973a, 1973b) reporting of individuals who lacked voluntary visual imagery is not considered by present day researchers.

Another case of people who reported having no imagery on the VVIQ comes from Isaac and Marks (1994), who ran five studies with especially selected groups of participants. Study 1 involved a representative sample of 345 children aged from 7 to 16 years selected from four primary, intermediate and high schools in Dunedin, New Zealand. Isaac and Marks (1994) reported that: "0.3 per cent reported minimum vividness of visual imagery scores (criterion total score = 160) and .4 per cent minimum vividness of movement imagery scores (criterion total score = 240)." This finding of extremely high vividness and zero vividness VVIQ scorers had become a routine result in every study conducted by the team. To the best of our knowledge, in contemporary research on aphantasia, children have not yet been tested.

Consulting the past literature on participants with no imagery at times paints a different picture to modern day research and it hence deserves attention. As mentioned, several controlled studies compared the performance of people reporting high vividness on the VVIQ and people who reported low vividness on the VVIQ, and found that the former out-performed the latter on tasks involving picture recall and perception (e.g., Marks, 1972, 1973a, 1973b, 1995; Gur & Hilgard, 1975; Berger and Gaunitz, 1979; McKelvie, 1995). For example, Marks and Isaac were able to correlate self-reported VVIQ vividness with objective measures such as picture recall, picture recognition and associative learning. Research participants' VVIQ scores were found to be distributed from the maximum ($32 \times 5 = 160$) to the minimum ($32 \times 1 = 32$). In one experiment (Marks, 1973b) from an initial sample of 74 students, the 18 highest vividness VVIQ scorers (mean rating = 1.64) and the 18 lowest vividness VVIQ scorers (mean rating = 3.25) were selected to form two experimental groups. As predicted, in the picture recall tasks, groups of vivid VVIQ imagers produced significantly higher recall accuracy than non-vivid VVIQ imagers. Additionally, females scored more highly than males. A meta-analysis of the VVIQ as a psychometric instrument was published by McKelvie (1995). The findings from this period, which are mainly overlooked, have great relevance to the current attempts to differentiate the cognitive abilities of aphantasics from non-aphantasics (e.g., Pounder et al., 2022; Dance et al., 2021b; Dance et al., 2023; Dawes et al., 2022). These early studies clearly indicate reliable individual differences in recall of pictorial information between 'aphantasics' (loosely defined as having VVIQ scores <32) and non-aphantasics.

The picture recall task used by Marks (1972, 1973) successfully distinguished vivid and non-vivid VVIQ imagers, but other tasks also targeted voluntary visual imagery. From the widely-cited work of Allan Paivio (e.g., Paivio, 1969; Paivio et al., 1966) it was evident that, notwithstanding the wide individual differences in VVIQ vividness, there were well-established methods for visual imagery production and memory that were of universal application (e.g., see Luria, 1987; Yates and Kamboj, 2017). In associative memory, for example, where participants were required to associate a list of words with a sequence of places on a mnemonic walk, performance of selected groups of 'vivid' and 'non-vivid' VVIQ imagers (i.e., low visual imagers) were observed not to differ (Marks, 1972). Both vivid and non-vivid VVIQ participants were almost equally able to produce associative connections and recall words and without any significant gender differences. Thomas et al. (2022) have confirmed that aphantasic participants benefited from interactive imagery instructions as much as controls.

Hence, despite the recent literature suggesting that the individual difference of the self-reported inability to generate voluntary visual imagery in healthy individuals, which Galton discovered in 1880, was re-discovered in 2015, this phenomenon was in fact reported by several previous investigators including McKellar (1957), Marks (1972) and Isaac and Marks (1994), albeit without using the term 'aphantasia', which had not yet been coined.

3. Defining 'aphantasia'

Aphantasia is often defined as 'the inability to create visual images in one's mind' (Keogh & Pearson, 2018), the 'lifelong absence of visualisation' (Fulford et al., 2018), or the 'lifelong absence of mind's eye' (Zeman, 2020). With few exceptions (e.g., Zeman et al., 2015), aphantasia is defined as an a 'condition' where people have a *total lack* of voluntary visual imagery. There are several problems with these definitions. A recent paper by Blomkvist (2022) raises some of them, such as whether to define aphantasia as affecting all kinds of mental imagery (not just visual imagery) or whether aphantasia only applies to voluntary imagery (rather than involuntary imagery). Here, we call attention to further problems regarding the definition, and, as our focus is on how the VVIQ is used in

aphantasia research, will limit our discussion to how aphantasia is defined as affecting voluntary visual imagery.[3]

As aphantasia is most commonly defined as a 'condition' where people have a total lack of voluntary visual imagery, we would expect it to map onto the lowest score on the VVIQ as this denotes 'no image at all' (that is, '1' or '5', depending on whether a reversed scale is used). Hence, to study aphantasia, researchers ought to screen and select a sample of participants with the lowest score on the VVIQ and compare them to a control group with scores within the mid-range or high-range (depending on aim). Indeed, this is the pattern that the research initially followed. An oft-cited case of acquired aphantasia in recent times is the study of MX, who reported having no visual imagery at all, scoring 16 on the Z-varVVIQ, Zeman's modified version of the VVIQ (Zeman et al., 2010) (see §4 for details on Z-varVVIQ). Subsequently, the first study reporting on congenital aphantasia was a study of mostly participants who also scored 16 on the Z-varVVIQ (12/21) (Zeman et al., 2015).[4]

But in most aphantasia studies there has been a shift in how the sample is defined (an exception is Fulford et al., 2018) without a shift in the definitions of aphantasia offered. In particular, participants who score higher than 16 on the VVIQ are now included in the aphantasia sample. For example, Zeman (2020) and Milton (2021) include participants who score ≤23, Bainbridge et al.'s (2021) study includes participants scoring ≤25, and Keogh et al.'s (2021) and Wicken et al.'s (2021, p. 288) studies include participants who score ≤32. We contend that this is a problematic practice because there is no longer any indication that these participants have *no voluntary visual imagery*. In fact, they are reporting the opposite; they are reporting having visual imagery, albeit imagery that is dim and vague.

This is problematic for two reasons. Firstly, it is misleading as the definitions commonly state that aphantasia is a 'condition' where people have a complete lack of voluntary visual imagery. In many cases, samples do not line up with how the researchers themselves describe aphantasia in their own articles. For example, the definitions above all use phrases such as 'absence', 'inability' and 'lack', which suggests that the sample that they are interested in studying is the sample of people with *no* voluntary visual imagery. But the samples they have in fact studied include participants who report low voluntary visual imagery. Hence, they have studied participants who actually *can* create visual imagery.

In some cases, separate analyses have been conducted for the no-imagery group and the low-imagery group. For example, Dance et al., 2021a, b investigated the relationship between aphantasia and synaesthesia, recruiting participants from the online Synaesthesia Battery (Eagleman et al., 2007) which includes the VVIQ-2 (Marks, 1995). Here, participants scoring ≤64 on the VVIQ-2 (corresponding to having vague and dim imagery) were included in the sample. Results here showed that there were 196 people scoring ≤64 in the sample, and amongst these, there were 144 synaesthetes. For graphene-colour synaesthesia, there was a prevalence of 73.5% in the aphantasia group, and 77.5% in non-aphantasia group – a non-significant difference. But the study also independently analysed the results from the no imagery group and the low imagery group. In this case, there was also no significant difference between the groups when it came to the prevalence of synaesthesia.[5] Now, one could argue that given the fact that no significant difference was found between the groups here, we could collapse analysis across the two groups in general. We would caution against this, as these results only suggest that visual imagery does *not* play a role in graphene-colour synaesthesia, but they cannot inform us about whether the two groups would perform similarly on other tests.

