A Clinical, Neuropsychological, and Forensic Investigation of Congenital Aphantasia

Carla Jane Dance

Thesis submitted for the degree of Doctor of Philosophy

School of Psychology

University of Sussex

July 2022

Declaration

The thesis conforms to an 'article format' in which Chapters 2-5 consist of discrete articles written in a style that is appropriate for publication in peer-reviewed journals in the field. Chapter 1 and Chapter 6 present an overview and discussion of this thesis.

Chapter 2 is published in *Consciousness and Cognition* as:

Dance, C. J., Ipser, A., & Simner, J. (2022). The prevalence of aphantasia (imagery weakness) in the general population. *Consciousness and cognition*, *97*, 103243.

The author contributions are as follows: Carla Dance was responsible for data collection (with assistance from Alberta Ipser), data analysis, and writing the manuscript. Julia Simner (supervisor) gave feedback on the study design and manuscript.

Chapter 3 is published in *Consciousness and Cognition* as:

Dance, C. J., Jaquiery, M., Eagleman, D. M., Porteous, D., Zeman, A., & Simner, J. (2021). What is the relationship between Aphantasia, Synaesthesia and Autism? *Consciousness and Cognition*, *89*, 103087.

The author contributions are as follows: Carla Dance was responsible for all data analysis, and for primary data collection on aphantasia and autistic traits (with additional data from Matt Jaquiery who provided 32 out of 236 aphants/controls). All analyses on synaesthesia were also conducted by Carla Dance, using synaesthesia scores collected by Julia Simner, David Eagleman [Synesthesia Battery], and David Porteous [Generation Scotland]. Carla Dance wrote the manuscript. Julia Simner (supervisor), Matt Jaquiery, Adam Zeman, David Eagleman, and David Porteous gave feedback on the manuscript.

Chapter 4 is published in *Perception* as:

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.

The author contributions are as follows: Carla Dance was responsible for data collection, data analysis, and writing the manuscript. Julia Simner (supervisor) gave feedback on the study design and manuscript. Jamie Ward also provided feedback on the manuscript.

Chapter 5 has been submitted for publication as:

Dance, C. J., Hole, G., & Simner, J. (2022). The role of visual imagery in face recognition, face matching, and the construction of facial composites. Evidence from Aphantasia. [Manuscript under review].

The author contributions are as follows: Carla Dance was responsible for data collection, data analysis, and writing the manuscript. Julia Simner (supervisor) and Graham Hole gave feedback on the study design and manuscript.

I hereby declare that this thesis has not been, and will not be, submitted in whole or in part to another University for the award of any other degree.

Carla Jane Dance
July 2022

Acknowledgements

I would like to take the chance to thank the people who have made this thesis possible. First and foremost, I would like to thank my supervisor Professor Julia Simner for the endless guidance and support she has provided throughout my PhD journey. She has helped me to grow into the researcher I am today, and I will be forever grateful to her for her excellent mentorship. It has been a true privilege to have been able to work beside you Jools – thank you.

Thank you also to my second supervisor Professor Jamie Ward, who has supported me through my thesis, and in my Leverhulme Doctoral Scholarship duties more broadly along with Professor Anil Seth. I would also like to thank my collaborators, including Dr Frances Meeten and Dr Graham Hole, who were instrumental in bringing my research ideas to life. I am grateful to you for your guidance and expertise. I would also like to thank the MULTISENSE and Imagery Lab. I have learned so much from you all, and I am so grateful to have been a part of such a truly wonderful lab group.

I would also like to extend thanks to the research participants who have made this thesis possible. I am so very grateful for your willingness to take part in our studies. One of the main focuses of my research is to help people understand themselves and others better. I hope that you find the results of my research helpful in some way, and again thank you for your support and participation.

Finally, thank you to my friends and family who have supported me over the last few years. I cannot thank you enough for always being there for me, and for having my back. Thank you to my parents, and my brothers, for always believing in me, and to my friends for keeping me motivated and upbeat. Thank you also to Alastair, my number one supporter throughout this whole process. I cannot have imagined completing this PhD without you all in my corner.

Abstract

Congenital aphantasia is a neurodevelopmental condition characterised by markedly impaired visual imagery determined from birth. For example, when asked to imagine a sunset, most people can 'see' an image of the scene within their mind's eye. People with aphantasia do not experience an internal image, despite knowing what a sunset looks like. Aphantasia only entered mainstream science in recent years, meaning our understanding of the condition is relatively limited. Through four experimental studies, this thesis will enhance our understanding of aphantasia from a clinical, neuropsychological, and forensic perspective. Chapter 2 asks how often aphantasia occurs in the general population, providing a reliable prevalence estimate. Chapter 3 examines how aphantasia intersects with the clinical condition of autism, and the neuropsychological condition of synaesthesia – both traits themselves characterised by imagery differences. Here, I show that aphantasics tend to experience high levels of autistic traits, and have an unusual profile of (certain types of) synaesthesia. Chapter 4 examines the broader sensory/imagery profile of aphantasia. Here, I show that aphantasia is often tied to low imagery beyond the visual domain (weak auditory imagery, olfactory imagery, etc.), and I coin the term *dysikonesia* to characterise multi-sensory imagery weakness (where aphantasia is the *visual* subtype). I also show that aphantasics tend to experience lower levels of sensory sensitivity (i.e., differences in responsivity to incoming sensation) across multiple senses. These findings demonstrate for the first time that mental imagery and sensory sensitivity are linked. Finally, Chapter 5 explores the forensic question of whether imagery influences face processing. I show that aphantasia is linked to weak face recognition and perception, but intact abilities when constructing face composites. Overall, this thesis shows that aphantasia is linked to a distinct psychological profile, with clinical, neuropsychological, and forensic implications. Together, this body of work serves to enhance our understanding of aphantasia, and mental imagery more broadly.

Table of Contents

Chapter 5

Chapter 6

Chapter 1
General Introduction

Visual imagery is the ability to build a picture-like representation in the mind. For example, when asked to imagine a sunset, most people can create an image of the scene within their mind's eye. The vividness of visual imagery varies from person to person. For some, it is exceptionally strong and nearly as vivid as real-life perception, whereas for others it is virtually or completely absent, a condition known as *aphantasia*. In this thesis I will provide an in-depth examination of the clinical, neuropsychological, and forensic profile of aphantasia. Through four experimental chapters, I will answer key questions about aphantasia relating to its prevalence, its links with other neuropsychological or clinical conditions (synaesthesia, and autism), its relationship with sensory sensitivity (see below), and its role in face processing. In this first chapter I will introduce aphantasia, providing scientific background to my research questions. Importantly, a more detailed literature review will be provided within each empirical chapter of this thesis, since all four chapters have been submitted for publication (at the time of submitting this thesis, three had already been published, and one is currently under review). This general introduction therefore provides an overview of key concepts, paving the way for the empirical chapters that follow.

What is Congenital Aphantasia?

Congenital aphantasia (henceforth aphantasia) is a neurodevelopmental condition characterised by markedly impaired visual imagery from birth. Simply put, people with aphantasia experience no visual imagery whatsoever, or imagery that is only dim, vague, or fleeting (e.g., Keogh & Pearson, 2018; Zeman, Dewar, & Della Sala, 2015). Although scientists have known for at least 100 years that some people have a blind 'mind's eye' (Galton, 1880), aphantasia only entered mainstream research in recent years since the term was coined in 2015 (Zeman et al., 2015). Before introducing aphantasia in further detail, it is important to highlight that this seminal paper by Zeman and colleagues defined aphantasia as a congenital '*condition*'. Here, the word condition is used to indicate that aphantasia is linked to a specific profile of characteristics, akin to the way that other sensory differences (e.g., synaesthesia; see below) are also defined as conditions. In *no sense* is the use of the word condition meant to be indicative of any kind of inherent 'deficit', but rather, it illustrates that aphantasia is one (of many) variances in human

sensory experience. I recognise, however, that there will be individual differences in how people with aphantasia define their imagery experience, and some may indeed prefer to refer to it differently (e.g., as a trait, sensory difference, or otherwise). Throughout this thesis, I therefore define aphantasia as a condition in line with current understandings, but acknowledge the importance of researchers being sensitive to the implications, limitations, and indeed consequences of the language used to define aphantasia and other sensory differences (I return to these issues within the General Discussion; see Chapter 6).

Within the existing literature, the most widely used approach to categorise aphantasia is to administer the *Vividness of Visual Imagery Questionnaire* (*VVIQ*; Marks, 1973, 1995). This 16-item measure asks participants to imagine different scenes (e.g., "A beach by the ocean on a warm summer's day"), and rate the vividness of their visual imagery. Rating takes place on a 5-point scale of 1 ("No image at all, you only "know" that you are thinking of the object"), 2 ("Vague and dim"), 3 ("Moderately clear and vivid"), 4 ("Clear and reasonably vivid"), and 5 ("Perfectly clear and as vivid as normal vision"). An overall test score of 16-32 captures the phenomenology of aphantasia, where imagery is either absent or vague/dim/fleeting (Zeman et al., 2015; Zeman, Dewar, & Della Sala, 2016). Below we will see that this self-report questionnaire is particularly useful in identifying aphantasia, and it converges with evidence from a range of behavioural tasks. Remarkably, although visual imagery is thought to be involved in a number of important functions relating to memory (D'Argembeau & Van der Linden, 2006; Keogh & Pearson, 2014), motivation (Knäuper et al., 2011), and mental wellbeing (Blackwell et al., 2013), people with aphantasia nevertheless tend to live 'normal' lives, often without realising they are in any way different (Zeman et al., 2016). Some scientists have even gone as far as suggesting that aphantasia may be no more than a problem with *metacognition*. In other words, rather than truly lacking visual imagery, aphantasics might just lack the ability to report it accurately, or be unable to access it at a conscious level (de Vito & Bartolomeo, 2016). There are, however, a growing number of studies showing that aphantasia is likely characterised by a *true* impairment in visual imagery, unrelated to self-reporting. In providing clear motivation for my investigation of aphantasia as a distinct condition, I now briefly summarise this literature in turn below.