One might think that this is much ado about nothing, and an easy remedy would be to broaden the definition of aphantasia to include weak visual imagery and hence define aphantasia on a spectrum from no visual imagery to low visual imagery. Indeed, this is in line with Zeman's original definition of aphantasia 'a condition of reduced or absent voluntary imagery' (Zeman et al., 2015). This brings us to the second problem, namely that we do not yet know the nature of the underlying mechanism that causes aphantasia, and so, we cannot know whether the mechanism that causes some people to lack voluntary visual imagery is the same mechanism as causes other people to experience low visual imagery (for an overview, see Craver & Tabery, 2019). Many capacities can exist on a spectrum, but different underlying mechanisms could be responsible for how these capacities manifest (Hornett, 2021). Take vision for example. We could model vision on a spectrum, from perfect vision to complete blindness. Naturally, people would fall on different parts of the spectrum, and many that fall towards the lower end would do so because of problems with the retina causing e.g., long-sightedness or short-sightedness. Some people would be classified as completely blind. But importantly, the underlying mechanism causing complete blindness need not be the same as the mechanism which causes someone to fall on the lower end of the spectrum (and there could even be different mechanisms causing complete blindness). We commonly distinguish between retinal blindness and cortical blindness, where one is due to issues with the retina, and the other issues with the visual cortices. These are different underlying mechanisms. But a simple spectrum from perfect vision to complete blindness does not capture this nuance, and could trick us into thinking that the mechanism is homogeneous, even though it is not.

What is the remedy? If we are interested in studying people who are completely blind, we ought *not* to include people that fall on the low-vision part of the spectrum in our sample, as this is a caused by a radically different mechanism compared to cases of cortical blindness. In fact, in this case, we ought to employ further methods of testing in order to ensure that

[3] Studies showing that aphantasia affects other kinds of imagery include Dawes et al. (2020) and Zeman et al. (2020).

[4] 9/21 participants reported substantially reduced imagery.

[5] The study also investigated whether projector or associator synaesthesia was more common in aphantasics compared to non-aphantasics, and found a similar pattern in both cases (that is, in both groups, around 90% were associators, and 10% were projectors). However, when treating projector and associator scores continuously, instead of as categories, it was found that scores from aphantasics were more associator-like than in non-aphantasics.

within the sample of blind participants, we can also distinguish between the ones who are cortically blind and the ones who are retinally blind. This illustrates how broadening a definition to include people that fall on different parts of a spectrum could inadvertently lead to studying a sample with heterogeneous underlying causes for their symptoms.

We contend that in the case of aphantasia it is even more important that we do *not* at this point adopt a lenient definition, because, unlike the case of blindness where we know about the different underlying causes, we are still agnostic as to the underlying causes of aphantasia (though some suggestions have been made: Blomkvist, 2022; Marks, 2023; Pearson, 2019).

4. Cross-study comparisons

Including people with low visual imagery in the aphantasia sample makes another problem apparent, namely, the arbitrary and varying VVIQ cut-off points for aphantasia. If we define aphantasia as a 'condition' (more accurately, an individual difference) where individuals have absent or *reduced* voluntary imagery, as suggested by Zeman in 2015, the question immediately arises of where the cut-off point for having 'reduced' imagery lies. We have found that researchers generally do not offer any independent reason for their selected cut-off points and that they can vary greatly (e.g., =16 in Zeman et al. (2015), ≤25 in Bainbridge et al. (2021), and ≤32 in Wicken et al. (2021, p. 288)), and we believe that this practice can hamper cross-study comparison.[6]

Having arbitrary varying cut-off points causes problems for cross-study comparisons. Multiple studies now claim that we know a fair bit about aphantasia, for example, aphantasia is accompanied lower recall of episodic memory details (Bainbridge et al., 2021; Dawes et al., 2020, 2022), largely retained working memory (Jacobs et al., 2018; Keogh & Pearson, 2021), and reduced emotional reaction to visually imagined scenarios (Wicken et al., 2021). But most of these studies use different cut-off points, raising the question of whether all these findings can be generalised to the population with aphantasia. Putting the concerns from §2 aside, there can be a 16-point difference between purported aphantasic participants from different studies, and yet, these studies claim to be studying the same 'condition' and findings are amalgamated in many literature reviews. Indeed, the difference in performance between groups have been documented in a study by Zeman et al. (2020) which distinguished between 'moderate aphantasia' (17–23) and 'extreme aphantasia' (=16) and found statistically significant differences between these groups in relation to the experience of brief flashes of visual imagery. An even bigger difference can be expected for samples that sit 16 points apart on the scale, and hence, we believe that we cannot easily amalgamate findings from different studies.

A further problem caused by the arbitrary cut-off points relates to determining the prevalence of aphantasia. Dance et al. (2022) note that different studies have reached different conclusions here, partly due to using different cut-off points. For example, Faw et al. (2009) who take aphantasia to mean no imagery, report a prevalence of 2.1% in a sample of 750 participants. Out of a sample of 1288 participants, Zeman et al. (2020) report that .7% have no imagery and 2.6% have low imagery (VVIQ score ≤23). In their own study, where they survey 502 participants and take aphantasia to be indicated by a score of ≤32 on the VVIQ, Dance et al. (2022) find 4.2% have aphantasia. We can clearly see that all these studies do not include equivalent samples and the differing results should come as no surprise. But this point can be easily overlooked as the studies all purportedly study aphantasia. Recently, despite the sample differences, these studies amongst others were all included in a meta-analysis into the prevalence of aphantasia, totalling 2693 participants (Monzel et al., 2022). This meta-analysis concludes that the prevalence of aphantasia is 3.5%. But, as noted, we cannot easily compare results from these studies, as samples have been demarcated in different ways. Participants who scored between 24 and 32 would not have counted as aphantasics in Zeman et al.'s study, but would have done so in Dance et al.'s study. As the sampling criteria differ, we cannot conduct a meta-analysis of the results to find out about the prevalence of aphantasia, and hence, the arbitrary and differing cut-off points hinder cross-study comparisons and the prospect of conducting meta-analyses. The research by Marks (1973a) and Isaac and Marks (1994) indicates that congenital aphantasia defined as the complete absence of voluntary imagery has a prevalence no higher than .05 percent (one in 200).

5. 'Diagnosing' aphantasia with variant-VVIQs

The problems of how to 'diagnose' aphantasia are not limited to its definition, but also include which version of the VVIQ to use. Variations of the VVIQ commonly used to 'diagnose' aphantasia, often occur outside research contexts. Aphantasia has received a lot of attention in media, and various ways of 'diagnosing' aphantasia are available online. In particular, the Aphantasia Network website (AN) (https://aphantasia.com/vviq/) claims that their variant VVIQ (AN-varVVIQ) can tell a person whether they have aphantasia. According to the AN, close to 500,000 tests have been taken, but their version of the VVIQ departs from the original VVIQ in multiple ways. We believe that these ways are likely to deleteriously influence the validity of the measure, and to the best of the authors' knowledge, the Aphantasia Network's questionnaire has not been validated.[7] *The authors are also aware that the AN's version of the VVIQ (the AN-varVVIQ) has been altered on several occasions, which means that the 'aphantasic' population contains individuals who have been 'diagnosed' using different criteria depending on when they completed the AN-varVVIQ.* To highlight the problems of AN-varVVIQ, a summary of the main differences between the VVIQ and the AN-varVVIQ follows.

[6] Note that some have cited past studies in support and pointed out that their chosen cut-off point corresponds to imagery reported as 'dim and vague' on the VVIQ (Dance et al., 2021).

[7] The basic psychometric requirements for questionnaires and surveys are described in Chapter 8 of Marks and Yardley (2004).

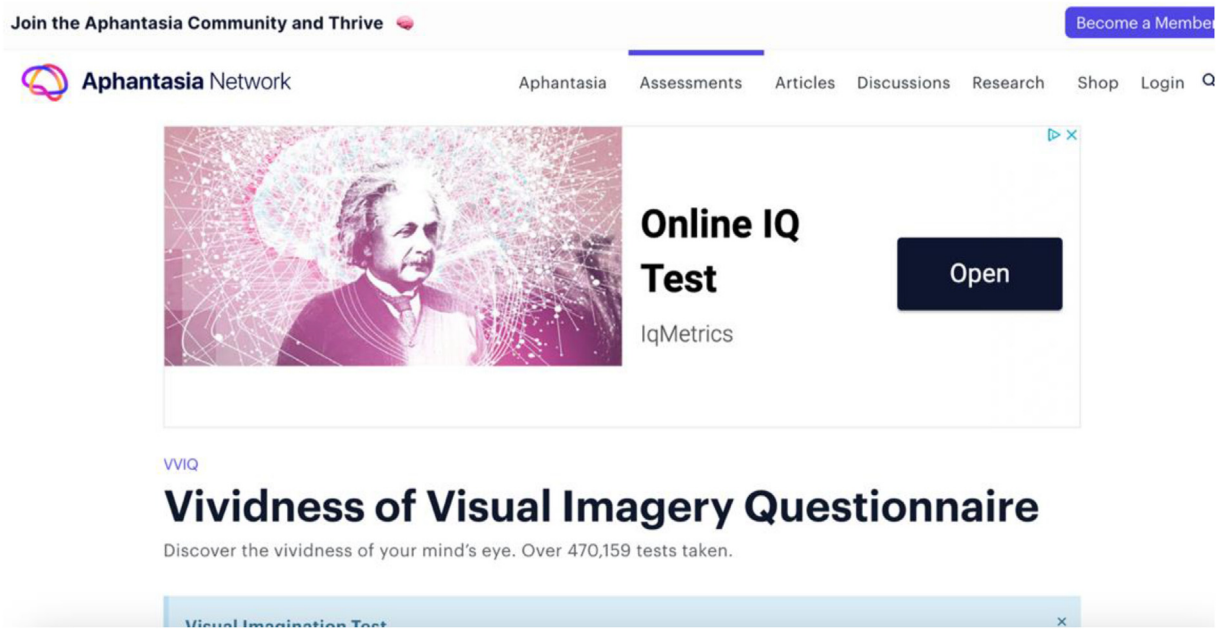

Fig. 1 – The representation of the Aphantasia Network's variant of the VVIQ as the "Vividness of Visual Imagery Questionnaire" on the Aphantasia Network website at https://aphantasia.com/vviq/ Accessed on 3 November 2023 (the image of Einstein is still there!).

Firstly, the framing of the AN-varVVIQ has connotations that are inconsistent with the unassuming and prosaic introduction to the original VVIQ, which states:

> "Visual imagery refers to the ability to visualize, that is, the ability to form mental pictures, or to 'see in the mind's eye'. Marked individual differences are found in the strength and clarity of reported visual imagery and these differences are of considerable psychological interest."

Visiting the AN website, which hosts their test, one can see marked differences that could potentially amount to confounding and biasing of scores. Since AN-varVVIQ is accessible via a website that aims to make profit, advertisements are shown before a participant takes the test. These are often focused on psychological tests, such as personality tests or IQ tests. In our recent visits to the site (3/11/2023), we were for example greeted by an advertisement for an online IQ test depicting a picture of Albert Einstein (Fig. 1). The association of the AN-modified VVIQ with an image of Einstein, a well-known genius, is unfortunate and could potentially affect a person's results on the test, for example by causing stereotype threat (Spencer et al., 2016), and we currently know nothing about the effects of this potentially confounding factor.

Moreover, if clicking through to the AN-varVVIQ, we can see that this also displays arbitrary boundaries between allegedly different 'conditions', similar to that seen in aphantasia research more widely, where the categories favoured by the AN are:

- Visual Aphantasia or *image-free imagination*
- Visual Hypophantasia or *mostly image-free imagination*
- Visual Phantasia or *vivid visual imagination*
- Visual Hyperphantasia *or extremely vivid visual imagery*

The above four categories appear to be unique to the AN and have never appeared in the VVIQ itself or in any related publications by its author. In particular, a category for the large, majority of people who fall in between the second and third of those listed is missing. In fact, the entire case for these 'conditions' evaporates when we simply refer to them all collectively as 'individual differences'.

Another notable difference is that the AN-varVVIQ test describes the VVIQ as a "Visual Imagination Test". Yet, as noted, the VVIQ makes no reference to the "imagination", and the consensus in both the philosophical and psychological communities is that visual imagery and imagination are importantly different (Kind, 2005). Imagination is commonly thought to be separable from visual imagery, for example when imagining hypothetical scenarios or purely semantic content, which do not rely on the generation of visual imagery. Equating the participants' visual imagery vividness with "Visual Imagination" from the get-go does not appear an optimum way to obtain a bias-free distribution of vividness scores, and a low score on the AN-varVVIQ could lead participants to falsely believe that not only is their visual imagery impaired, but their image-free imagination capacity is too. This in turn could lead to false conclusions about how creative they are or which kinds of professions would suit them.[8]

The questionnaire also differ with respect to procedural factors including the different instructions and different

[8] There is some evidence suggesting that people with aphantasia choose less traditionally creative occupations, and that the naming of 'aphantasia' did not affect these choices as they were made before the term was coined (Zeman et al., 2020). Nevertheless, people could still have been aware of their own reduced imagery ability before making these choices.

Table 2 – A comparison between VVIQ and AN-varVVIQ.

Original VVIQ	Aphantasia Network Website	Changes
Perfectly clear and as vivid as normal vision	Perfectly realistic, as vivid as real seeing	i) 'clear' changed to 'realistic' ' ii) 'normal vision' changed to 'real seeing'
Clear and reasonably vivid	Realistic and reasonably vivid	iii) 'Clear' changed to 'Realistic'
Moderately clear and vivid	Moderately realistic and vivid	iv) 'clear' changed to 'realistic'
Vague and dim	Dim and vague image	v) 'Vague and dim' changed to 'Dim and vague image' [in a previous version, a new descriptor 'flat' had been inserted]
No image at all, you only "know" that you are thinking of an object	No image at all, I only "know" I am thinking of the object	vi) "you" changed to "I" vii) "an" changed to "the'

rating scales employed by the AN-varVVIQ.[9] The descriptors of the five rating scale points differ in significant ways from the original VVIQ formulation of Marks (1973a). The AN-varVVIQ states:

> *"VVIQ Instructions*
>
> *For each scenario try to form a mental picture of the people, objects, or setting. Consider carefully the vividness of your visual imagery experience. Does some type of image come to mind? Rate how vivid the image is using the 5-point scale. If you do not have a visual image, rate vividness as '1'. Only use '5' for images that are as lively and vivid as real seeing. The rating scale is as follows:*
>
> 1. *No image at all, I only "know" I am thinking of the object*
> 2. *Dim and vague image*
> 3. *Moderately realistic and vivid*
> 4. *Realistic and reasonably vivid*
> 5. *Perfectly realistic, as vivid as real seeing"*

These are not the VVIQ instructions (see Supplementary Material), and the rating scale in the AN-varVVIQ has seven differences in the rating descriptors from the original VVIQ, which we detail below (see Table 2):