Besides showing differences on self-report measures of imagery (e.g., *VVIQ*; Marks, 1973, 1995), previous research has also found that people with aphantasia differ from people with intact visual imagery (henceforth 'imaging controls') on *objective* imagery tasks. For example, Keogh and Pearson (2018) found differences in an imagery-based binocular rivalry paradigm. In this task, participants were first asked to visualise a colour (i.e., make a mental image of the colour in their mind's eye). They were then presented with a binocular rivalry display (i.e., one colour to each eye), where one of the presented colours matched the colour previously visualised. The typical phenomenological experience of a binocular display is a sense of 'rivalry'; i.e., the colours do not fuse but slowly flicker from one to the other in an alternating display. However, from earlier studies we know that visualising a colour in advance means that this colour tends to become dominant in the binocular display (e.g., if green is visualised in the mind's eye, then green becomes dominant in a subsequent binocular presentation of green and red). The extent of this dominance is indicative of visual imagery strength (Pearson, Clifford, & Tong, 2008). Strikingly, people with aphantasia do not show this dominance effect (Keogh & Pearson, 2018), suggesting they do indeed lack the ability to generate visual imagery. This is further supported by the fact that dominance scores on the imagery binocular rivalry paradigm also correlate with the spectrum of scores on the *VVIQ* (Keogh & Pearson, 2018; Pearson, Rademaker, & Tong, 2011). People, therefore, seem to have relatively good metacognition about their own visual imagery abilities.

Consistent with this, other research provides yet further evidence for aphantasia being characterised by a true absence of visual imagery. A recent study by Kay, Keogh, Andrillon and Pearson (2021) found differences between aphantasics and imaging controls in their pupillary light response. In 2x2 design, participants were asked to perceive or visualise (i.e., make an image in their mind's eye) either bright or dark triangles. The pupils of both aphantasics and imaging controls constricted when they perceived bright relative to dark triangles (this pupillary light response is a physiological reflex that serves to protect the eye in response to bright light; Kay et al., 2021). Interestingly, this effect was also seen when the imagers were asked to *visualise* the triangles, which we can take as a physiological manifestation of visual imagery. In contrast, aphantasic participants showed no pupillary light response when they were asked to visualise the triangles. These findings therefore provide physiological evidence for a lack of visual imagery in aphantasia.

Neuroimaging research, too, provides further support for aphantasia being driven by a genuine impairment in visual imagery. A study by Milton et al. (2021) found that people with aphantasia have reduced connectivity (resting state fMRI) between visual-occipital areas and areas involved in top-down control (prefrontal cortex) relative to very vivid imagers (otherwise known as *hyperphantasics*; Milton et al., 2021). Aphantasics also had significantly less BOLD activity in the left anterior parietal cortex when asked to imagine (as opposed to perceive) famous places/faces, relative to both very vivid, and average imagers (Milton et al., 2021). Given that visual imagery tends to recruit areas involved in vision and visual attention (e.g., visual cortex, parietal cortex) (Cui, Jeter, Yang, Montague, & David, 2007; Dijkstra, Bosch, & van Gerven, 2017; McGeown et al., 2012; Winlove et al., 2018), but also frontal areas linked to attention and information integration (see Dijkstra et al., 2017, 2019; Pearson et al., 2015), these findings suggest that aphantasia may be driven – at least to some degree – by reduced connectivity between visual and frontal areas of the brain, but also that aphantasics may rely less on visual areas when completing 'imagery-based' tasks. These neural differences therefore provide further evidence for visual imagery being impaired in aphantasia.

In sum, this brief literature review has introduced aphantasia as a sensory difference characterised by markedly impaired visual imagery (as opposed to an issue with metacognition). Like other sensory traits that have come before it, aphantasia may be seen as an *individual difference*, providing a different way of thinking, and of processing the world. Next, I summarise what previous research can tell us about the characteristics associated with aphantasia, before presenting the questions that this thesis aims to answer. In other words, in what ways does aphantasia influence the lives of people who experience it, and what questions remain unanswered?

Aphantasia as an individual difference

Previous research has started to reveal how aphantasia might affect day-to-day life. For example, aphantasic individuals experience less detailed autobiographical memories and imagined events[1] (Dawes, Keogh, Andrillon, & Pearson, 2020; Dawes, Keogh, Robuck,

[1] We note here that imagination and imagery are related, but they are *not* the same thing (Dance, Jaquiery et al., 2021). Despite their lack of visual imagery, people with aphantasia can have very strong imagination indeed (for example, see Zeman et al., 2019 for art creations by aphantasic people). I will explore this in more detail in Chapter 3.

& Pearson, 2022; Milton et al., 2021), and report having less sensorily detailed night-time dreams than imaging controls (Dawes et al., 2020). Aphantasia also has implications for career pathways, since aphantasics are more likely to enter professions in science and maths as opposed to the arts (Zeman et al., 2020). In other ways, however, aphantasic individuals are no different than people with intact visual imagery. For example, aphantasics are just as accurate as imaging controls at remembering and then recognising series of words and patterns (Pounder et al., 2022), and at remembering the details of a short story (Milton et al., 2021). They are also just as accurate in tests of number, spatial, and visual working memory (Keogh, Wicken, & Pearson, 2021; Knight, Milton, & Zeman, 2022), although they tend to use *different (non-visual) strategies* (Keogh et al., 2021). Interestingly, they also perform just as well as imaging controls in tests once thought to measure visual imagery. For example, they perform as accurately as controls in the Curved Segments test, where participants indicate whether an imagined upper-case letter in the alphabet has a curved line segment (e.g., P = yes, N = no; Milton et al., 2021). They also perform as accurately as controls in Shepard and Metzler's (1971) rotation task, where participants mentally rotate a 3D shape and identify whether it is identical or different to another 3D target shape (Pounder et al., 2022). However, aphantasics again seem to use different cognitive strategies relative to imaging controls to complete these kinds of mental rotation tasks, since their response times (particularly for those with *no visual imagery whatsoever*) tend to be significantly slower (Pounder et al., 2022).