The effect of the changes shown in Table 2 on the distribution of AN-varVVIQ scores is unknown. To the best of the authors' knowledge, the various different AN-varVVIQ versions have not been empirically compared to the original VVIQ and they have never been psychometrically validated. The variants have unknown reliability and validity, and there is no way of knowing what the AN scores mean in comparison to the distribution of the original, validated VVIQ scores, which have previously been used in hundreds of different studies. Thus, close to half a million people have been issued with what they are led to assume are genuine VVIQ scores that are likely to be inaccurate. As noted, the AN is falsely branding its AN-varVVIQ questionnaire as the real VVIQ, stating that "[the VVIQ] was created in 1973 by British psychologist David Marks and is proven to be an accurate test of the vividness with which you can see people, objects, or settings in your mind's eye". This means that people are leaving the site falsely believing that they are taking the real VVIQ when in fact they have taken an altered, unapproved and unvalidated version. We set aside ethical and legal issues here, but the misrepresentation of the AN-varVVIQ as the VVIQ has possibly led hundreds of thousands of people to believe they have a form of aphantasia when, in fact, this cannot be known from an unvalidated measure with potentially biasing factors affecting their results.

Thus far, we have been discussing the misleading use of the AN-varVVIQ on a commercial site. Unfortunately, we also see uses of variant-VVIQs in peer-reviewed research, which is a matter that raises questions of probity and validity that could blot the scientific record. An altered version of the VVIQ is used by Adam Zeman, his collaborators and other researchers for investigations into the nature of aphantasia, e.g., see Zeman et al. (2020; Appendix 1). Zeman et al.'s (2020) modified variant of the VVIQ (Z-varVVIQ) differs from the original VVIQ in multiple ways that are enumerated below (see Fig. 2).

Comparing Zeman's scale with that of the original VVIQ, one can observe multiple differences that appear unnecessary and may bias the participants' scores in unknown ways. Firstly, unwarranted changes have been made to the original VVIQ instructions, e.g., the insertion of the statements: "If you do not have a visual image, rate vividness as '1'"; "Only use '5' for images that are truly as lively and vivid as real seeing"; "it is not necessarily desirable to experience imagery or, if you do, to have more vivid imagery". These additional instructions could skew participants' scores towards the low vividness end of the rating scale, as they emphasise the possibility of having no visual imagery and seem to discourage the highest rating of 5.

Secondly, for the first rating point, the description has been changed from "Perfectly clear and as vivid as normal vision" to "Perfectly clear and vivid as real seeing". We worry that 'real seeing' cannot be substituted for 'normal vision' without affecting participants' scores. To see why, it is useful to consider the contrast classes of 'normal vision' versus 'real seeing'. 'Normal vision' most naturally refers to the participant's normal vision. So for a participant with perfect vision, to rate a visual image as "Perfectly clear and as vivid as normal vision" on the VVIQ, they would need to perceive the visual image as crisply as they see the real world. But a participant with, say, astigmatism, will be rating their visual imagery in

[9] The AN altered its 'AN-varVVIQ' rating scale during the first two months of 2023 while this paper was being drafted. The text quoted in Table 1 was accessed on 7 March 2023. Each time the AN alters its AN-varVVIQ, it muddies the waters about how the current population of 'self-diagnosed' aphantasics are to be defined.

1 VIVIDNESS OF VISUAL IMAGERY QUESTIONNAIRE (VVIQ)

For each item on this questionnaire, try to form a visual image, and consider your experience carefully. For any image that you do experience, rate how vivid it is using the five-point scale described below. If you do not have a visual image, rate vividness as '1'. Only use '5' for images that are truly as lively and vivid as real seeing. Please note that there are no right or wrong answers to the questions, and that it is not necessarily desirable to experience imagery or, if you do, to have more vivid imagery.

Perfectly clear and vivid as real seeing	5
Clear and reasonably vivid	4
Moderately clear and lively	3
Vague and dim	2
No image at all, you only "know" that you are thinking of the object	1

Fig. 2 – The Z-varVVIQ.

relation to *their* normal vision, and perhaps the visual image is not as crisp as the first hypothetical participant's. But 'real seeing' has a different contrast class, namely 'unreal seeing'. Rather than just implying that things might appear a little blurry, this has connotations of hallucination, and hence a participant might think to rate their imagery on this level, it would need to involve an extreme type of hallucinatory phenomenology. Another problem with 'real seeing' is that not all real seeing is clear and vivid (Macpherson, 2018). Many real visual experiences are not, such seeing something under water or in the mist (both unclear), or seeing objects on a dark night or a stationary scene such as a grey clouded sky (not vivid).

Thirdly, for the third rating point, the description has been changed by substituting the word 'lively' for 'vivid'. Although 'lively' correctly implies something that is 'vivid', we simply do not know how the use of 'lively' instead of 'vivid' could influence the ratings on the questionnaire.

The use of an instrument with unknown psychometric properties can lead to mistaken inferences and scientific error. The above noted changes mean that scores from the Z-var-VVIQ and inferences based on them could be biased to an unknown degree and the findings might not be comparable to scores obtained in studies using the VVIQ or VVIQ-2 instruments in their original formats.

6. Could a 'diagnosis' of aphantasia create distress and stigmatization?

We worry about the above problems because of the potential impact on people's lives, as people can be led to believe that they have been 'diagnosed' (labelled) with a 'condition' (a difference) called 'aphantasia' by a scientifically validated procedure. The effect this might have on a person's life is currently largely unknown. Even though there is no well-established basis to classify aphantasia as any kind of neurological, cognitive or psychological condition (Monzel et al., 2022), aphantasia is still commonly labelled as a 'condition' (Zeman et al., 2020) and researchers also refer to participants as being 'diagnosed' with aphantasia (Dance et al., 2021a, b; 2022; Keogh & Pearson, 2018). Receiving a 'diagnosis' of a 'condition' labelled 'aphantasia' might well trigger health anxiety, stigma and other psychological concerns, such as worries about learning or memory impairments, which could cause significant distress for a person.

'Stigmatization' refers to the action of describing or regarding someone or something as worthy of disgrace or disapproval. Stigma is an adverse reaction to the perception of a negatively evaluated difference (Susman 1994). We are concerned that people receiving a 'diagnosis' of aphantasia from a trusted source, such as a psychologist or neurologist or even the Aphantasia Network, could be led to worry about the implications of this 'diagnosis', and experience increased anxiety, feelings of stigma and possibly also, upon disclosure of the 'condition' to others, be discriminated against. Aphantasia could be experienced as a disability, albeit a hidden one.

Hypothetically, four questions that may face 'diagnosed' aphantasics could be: (i) Is my aphantasia a disability or medical condition? (ii) If so, what are the implications, should I be worried? (iii) Should I self-label as a person with a disability? (iv) Should I disclose the disability to others? Indeed, the first question and similar ones are oft-searched queries on Google (see Fig. 3). It appears possible that a positive 'diagnosis' from a trusted website or test could elicit a degree of distress associated with being categorised as 'different'. In people already experiencing above average trait levels of anxiety, the extra worry created by the label might be concerning. If a person perceives their aphantasia as a disability, it can be problematic for individuals with a hidden disability deciding whether to disclose the disability because the boundary between being disabled and non-disabled is less clear than for those with visible disabilities (Hendry et al., 2022). 'Diagnosed' aphantasics have the possibility to pass as not having a disability which is less of an option for individuals with a visible disability. However, "this can have a

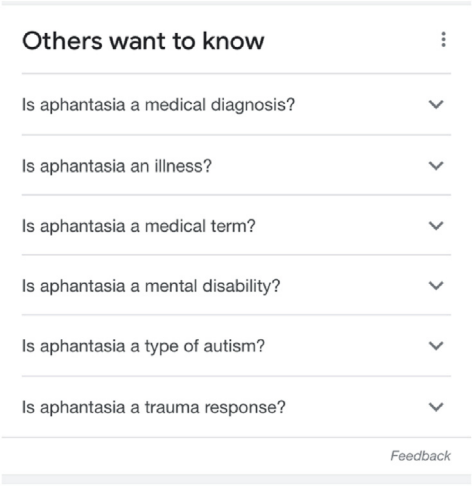

Fig. 3 – Google searches related to 'is aphantasia a medical condition?', searched on 13/3/2023.

negative impact upon self-concept and identity, as the individual must continually decide when and where to disclose the disability, which in turn can lead to discrimination" (Hendry et al., 2022).

To date, our knowledge of the wellbeing of aphantasics is limited, as is our knowledge of how a 'diagnosis' of aphantasia affects a person. The only study which has reported on the wellbeing in people with aphantasia was conducted by Monzel et al. (2022), who tested a sample of 156 participants (VVIQ score $\leq$23) on a new questionnaire – *The Aphantasia Distress Questionnaire* – which surveys people on their wellbeing in relation to their aphantasia. Although this study did not evaluate the impact of diagnosis on distress, statements in the questionnaire are based on interviews conducted with other aphantasics, and include "I had a feeling that I was inferior to other people because of my (lack of) mental imagery" and "I have a feeling that my (lack of) mental imagery was putting a strain on my personal relationships." Results showed that 34.7% experienced distress as a result of their aphantasia. The study did not test whether finding out about having aphantasia further negatively affected the wellbeing of these participants, but we know that stressful life events often lead to distress (Kendler & Prescott, 2006), and Monzel et al. suggest that finding out that one has aphantasia could indeed be a stressful life event.

This is echoed by comments from aphantasics in online communities who often report distress as a result of finding out about their 'condition'. At present, we only have anecdotal evidence on the impact of an aphantasia 'diagnosis' and no objective studies on this topic. Anecdotal evidence from social media sites including Facebook and YouTube indicates that 'diagnosed' aphantasics have been expressing a range of concerns about their label and 'diagnosis'. A Facebook site aiming to support aphantasics contains multiple posts expressing concerns about issues such as the following: feelings of 'shock' or 'devastation' on receiving the aphantasia 'diagnosis'; the possibility of having attention-deficit/hyperactivity disorder (ADHD), and/or severely deficient autobiographical memory (SDAM), and/or prosopagnosia, and/or dyslexia; having poor memory, especially for names; lack of colour memory; a bad sense of direction; poor spelling ability; poor mathematical ability, especially mental arithmetic; the possibility of having 'multi-sensory-aphantasia'; whether to disclose; the reactions of family and friends. These are not trivial concerns. Especially for adolescents, these concerns are potentially life-changing and may influence an individual's identity, perceived self-worth, self-efficacy, and self-esteem. One can hypothesise that these factors could impact educational and occupational aspirations, place boundaries on social engagement, bring stigma and discrimination that would not have been the case in the absence of the 'diagnosis' of aphantasia.

We think that the lack of research on the impact of a 'diagnosis' warrants more caution from investigators in conducting research with respect to aphantasia. Avoiding false positives, where a non-aphantasic person is wrongly characterised as aphantasic is especially important, as this could cause considerable distress. False negatives, or misses, where an aphantasic person is branded as 'non-aphantasic' could also have deleterious consequences. Hence, in line with our argumentation throughout this paper, our recommendation is to be more conservative in how we draw the line between aphantasia and non-aphantasia such that people who e.g., report merely low imagery abilities on the VVIQ are *not* classified as aphantasics. In line with how conditions are normally assessed by a specifically developed test, or battery of tests, we also recommend developing such tools, as the VVIQ was not designed to be a diagnostic test (see §6). Finally, despite our words of caution, we would like to point out that finding out that one has aphantasia could also have a positive effect one someone's life, for example by making sense of cognitive aspects which did not previously make sense (e.g., difficulty recalling episodic memories), or by enabling an individual to connect with a new community of like-minded people. These positive aspects should not be diminished and we ought to also further understand these positive effects. Nevertheless, as is the case when identifying other conditions, our point is that when doing so, we ought to use reliable and valid measures in order to avoid both false positives and false negatives.

7. Future directions

Based on the arguments in this article, we make a few recommendations aimed at moving aphantasia research forward, especially research into the nature of aphantasia, its potential underlying causes, and the development of a new psychometric measure.

We believe that fundamental questions need to receive more attention, as answering these questions will impact how we think about aphantasia and its effect on people. A pressing fundamental question is the following. Is aphantasia a disorder, or does it merely represent a natural variation of mental imagery abilities? Zeman et al. (2020) have taken the approach that aphantasia is only a natural variation in an ability, but more recently, Monzel et al. (2022) investigated whether aphantasia could be a mental disorder. They tested this with respect to a framework developed by Davison et al. (2016) whereby a mental disorder needs to meet the following

criteria: (1) statistical rarity, (2) violation of social norms and inappropriate behaviour, (3) impairments in activities of daily living, (4) personal distress. They took (1) as a necessary criterion, and either of (1)–(3) as sufficient. They found that aphantasia met (1) and though a subset of their sample also met (4), the sample as a whole did not meet any sufficient criterion. At first glance, this indicates that aphantasia should not be classified as a mental disorder. However, limitations of the study means that the answer is not final. For example, in investigating (2), they only tested for an impairment in theory of mind, and investigated this using a Reading the Mind in the Eyes Test (RMET) (Baron-Cohen et al., 2001). This only tests one aspect of theory of mind, namely the ability to attribute emotional states from perceiving an agent's eyes. It does not test for the ability to attribute other mental states, such as beliefs or desires (Baron-Cohen et al., 1985; Wellman & Phillips, 2001), and it also does not test for the ability to predict future mental states and behaviours (Goldman, 2006). These are also crucial aspects of what it means to have a theory of mind, and it is possible that aphantasics could experience deficits with respect to these aspects even if they do not experience deficits with respect to the RMET, especially as RMET does not involve the use of visual imagery, but other aspects of theory of mind do (Goldman, 2006). Hence, potential deficits in a theory of mind have not been sufficiently tested. Secondly, the sample tested by Monzel et al. was comprised of people with no imagery and low imagery (VVIQ $\leq$23), and this could have impacted results in unknown ways. If our argument in §2 is right, the sample should not include participants with low voluntary visual imagery. Testing such a sample could potentially lead to different results on (1)–(4) that could indicate that aphantasia ought to be classified as a disorder.