Similarly, other research has found that aphantasics have similar attentional abilities as imaging controls, despite using different strategies. In a study by Monzel, Keidel, and Reuter (2021), aphantasics and imaging controls were asked to visualise a target fruit (e.g., banana) and then identify the target out of pairs of subsequently presented fruits (e.g., banana, tomato). There was no difference (in accuracy or reaction time) between groups when the pairs of fruit were presented as words. However, aphantasics were slower than controls (though equally accurate) when the pairs of fruit were presented as pictures. These findings suggest that attentional strategies for aphantasics may be less susceptible to being primed by mental images, yet this does not generally impair their attentional ability to search for and find a target overall (i.e., their accuracy remains intact; see also Keogh & Pearson, 2021). Overall, then, in some ways aphantasics experience slight weaknesses relative to people with visual imagery (e.g., in autobiographical

memory, and imagination), yet in other ways they appear no different (although they might use different cognitive strategies).

Another area of research has examined whether aphantasia influences the ability to remember and then re-create pictures from memory. In a study by Bainbridge, Pounder, Eardley, and Baker (2021), aphantasics and imaging controls were presented with a series of pictures (e.g., a kitchen scene), and were subsequently asked to draw the pictures from memory. Drawings by aphantasics contained fewer objects (e.g., omitted a chair) and less colour, but were equally accurate in object details (e.g., texture), size, and spatial location. There were, however, no differences between the two groups in drawings produced when the picture was visible on screen (i.e., not from memory), nor in the ability to recognise the pictures. These findings show that aphantasics may have slight weaknesses in recalling objects from memory, but their object-detail and spatial abilities remain intact (see also Dawes et al., 2020; Keogh & Pearson, 2018; Palermo, Boccia, Piccardi, & Nori, 2022), as do their abilities to recognise and copy pictures. Similarly, another study by Milton et al. (2021) found that aphantasics were just as accurate as imaging controls at drawing a complex line-drawing figure from memory (Rey-Osterrieth Figure; Osterreith & Rey, 1944). However, their memory performance in these kinds of tasks becomes less accurate than controls when memory demands become very difficult (see Jacobs, Schwarzkopf, & Silvanto, 2018; Monzel et al., 2021). Overall, these findings suggest that aphantasics might experience impairments in long-term object memory, though these weaknesses seem to be relatively minimal (i.e., they can nevertheless re-construct the details of scenes rather well, with slight impairments tending to appear more prominently only when memory demands become challenging).

Finally, another line of research has shown that aphantasia may be *protective* against certain clinical psychopathological traits. For example, aphantasics tend to experience less intrusive imagery of past traumatic events compared to imaging controls (Dawes et al., 2020), and less physiological fear responses when asked to read and imagine fear evoking passages (Wicken, Keogh, & Pearson, 2021). However, their fear responses are no different from imaging controls when they instead *perceive* fearful stimuli in pictures (Wicken et al., 2021). These findings suggest that aphantasia might be protective against experiencing fearful responses induced by imagining scenarios, and intrusive imagery characteristic of anxiety conditions such as post-traumatic stress disorder (Hirsch &

Holmes, 2007). More broadly, these studies highlight the powerful role of imagery in inducing emotive responses (Holmes & Mathews, 2005, 2010). Other research has shown that aphantasics are also less susceptible to artificially induced pseudo-hallucinations (Königsmark, Bergmann, & Reeder, 2021). This is consistent with findings showing that strong imagery is linked to a higher incidence of visual hallucinations (Shine et al., 2015), and suggests that aphantasics might be less susceptible to hallucinations in the real world. Taken together, these findings illustrate that aphantasia may be protective against particular clinical psychopathologies.

In sum, this literature review shows how aphantasia might influence the lives of people who experience it. In some ways, aphantasics are no different from people with intact visual imagery (e.g., short-term memory, and attention), though they may use different cognitive strategies. In other domains, aphantasia may be associated with slight weaknesses (e.g., in autobiographical memory, and imagination), yet in other ways altogether aphantasia may be *beneficial* (e.g., protecting against imagery-induced fear). Despite this growing understanding of aphantasia, there are, however, many questions that remain unanswered. In this thesis, I endeavour to build on this existing literature, by providing an in-depth examination of the clinical, neuropsychological, and forensic profile in aphantasia. In doing so, I will further our knowledge of aphantasia, and its associated characteristics.

Before outlining the questions that this thesis intends to answer, it is also important to consider the implications of aphantasia for mental imagery research more broadly. In particular, the existence of aphantasia has potential implications for our understanding of the very *function* of visual imagery. The question of whether visual imagery has an important role in 'typical' cognitive functioning has been a long debated question within what is termed 'The Imagery Debate' (Kosslyn, Thompson, & Ganis, 2006; Pearson & Kosslyn, 2015; Pylyshyn, 1973). One side of this debate suggests that information is stored propositionally, in a descriptive or symbolic like state (Pylyshyn, 1973, 2002). Here, visual imagery is not deemed to be important or necessary to typical cognitive functioning. Alternatively, the other side of this debate posits that information is stored in a depictive, or pictorial format (Kosslyn et al., 2006; Pearson & Kosslyn, 2015), where mental images are seen to be key to cognition. The existence of aphantasia may suggest that information can be stored propositionally, and that imagery is *not necessary* for

typical cognitive functioning (Zeman et al., 2010). Indeed, there may instead be *individual differences* in the way information is stored and processed within the brain, including depictive formats, propositional formats, or a mixture of the two (Pearson & Kosslyn, 2015). I return to these ideas within the General Discussion (Chapter 6), but highlight for now that placed within this broader theoretical context, aphantasia research not only seeks to improve our understanding of the causes and consequences of imagery absence, but it also has key implications for our understanding of mental imagery and cognition more widely.