Further, when asking whether aphantasia is a disorder, we also need to be concerned with more the fundamental question in philosophy of psychiatry of how to individuate disorders. The practice which is common in diagnostic manuals, such as the DSM-V (American Psychiatric Association, 2013), involves conceiving of disorders as clusters of co-occurring symptoms, where a sufficient number of these symptoms needs to be present for a diagnosis but generally no one symptom is taken to be necessary. When it comes to aphantasia, some research could be taken to have established clusters of co-occurring symptoms (Dawes et al., 2020) (but note that this study involved a sample who scored $\leq$32 on the VVIQ). Importantly, finding a cluster of co-occurring symptoms for aphantasia will takes us beyond using the VVIQ, as further questionnaires and methods are needed to test for other symptoms.[10]

But a cluster of co-occurring symptoms might not be sufficient for individuating something as a distinct disorder, since there could be heterogenous causes for these symptoms. We might think that the symptoms must be driven by the malfunctioning of one underlying mechanism (McHugh & Slavney, 1998). To find this underlying mechanism, we need to not only document the cluster of co-occurring symptoms, but also develop and test hypotheses about what mechanism could cause this clustering. In aphantasia research, Blomkvist (2022) has argued that malfunctionings of the episodic system could potentially cause the co-occurring cluster of symptoms that we see in aphantasia. An alternative hypothesis is that aphantasia represents a deficit in the perceptual system, more specifically an inability to activate top-down processing in the ventral pathway (Pearson, 2019).[11] If so, findings about aphantasia could shed light on the neural areas which play a role in the formation of visual imagery, such as high-level and low-level visual areas (Bartolomeo et al., 2020; Pearson, 2019). Low-level visual areas are thought to play a role in the formation of visual imagery, but to date, studies of aphantasia have shown that there are no differences in low-level visual areas between aphantasics and controls, which is surprising as these areas have been hypothesised to play a crucial role in the formation of visual imagery (Pearson & Kosslyn, 2015). For example, a recent study using a working visual memory paradigm shows that visual images can be decoded from V1 in participants who score $\leq$32 on the VVIQ (Weber et al., 2023).[12] Finding the neural mechanisms of aphantasia is crucial for advancing the question of how aphantasia ought to be defined.

In order to test hypotheses about the underlying causes for aphantasia to ascertain whether there is a homogeneous cause in a neurological system, we need to move beyond behavioural tests. Brain imaging methods, such as fMRI, represent one way of doing this, where we should expect more homogeneous results if the clustering of symptoms result from one underlying mechanism, such as a malfunction in a particular system. So far, only one fMRI study has been conducted on aphantasia with participants that score $=$16 on the VVIQ (Fulford et al., 2018), but this method is a promising way forward.[13] In general, increased use of brain scanning techniques and more neuroscientific studies of clinical cases of lost mental imagery should be explored in tandem with behavioural tests (similar to that of Thorudottir et al., 2020).

But even if we find both a clustering of co-occurring symptoms and a common underlying mechanism in aphantasia, we should also consider social factors in deciding whether to classify aphantasia as a disorder, such as whether people experience distress or impairments of daily activities (Monzel et al., 2022), and whether a diagnosis would be helpful, e.g., for accessing treatment or therapy. Monzel et al.'s

[10] For example, when diagnosing autism five symptoms are specified in DSM-V (American Psychiatric Association, 2013): difficulty interpreting verbal and non-verbal language; difficulty 'reading' other people and expressing their own emotions; repetitive behaviour and routines; behaviour that challenges, such as episodes of frustration or in some cases violent behaviour. Any one of these criteria alone would be considered insufficient for a diagnosis (Mencap, 2023).

[11] Note that, as argued by Blomkvist (2022), this hypothesis would not explain all symptoms of aphantasia – for example, it would not explain the reduction of non-visual imagery experienced by a majority of participants with aphantasia (Zeman, 2020; Dawes, 2020) – and would be in need of further auxiliary hypotheses to do so.

[12] This study only included one participant who scored $=$16 on the VVIQ. We believe that an interesting extension of the study could involve contrasting participants who score $=$16 with those who score $\leq$32 on the same tasks.

[13] Another fMRI study was carried out by Milton et al. (2021), but this study defined the aphantasia sample as participants scoring $\leq$23 on the VVIQ.

new Aphantasia Distress Questionnaire could be beneficial to use for establishing how aphantasia might negatively impact people in their social lives. However, we note the conclusion of Monzel et al. (2022) that: "the impact on activities of daily living and personal distress is too weak to justify a classification as a mental disorder" (p. 314). We also ought to investigate whether any stigma is associated with aphantasia as this too could have an impact on wellbeing. Until we have reached a scientific consensus on whether aphantasia ought to be classified as a disorder, we should be cautious not to use medical language such as 'condition', 'diagnosis' or 'disorder' to refer to aphantasia, as the label itself could cause people unnecessary distress.

Finally, we would like to address to use of the VVIQ in future aphantasia research. Though the VVIQ is a quick, reliable and valid measure to use for first screening participants, we believe that a specialized aphantasia measure ought to be developed. The VVIQ was never designed as a clinical test and is limited in several ways when it comes to investigating aspects relevant to aphantasia. The VVIQ assesses participants' voluntary visual imagery abilities in line with the original description of aphantasia, but recent studies have found that people with aphantasia often differ in imagery abilities across the modalities (Zeman, 2020; Dawes, 2020), suggesting that we should use a more all-encompassing measure (for example, the Plymouth Sensory Imagery Questionnaire (Andrade et al., 2014), which has been used, though not for initial screening (Dance et al., 2021)). This does not imply that we ought to only study participants with reduced multisensory imagery. To further the study of aphantasia and its effect on people's lives, we ought to investigate both the potential interactions between different kinds of imagery and individual differences in imagery across modalities. We might also want to be able to distinguish between sensory imagery and other kinds of imagery, such as spatial imagery, as it is possible that a participant has a 'feeling' of the spatial layout of a scene, whilst not experiencing visual imagery of the same scene, and there is not option on current versions of the VVIQ matching this experience. A future measure ought to keep these distinct.[14] Further, it has been suggested that aphantasia could be associated with differences in metacognitive access to the visual imagery representations (Liu & Bartolomeo, 2023), and that differences in metacognitive access could bias participants' responses to questions on the VVIQ. This study compared participants with low visual imagery to those with average or high visual imagery, showing that the former display slower response times with retained accuracy compared to the other groups on visual imagery and visual perception tasks. The authors suggest that this is due to a slowing of visual processing. This is an interesting suggestion, as metacognitive monitoring has been demonstrated to be sensitive to processing fluency in non-aphantasic participants (Murphy et al., 2022). We take from this that a psychometric measure of aphantasia ought to also assess metacognition, which is not assessed by the VVIQ. Finally, the new measure ought to also test memory, as systematic differences between aphantasics and controls have also been found in episodic memory (Bainbridge, 2021; Dawes, 2020), and some differences have been documented in working memory tasks (Keogh & Pearson, 2021; Pounder, 2021). Again, these aspects are not assessed by the VVIQ. For these reasons, we believe that a specialized psychometric measure ought to be developed for aphantasia, and that we should not rely solely on the VVIQ to 'diagnose' aphantasia. As we are still discovering further aspects of aphantasia, it should be noted that this discussion is not meant to capture the exhaustive list of what a psychometric measure should test, but we hope that it will point the research field in the right direction.[15]

In conclusion, we believe this article serves to bring aphantasia research forward by pointing out weaknesses in the methodology in current aphantasia research and by indicating ways to improve methods for investigating aphantasia. We have suggested that the claim that congenital aphantasia is a 'condition' is premature, requiring much more detailed and rigorous investigation with improved methodology. Also, to date, no evidence has been brought forward to suggest a single psychiatric sign or symptom of the most common, congenital form of aphantasia as a special condition. Hence, for the time being, we prefer the parsimonious hypothesis that congenital aphantasia is an individual difference at one end of a normal distribution of imagery vividness differences. At the same time, we believe that researching aphantasia would be improved by the development of a new psychometric measure for the identification of aphantasia, as the current use of the VVIQ does not capture all aspects of aphantasia, and makes difficult cross-study comparisons. To further investigate whether aphantasia should be labelled a 'condition', we suggest that we ought to identify a reliable clustering of co-occurring symptoms which would warrant the testing of hypotheses about the underlying causes of congenital aphantasia. Ultimately, when investigating aphantasia, we also ought to not lose sight of the wellbeing of the people whose lives are affected by being given a label indicating an extreme individual difference.