Aphantasia: the unanswered questions

This thesis aims to answer a series of key questions about aphantasia, linked to its clinical, neuropsychological, and forensic profile. The first of these questions asks how often aphantasia occurs in the general population (Chapter 2). Surprisingly, although we have a growing understanding of how aphantasia affects people in their day-to-day lives, we still know relatively little about how many people world-wide experience this imagery difference. In the past there have been attempts to estimate the prevalence of aphantasia, but they are limited in their accuracy. The most commonly cited prevalence for aphantasia is approximately 2% (as cited in Fulford et al., 2018; Watkins, 2018; Zeman et al., 2015, 2020, etc.), provided first by Betts (1909), then later by Faw (2009). One substantial problem with this 2% estimate is that it does not estimate the prevalence of aphantasia as it is known today. Instead, it purports to reflect the prevalence of a *total absence of visual imagery*, which is not by definition aphantasia as we recognise it now (i.e., imagery is absent *or dim/vague*; Keogh & Pearson, 2018; Zeman et al., 2015). In a more recent study, Zeman et al. (2020) attempted to estimate the prevalence of aphantasia using the *VVIQ* (Marks, 1973). Zeman and colleagues found that 0.7% of 1288 individuals surveyed experienced a *complete lack* of visual imagery, and 2.6% experienced what the authors termed '*moderate aphantasia*' (p. 428), which they classified as a VVIQ score of 16-23. Again, these rates do not reflect aphantasia as we recognise it today (i.e., absent/vague/dim imagery; *VVIQ* = 16-32). Another limitation of this study by Zeman et al. (2020) is that visual imagery was explicitly mentioned in the recruitment materials (personal communication), which is problematic for any epidemiological study, because this can encourage people with especially high (or low) imagery to take part. Taken together, it is clear from previous research that the true prevalence rate for aphantasia remains unknown.

In this thesis I will therefore first set out to provide a reliable prevalence estimate for aphantasia (see Chapter 2). Using methodology designed to minimise recruitment biases (i.e., no mention of visual imagery or aphantasia in any recruitment materials), and an appropriate threshold on the measure commonly used to diagnose aphantasia (*VVIQ*; Marks, 1973), I will provide a prevalence estimate for aphantasia according to current definitions (i.e., where visual imagery is *absent or vague/dim*). I will additionally provide a prevalence estimate for the most severe subtype of aphantasia, where visual imagery is *completely absent*. I endeavour to provide answers about how often aphantasia occurs in the general population, providing useful prevalence estimates that can be used by future researchers.

Besides the question of how often aphantasia occurs in the general population, another question that remains unanswered is whether aphantasia is linked with other neurodevelopmental traits. In Chapter 3, I aim to answer this question by examining the relationship between aphantasia and two other conditions: *synaesthesia* (see below), and *autism spectrum conditions* (henceforth *autism*). Why might we expect there to be a link between aphantasia, synaesthesia, and autism? This is because all three conditions have been linked to atypical imagery phenomenology – aphantasia and autism to *weak* imagery, and synaesthesia to *strong* imagery. All three conditions have also been linked to atypical sensory phenomenology: I will describe these sensory differences in more detail throughout this thesis, but for now it is important to understand that mental imagery has sensory qualities simply because it represents a mental simulation of the sensory world. For example, when we create a visual image of the house we grew up in, we re-instantiate the sensory visual qualities of the house (e.g., our mental image represents its size, shape, visual textures, and so on). I return to the sensory qualities of imagery throughout this thesis. For now, I briefly describe synaesthesia and autism, and the ways these traits might relate to aphantasia.

Synaesthesia is a neuropsychological condition characterised by a merging of the senses, where everyday stimuli trigger unusual additional sensations or experiences (Simner, 2019; Simner & Hubbard, 2013). For example, the experience of colours is triggered by listening to music in *sound-colour synaesthesia* (Ward, Huckstep, & Tsakanikos, 2006), or triggered by numbers and letters in *grapheme-colour synaesthesia* (Simner, Glover, &

Mowat, 2006; Simner, Mulvenna, et al., 2006). Important to the purposes of this thesis, some researchers have suggested that strong visual imagery may be required to experience synaesthesia (Barnett & Newell, 2008; Price, 2009; Ward, 2019b), since synaesthetic experiences are often 'like mental images', albeit far stronger (Ward, 2019b, p.2). In support of this, many studies have shown that synaesthetes self-report higher levels of imagery than non-synaesthete controls (e.g., Barnett & Newell, 2008; Chiou, Rich, Rogers, & Pearson, 2018; Chun & Hupé, 2016; Mealor, Simner, Rothen, Carmichael, & Ward, 2016; Spiller, Jonas, Simner, & Jansari, 2015). Synaesthetes also show stronger priming than non-synaesthetes on imagery-based binocular rivalry tests (see above for the methodology of these tasks) (Chiou et al., 2018), and are faster to make decisions about letters held in mind as visual images (Spiller & Jansari, 2008). These links between imagery and synaesthesia suggest that aphantasics – given their *low* visual imagery – might be less likely to experience synaesthesia, or indeed unable to experience it at all.