Funding

AB's work is part of a project supported by a UKRI Future Leaders Fellowship, Grant Ref: MR/W00741X/1.

CRediT authorship contribution statement

Andrea Blomkvist: Conceptualization, Investigation, Writing – original draft, Writing – review & editing. **David F. Marks:** Conceptualization, Investigation, Writing – original draft, Writing – review & editing.

Declaration of competing interest

None.

[14] Thank you to an anonymous reviewer for raising the point about a measurement's ability to distinguish spatial imagery from visual imagery.

[15] To clarify our position on a new measure of aphantasia.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cortex.2023.09.004.

REFERENCES

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders [DSM-5]* (5th ed.). American Psychiatric Association.

Andrade, J., May, J., Deeprose, C., Baugh, S. J., & Ganis, G. (2014). Assessing vividness of mental imagery: The Plymouth sensory imagery questionnaire. *British Journal of Psychology, 105*(4), 547–563. https://doi.org/10.1111/bjop.12050. Epub 2013 Oct. 7. PMID: 24117327.

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; a Journal Devoted To the Study of the Nervous System and Behavior, 135*, 159–172. https://doi.org/10.1016/j.cortex.2020.11.014

Barber, T. X. (1959). The" eidetic image" and" hallucinatory" behavior: A suggestion for further research. *Psychological Bulletin, 56*(3), 236.

Baron-Cohen, S., Leslie, A. M., & Frith, U. T. A. (1985). Does the autistic child have a "theory of mind"? *Cognition, 21*, 37–46.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "' Reading the Mind in the Eyes "'. *Test Revised Version : A Study with Normal Adults, and Adults with Asperger Syndrome or High-functioning Autism., 42*(2), 241–251.

Bartolomeo, P. (2002). The relationship between visual perception and visual mental imagery: A reappraisal of the neuropsychological evidence. *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior, 38*(3), 357–378.

Bartolomeo, P., Hajhajate, D., Liu, J., et al. (2020). Assessing the causal role of early visual areas in visual mental imagery. *Nature Reviews. Neuroscience, 21*, 517.

Basso, A., Bisiach, E., & Luzzatti, C. (1980). Loss of mental imagery: A case study. *Neuropsychologia, 18*, 435–442.

Berger, G. H., & Gaunitz, S. C. (1979). Self-rated imagery and encoding strategies in visual memory. *British Journal of Psychology, 70*(1), 21–24.

Blomkvist, A. (2022). Aphantasia: In search of a theory. *Mind and Language, May*, 1–23. https://doi.org/10.1111/mila.12432

Brain, R. (1954). Loss of visualization. *Proceedings of the Royal Society of Medicine, 47*, 288–290.

Campos, A., & Pérez-Fabello, M. J. (2009). Psychometric quality of a revised version vividness of visual imagery questionnaire. *Perceptual and motor skills, 108*(3), 798–802.

Charcot, J. M., & Bernard, D. (1883). Un cas de suppression brusque et isolée de la vision mentale des signes et des objets (formes et couleurs). *Le Progrés Médical, 11*, 568–571.

Craver, C., & Tabery, J. (2019). Mechanisms in science. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy*.

de Vito, S., & Bartolomeo, P. (2016). Refusing to imagine? On the possibility of psychogenic aphantasia. A commentary on zeman et al. (2015). *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior, 74*, 334–335. https://doi.org/10.1016/j.cortex.2015.06.013

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, 22*(167), 318–334. https://doi.org/10.1016/j.cortex.2023.06.015. Epub ahead of print. PMID: 37597266.

Dance, C. J., Ipser, A., & Simner, J. (2022). The prevalence of aphantasia (imagery weakness) in the general population. *Consciousness and Cognition*. https://doi.org/10.1016/j.concog.2021.103243

Dance, C. J., Jaquiery, M., Eagleman, D. M., Porteous, D., Zeman, A., & Simner, J. (2021a). What is the relationship between aphantasia, synaesthesia and autism? *Consciousness and Cognition, 89*(January), Article 103087. https://doi.org/10.1016/j.concog.2021.103087

Dance, C. J., Ward, J., & Simner, J. (2021b). What is the link between mental imagery and sensory sensitivity? Insights from aphantasia. *Perception, 50*(9), 757–782. https://doi.org/10.1177/03010066211042186. Epub 2021 Aug. 31. PMID: 34463590; PMCID: PMC8438787.

Davison, G. C., Neale, J. M., & Hautzinger, M. (2016). Klinische Psychologie [clinical psychology]. *Bertz Verlag*.

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), 1–10. https://doi.org/10.1038/s41598-020-65705-7

Dawes, A. J., Keogh, R., Robuck, S., & Pearson, J. (2022). Memories with a blind mind: Remembering the past and imagining the future with aphantasia. *Cognition, 227*(July 2021), Article 105192. https://doi.org/10.1016/j.cognition.2022.105192

Eagleman, D. M., Kagan, A. D., Nelson, S. S., Sagaram, D., & Sarma, A. K. (2007). A standardized test battery for the study of synesthesia. *Journal of Neuroscience Methods, 159*(1), 139–145. https://doi.org/10.1016/j.jneumeth.2006.07.012

Farah, M. J., Levine, D. N., & Calvanio, R. (1988). A case study of mental imagery deficit. *Brain and Cognition, 8*(2), 147–164. https://doi.org/10.1016/0278-2626(88)90046-2. PMID: 3196480.

Faw, B. (2009). Conflicting intuitions may be based on differing abilities - evidence from mental imaging research. Journal. *Journal of Consciousness Studies, 16*(4), 45–68.

Fulford, J., Milton, F., Salas, D., Smith, A., Simler, A., Winlove, C., & Zeman, A. (2018). The neural correlates of visual imagery vividness – an fMRI study and literature review. *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior, 105*, 26–40. https://doi.org/10.1016/j.cortex.2017.09.014

Goldman, A. (2006). *Simulating minds: The philosophy, psychology, and Neuroscience of mindreading (issue 3)*. Oxford University Press USA.

Greenberg, D. L., & Knowlton, B. J. (2014). The role of visual imagery in autobiographical memory. *Memory & Cognition, 42*(6), 922–934. https://doi.org/10.3758/s13421-014-0402-5

Gur, R. C., & Hilgard, E. R. (1975). Visual imagery and the discrimination of differences between altered pictures simultaneously and successively presented. *British Journal of Psychology, 66*(3), 341–345.

Haber, R. N. (1969). Eidetic images. *Scientific American, 220*(4), 36–47.

Hatakeyama, T. (1974). The process of having identified an adult eidetic person and her eidetic experiences in daily life. *Tohoku Psychologica Folia*.

Hebb, D. O. (1968). Concerning imagery. *Psychological review, 75*(6), 466.

Hendry, G., Wilson, C., Orr, M., & Scullion, R. (2022). "I just stay in the house so I don't need to explain": A qualitative investigation of persons with invisible disabilities. *Disabilities, 2*(1), 145–163.

Holt, R. R. (1964). Imagery: The return of the ostracized. *American Psychologist, 19*(4), 254–264. https://doi.org/10.1037/h0046316.

Hornett, W. (2021). Perceptual capacities, discrimination, and the senses. *Synthese, 199*(5), 14063–14085. https://doi.org/10.1007/s11229-021-03410-2

Horowitz, M. J. (1970). *Image formation and cognition*. New York: Appleton-Century-Crofts.

Isaac, A. R., & Marks, D. F. (1994). Individual differences in mental imagery experience: Developmental changes and specialization. *British Journal of Psychology, 85*(4), 479–500.