Alternatively, another line of reasoning suggests that imagery might *not* be essential for experiencing synaesthesia. This view proposes that heightened imagery previously described in studies of synaesthesia might, theoretically, be driven by a methodological recruitment confound (Simner, 2013). This theory points out that most studies tend to test self-referred synaesthetes, and it may be exactly these synaesthetes who have the *strongest* imagery. Simply put, individuals with strong imagery (and therefore vivid synaesthesia) might be more aware of their synaesthetic experiences, and therefore more likely to self-refer (Simner, 2013). This view is supported by other studies showing no difference between synaesthetes and non-synaesthete controls in self-reported visual imagery (Seron, Pesenti, Nöel, Deloche, & Cornet, 1992; Spiller & Jansari, 2008; Ward, Ipser, et al., 2018). If true, this theory predicts that aphantasia and synaesthesia should be able to co-occur in the same individual. In this thesis, I test these opposing theoretical standpoints by examining the link between aphantasia and synaesthesia (see Chapter 3). If visual imagery is necessary for the development of synaesthesia, then I should find no occurrences of synaesthesia within an aphantasic population (and vice versa). However, if visual imagery is *not* necessary for experiencing synaesthesia, then I anticipate finding that synaesthesia and aphantasia can co-occur in the same individuals.

Another related question addressed in Chapter 3 is whether aphantasia might influence the *type* of synaesthesia experienced. This question is linked to a distinction between *associator* and *projector* synaesthesia (Dixon, Smilek, & Merikle, 2004). Associator synaesthetes experience their colours internally, as 'known' associations, or as visual imagery within the mind's eye. Projectors, on the other hand, experience their colours vividly as part of the outside world (e.g., projected onto the letters of written words in grapheme-colour synaesthesia) (Dixon et al., 2004). One school of thought, put forward by Simner (2013), suggests that associator and projector synaesthetes may differ in their visual imagery strength, with the latter having stronger imagery. This theory suggests that visual imagery may be particularly strong in projector synaesthetes because it provides a platform by which synaesthetic colours can be projected into the real word (Simner, 2013). Indeed, projectors tend to report stronger levels of visual imagery than associators (Amsel, Kutas, & Coulson, 2017). If true, this theory predicts that aphantasic synaesthetes – given their low imagery – will experience more associator-like than projector-like traits (if aphantasia and synaesthesia can indeed co-occur; see above). In this thesis I test this hypothesis (see Chapter 3), and in doing so, I endeavour to clarify the relationship between synaesthesia, aphantasia, and mental imagery.

Another neurodevelopmental condition that might be theoretically related to aphantasia is autism. People with autism experience a range of sensory and developmental differences, for example in communication and social skills (e.g., Holt & Yuill, 2014; Hopkins, Yuill, & Branigan, 2022), attention (e.g., heightened attention-to-detail; Baron-Cohen, Ashwin, Ashwin, Tavassoli, & Chakrabarti, 2009), behaviour (e.g., repetitive behaviours, interests and activities; Zandt, Prior, & Kyrios, 2007), and sensory sensitivity (e.g., sensory overwhelm; Robertson & Simmons, 2015; see below). Important to the purposes of this thesis, people with autism also experience deficits in *imagination* (American Psychiatric Association, 2013). Although imagination and imagery are not the same thing, they may be linked. We know they are not the same thing entirely, since people without mental imagery can still imagine, that is, they can still mentally generate events and objects from the past, present, or future, or indeed events and objects that cannot even exist in the real world (but they do this *without* experiencing a quasi-perceptual visual image). People with aphantasia can also be highly creative, and will often fiercely reject the suggestion from imagers that aphantasics cannot 'imagine' (Julia Simner, personal communication). However, imagery and imagination may yet be linked

in subtle ways. Visual imagery vividness (measured by the *VVIQ*; Marks, 1973) can predict the amount of detail people produce about imagined future events (e.g., in terms of sensory, spatial and emotional information) (D'Argembeau & Van der Linden, 2006). Consistent with this, people with aphantasia – as a group – report less sensory detail when they describe imagined events, compared to imaging controls (Dawes et al., 2022; Milton et al., 2021). Visual imagery strength might therefore be inherently linked to imagination, in many (but not all) individuals. I return to these ideas in Chapter 3.

Given the potential link between imagery and imagination, and the lower imagination reported in autism, I hypothesise that autism and aphantasia may be tied in some way via weak visual imagery. This hypothesis is further supported by the fact that the *Autism Spectrum Quotient* (*AQ*; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001), a questionnaire commonly used to measure autistic traits, contains a question asking specifically about the strength of visual imagery within its 'imagination' subscale. In this thesis, I therefore ask whether people with aphantasia have heightened autistic traits (see Chapter 3) relative to people with intact visual imagery, perhaps with particular weaknesses in the imagination domain of autistic symptomatology (even without the AQ question linked to aphantasia per se).

Beyond investigating how aphantasia might intersect with synaesthesia and autism, this thesis also aims to examine the *broader sensory profile* of aphantasia. First, I will examine whether aphantasics experience low imagery in other (non-visual) sense domains. Following this, I investigate whether imagery levels predict levels of sensory sensitivity across the senses (see below, and Chapter 4). In past literature, aphantasia has been defined exclusively as an absence of *visual* imagery (Milton et al., 2021; Zeman et al., 2015, 2016, 2020). However, strictly speaking, mental imagery also encompasses the other sensory modalities, including, for example, auditory imagery, gustatory imagery, olfactory imagery, tactile imagery, motor imagery, bodily imagery, and feeling imagery. In this thesis, I ask whether the visual imagery weakness seen in aphantasia extends to other sense domains. This question is motivated by personal reports from aphantasic writers who suggest that multiple modalities may be affected (Watkins, 2018), and findings from a study which showed that 54% of 2000 aphantasics surveyed reported weak/absent imagery in *all* sensory modalities, when asked one question about their imagery across senses (Zeman et al., 2020). In this thesis, I build on this literature to

provide an in-depth investigation of the sensory imagery profile in aphantasia using a series of standardised contemporary questionnaires (see Chapter 4). This investigation will enhance our understanding of the sensory phenomenology of aphantasia, and shed light more broadly on how imagery might be linked across the senses.

In examining the sensory profile of aphantasia, this thesis will additionally set out to investigate whether levels of mental imagery map onto levels of *sensory sensitivity* within the different sense domains (Chapter 4). Sensory sensitivity is a trait characterised by under- or over-reactivity to incoming sensory stimulation from the outside world (e.g., Robertson & Simmons, 2013). For example, someone who is visually hypersensitive might find lights glaring and seek to avoid them, whereas someone who is visually *hypo*sensitive might experience low responsivity to bright lights (known as 'sensory dampening') and actively *seek out* visual simulation (e.g., flicking their fingers in front of their eyes) (Ben-Sasson et al., 2009; Bogdashina, 2003). Sensory sensitivities are clinically relevant since they are commonly characteristic of autism (Bogdashina, 2003; Robertson & Simmons, 2015), but they also vary on a continuum within the general population (Robertson & Simmons, 2013). As well as reporting sensitivities in day-to-day life which can be measured using standardised questionnaires (e.g., *Glasgow Sensory Questionnaire*; Robertson & Simmons, 2013), people with sensory sensitivities also tend to experience more visual effects and visual discomfort ('pattern glare') in the *Pattern Glare Task*. In this task, alternating black and white parallel lines (known as 'gratings') are shown at particular spatial frequencies, some of which trigger visual sensitivity (Braithwaite et al., 2013; Ward et al., 2017; see Chapter 4). In this thesis, I will use the Pattern Glare Task along with standardised questionnaires to ask for the very first time whether the way we mentally image sensory information in our mind (mental imagery) maps onto the way we experience sensory information in the real world (sensory sensitivity). In other words, if aphantasics tend to experience weaker imagery, do they also experience lower sensory sensitivity? And if so, will this relationship between imagery and sensory sensitivity manifest within different (multiple) sense domains?

Finally, the last empirical question of this thesis asks whether aphantasia influences face processing (Chapter 5). This is a forensically important question, since face recognition is often crucial in criminal investigations when witnesses identify an offender (Frowd et al., 2019; Home Office, 2017). In this thesis, I ask the question of whether visual imagery

facilitates face recognition. In other words, does being able to visualise a face mean you are more likely to recognise it? Inconsistent findings from prior literature mean that the answer to this question is surprisingly not yet known. Some previous studies have shown no difference in face recognition for people low and high in imagery (measured by the *VVIQ*; Marks, 1973), in tasks that require participants to inspect target faces then subsequently identify them from among distractors (McKelvie, 1994). In contrast, other previous research shows that people with the clinical condition congenital prosopagnosia (who experience poor facial recognition from birth; Behrmann & Avidan, 2005) tend to experience low levels of visual imagery (Grueter et al., 2007; Grüter, Grüter, Bell, & Carbon, 2009), with particular weaknesses for face imagery (Tree & Wilkie, 2010). Important to the purposes of the present thesis, these findings from the prosopagnosia literature suggest that imagery may – at least to some degree – aid face recognition.

A small number of previous studies have already examined face recognition in aphantasia, in some preliminary ways. In one study, people with aphantasia self-reported more problems with face recognition than imaging controls, based on a single question about their face recognition abilities (Zeman et al., 2020). In another study, aphantasics also scored higher on the *Prosopagnosia Index* (Shah, Gaule, Sowden, Bird, & Cook, 2015), a questionnaire that measures difficulties with facial recognition in day-to-day life (Milton et al., 2021). Away from self-report, however, aphantasics performed no differently from controls in *behavioural* tests of face recognition, such as the *Famous Face Test* (where participants identified famous people such as Elton John from four similar faces; Milton et al., 2021). Taken together, this brief and conflicting literature review highlights that the relationship between visual imagery and face recognition remains somewhat unclear. In this thesis I therefore aim to investigate this relationship in further detail (see Chapter 5), by examining face recognition in people with and without aphantasia, using both standardised self-report questionnaires and behavioural measures. If visual imagery does indeed aid facial recognition, then I predict aphantasics to experience more problems with face recognition than imaging controls.

Another related question addressed in Chapter 5 is whether aphantasia affects the ability to build facial composites (using specialised computer software "EFIT6"; Visionmetric Ltd, 2019a). This is a forensically significant question, since composites of criminal suspects are often created by witnesses in criminal investigations (Frowd et al., 2019;

Home Office, 2017). Given their self-reported face recognition difficulties (Milton et al., 2021; Zeman et al., 2020), aphantasic individuals might construct less accurate facial composites of a target face from memory than people with intact visual imagery. However, aphantasics are less susceptible to producing '*false memories*' than imagers (i.e., they are less likely to draw objects from memory that were not actually present in an original scene; Bainbridge et al., 2021), and more broadly (high) imagery has been linked to a greater incidence of recalling the details of memories incorrectly (e.g., Dobson & Markham, 1993; Hyman & Pentland, 1996; Tomes & Katz, 1997). Their reduced proneness to produce false memories may therefore mean that aphantasics are unimpaired in their ability to create facial composites. In the present thesis, I therefore investigate for the very first time whether people with and without aphantasia differ in their ability to construct face composites of a target face from memory (Chapter 5).