Jacobs, C., Schwarzkopf, D. S., & Silvanto, J. (2018). Visual working memory performance in aphantasia. *Cortex; a Journal Devoted*

To the Study of the Nervous System and Behavior, 105, 61–73. https://doi.org/10.1016/j.cortex.2017.10.014

Kendler, K. S., & Prescott, C. A. (2006). *Genes, environment, and psychopathology: Understanding the causes of psychiatric and substance use disorder*. The Guilford Press.

Keogh, R., & Pearson, J. (2018). The blind mind: No sensory visual imagery in aphantasia. *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior, 105*(2015), 53–60. https://doi.org/10.1016/j.cortex.2017.10.012

Keogh, R., & Pearson, J. (2021). Attention driven phantom vision: Measuring the sensory strength of attentional templates and their relation to visual mental imagery and aphantasia: Measuring attentional templates. *Philosophical Transactions of the Royal Society B: Biological Sciences, 376*(1817). https://doi.org/10.1098/rstb.2019.0688rstb20190688

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

Kind, A. (2005). Imagery and imagination. *In Internet Encyclopedia of Philosophy*.

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

Liu, J., Spagna, A., & Bartolomeo, P. (2022). Hemispheric asymmetries in visual mental imagery. *Brain Structure & Function, 227*, 697–708.

Luria, A. R. (1987). *The mind of a Mnemonist: a little book about a vast memory, with a new foreword by Jerome S. Bruner*. Harvard University Press.

Macpherson, F. (2018). What is it like to have visual imagery? In S. Aldworth, & M. MacKisack (Eds.), *Extreme imagination: Inside the eye's mind*. University of Exeter.

Marks, D. F. (1972). In P. W. Sheehan (Ed.), *Individual differences in the vividness of visual imagery and their effect on function. The function and nature of imagery* (pp. 83–108). New York: Academic Press.

Marks, D. F. (1973a). 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

Marks, D. F. (1973b). Visual imagery differences and eye movements in the recall of pictures. *Perception & Psychophysics, 14*(3), 407–412.

Marks, D. F. (1986). *Theories of image formation. New York: brandon house*.

Marks, D. F. (1995). New directions for mental imagery research. *Journal of Mental Imagery, 19*(3–4), 153–167.

Marks, D. F. (2023). The action cycle theory of perception and mental imagery. *Vision, 7*(1), 12.

Marks, D. F., & Yardley, L. (2004). *Research methods for clinical and health psychology*. London: Sage.

McHugh, P., & Slavney, T. (1998). *The perspectives of psychiatry*. John Hopkins University Press.

McKellar, P. (1957). *Imagination and thinking*. New York: Basic Books.

McKelvie, S. J. (1995). The VVIQ and beyond: Vividness and its measurement. *Journal of Mental Imagery, 19*(3–4), 197–252.

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), 1–15. https://doi.org/10.1093/texcom/tgab035

Monzel, M., Vetterlein, A., & Reuter, M. (2022). No general pathological significance of aphantasia: An evaluation based on criteria for mental disorders. *Scandinavian Journal of Psychology*. https://doi.org/10.1111/sjop.12887

Moro, V., Berlucchi, G., Lerch, J., Tomaiuolo, F., & Aglioti, S. M. (2008). Selective deficit of mental visual imagery with intact primary visual cortex and visual perception. *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior, 44*(2), 109–118.

Murphy, D. H., Huckins, S. C., Rhodes, M. G., & Castel, A. D. (2022). The effect of perceptual processing fluency and value on metacognition and remembering. *Psychonomic Bulletin & Review, 29*, 910–921. https://doi.org/10.3758/s13423-021-02030-8

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological review, 76*(3), 241.

Paivio, A., Yuille, J. C., & Smythe, P. C. (1966). Stimulus and response abstractness, imagery, and meaningfulness, and reported mediators in paired-associate learning. *Canadian Journal of Psychology/Revue canadienne de psychologie, 20*(4), 362.

Pearson, J. (2019). The human imagination: The cognitive neuroscience of visual mental imagery. *Nature Reviews Neuroscience, 20*(10), 624–634. https://doi.org/10.1038/s41583-019-0202-9

Pearson, J., & Kosslyn, S. M. (2015). The heterogeneity of mental representation: Ending the imagery debate. *Proc Natl Acad Sci U S A, 112*(33), 10089–10092. https://doi.org/10.1073/pnas.1504933112

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; a Journal Devoted To the Study of the Nervous System and Behavior, 148*, 180–192. https://doi.org/10.1016/j.cortex.2021.12.010. Epub 2022 Jan. 24. PMID: 35180481.

Richardson, A. (1969). *Mental imagery*. Springer. https://doi.org/10.1007/978-3-662-37817-5.

Riddoch, M. J. (1990). Loss of visual imagery: A generation deficit. *Cognitive Neuropsychology, 7*(4), 249–273. https://doi.org/10.1080/02643299008253444

Sarbin, T. R., & Juhasz, J. B. (1967). The historical background of the concept of hallucination. *Journal of the History of the Behavioral Sciences*.

Segal, S., & Gordon, P-E. (1969). The Perky effect revisited: Blocking of visual signals by imagery. *Perceptual and Motor Skills, 28*, 791–797.

Sheehan, P. W. (1972). Role of imagery in incidental learning: Replication and extension of an effect. *Journal of Experimental Psychology, 95*(1), 226–228. https://doi.org/10.1037/h0033306.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*(3972), 701–703. https://doi.org/10.1126/science.171.3972.701.

Singer, J. L., & Antrobus, J. S. (1963). A factor-analytic study of daydreaming and conceptually-related cognitive and personality variables. *Perceptual and Motor Skills, 17*(1), 187–209. https://doi.org/10.2466/pms.1963.17.1.187.

Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype threat. *Annual Review of Psychology*. https://doi.org/10.1146/annurev-psych-073115-103235

Thomas, J. J., Ayuno, K. C., Kluger, F. E., & Caplan, J. B. (2022). The relationship between interactive-imagery instructions and association memory. *Memory & Cognition*, 1–20.

Thorudottir, S., Sigurdardottir, H. M., Rice, G. E., Kerry, S. J., Robotham, R. J., Leff, A. P., & Starrfelt, R. (2020). The architect who lost the ability to imagine: The cerebral basis of visual imagery. *Brain Sciences, 10*(2), 59.

Weber, S., Christophel, T., Görgen, K., Soch, J., & Haynes, J. D. (2023). Working memory and imagery in early visual cortex. *bioRxiv, 2023-02*.

Wellman, H. M., & Phillips, A. T. (2001). Developing intentional understandings. In *Intentions and intentionality: Foundations of social cognition*. (pp. 125–148). The MIT Press.

West, L. J. (1962). A general theory of hallucinations and dreams. In L. J. West (Ed.), *Hallucinations* (pp. 275–290). Grune & Stratton.

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 Royal Society: B, 1946*.

Yates, M., & Kamboj, S. K. (2017 Aug 31). Alcohol related mental imagery: The effects of a priming dose in at risk drinkers. *Addictive Behaviors Reports, 6*, 71–75. https://doi.org/10.1016/j.abrep.2017.08.003. PMID: 29450239; PMCID: PMC5800585.

Zeman, A. (2020). *Phantasia - the psychological significance of lifelong visual imagery vividness extremes - Supplementary Material*.

Zeman, A., 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; a Journal Devoted To the Study of the Nervous System and Behavior, 73*, 378–380. https://doi.org/10.1016/j.cortex.2015.05.019

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; a Journal Devoted To the Study of the Nervous System and Behavior, 130*, 426–440. https://doi.org/10.1016/j.cortex.2020.04.003