Summary

In this thesis, I build on the existing literature to further our understanding of congenital aphantasia, a life-long absence of visual imagery. In a series of four experimental chapters, I examine the profile of aphantasia from a clinical, neuropsychological, and forensic perspective. Each of my experimental chapters has been written in an 'article format', meaning they have either already been published in peer-reviewed journals (Chapters 2, 3 and 4) or are currently under review (Chapter 5). At the start of each experiment, I present a short summary to link each experimental chapter the one following. To start with, my first experimental chapter (Chapter 2) examines how often aphantasia occurs in the general population. Here, I provide a prevalence estimate for aphantasia according to contemporary definitions. In Chapter 3, I examine how aphantasia intersects with two other sensory differences linked to atypical imagery: *synaesthesia* and *autism*. Here, I test previous theories linking synaesthesia to strong imagery (Barnett & Newell, 2008; Price, 2009; Ward, 2019b), by asking whether synaesthesia and aphantasia can co-occur. I also ask whether aphantasia might influence the *type* of synaesthesia experienced (i.e., associator or projector synaesthesia). I also investigate whether people with aphantasia tend to experience high levels of autistic traits, given that both conditions might be linked via weak visual imagery (or at least weak imagination) (Baron-Cohen et al., 2001; Milton et al., 2021).

In Chapter 4 I will investigate the broader sensory profile in aphantasia, asking whether the visual imagery weakness characteristic of aphantasia extends to the other senses (e.g., weak auditory imagery, gustatory imagery, tactile imagery, etc.). I will also investigate – for the very first time – the link between mental imagery and sensory sensitivity, using both self-report and behavioural methods. I will ask whether aphantasics experience lower sensory sensitivity, given their low imagery. Finally, in Chapter 5 I will investigate the relationship between visual imagery and face processing. Using self-report and behavioural methods, I will compare facial composite creation and face recognition in people with and without aphantasia. If visual imagery aids face recognition, then I predict aphantasics to be less accurate at face recognition and composite creation than imaging controls. Alternatively, aphantasics may be just as accurate at constructing face composites as imaging controls given their resilience to 'false memories' (see above). Overall, Chapter 5 will address a forensically significant question, because face recognition and composite construction play a key role in witness testimony (Frowd et al., 2019; Home Office, 2017). In sum, this thesis will serve to enhance our understanding of aphantasia and its associated characteristics, and mental imagery more broadly.

Chapter 2

The prevalence of aphantasia (imagery weakness) in the general population

Chapter Summary

In my first empirical chapter I ask how often aphantasia occurs in the general population. This was an important question to start my experimental investigations with since previous attempts to estimate the prevalence of aphantasia have been inconsistent (Betts, 1909; Faw, 2009; Zeman et al., 2020). Hence, the true prevalence rate of aphantasia remained unknown at the start of my PhD. Using methodology suitable for a targeted prevalence study (i.e., designed to reduce recruitment biases), and the most widely used aphantasia diagnostic (*VVIQ*; Marks, 1973), I provide a prevalence estimate for aphantasia in line with contemporary definitions, where visual imagery is absent *or* vague/dim. Additionally, I also provide a prevalence estimate for the most extreme subtype of aphantasia, where visual imagery is *completely absent*. My aim here was to improve our understanding of how often aphantasia occurs, and provide robust prevalence estimates that can be used by future researchers.

Full citation

Dance, C. J., Ipser, A., & Simner, J. (2022). The prevalence of aphantasia (imagery weakness) in the general population. *Consciousness and cognition*, *97*, 103243.

Abstract

Visual mental imagery is the ability to create a quasi-perceptual visual picture in the mind's eye. For people with the rare trait of *aphantasia*, this ability is entirely absent or markedly impaired. Here, we aim to clarify the prevalence of aphantasia in the general population, while overcoming limitations of previous research (e.g., recruitment biases). In Experiment 1, we screened a cohort of undergraduate students (n502) using the Vividness of Visual Imagery Questionnaire (Marks, 1973) and found that 4.2% had aphantasia. To establish the reliability of our estimate, we then screened a new sample of people (n502) at an online crowdsourcing marketplace, again finding that approximately four percent (3.6%) had aphantasia. Overall, our combined prevalence from over a thousand people of 3.9% – which shows no gender bias – provides a useful index for how commonly aphantasia occurs, based on measures and diagnostic thresholds in line with contemporary aphantasia literature.

Introduction

When asked to visualise a sunset, most people can create a mental picture of it within their 'mind's eye'. However, the clarity and robustness of visual mental imagery varies from person to person. Some people report almost perception-like vividness while others – people with *aphantasia* – struggle to generate any visual imagery at all. Aphantasia is most often congenital and is defined as a life-long neurodevelopmental trait in which visual mental imagery is either entirely absent, or dim and vague (Zeman et al., 2015, 2020). Although the phenomenon was first recognised 140 years ago (Galton, 1880), it was only recently named and more formally acknowledged in the scientific literature (Zeman et al., 2015). Studies have since shown that people with aphantasia have their own set of comorbidities (e.g., Dance, Jaquiery, et al., 2021) and characteristics (see below; e.g., Bainbridge et al., 2021; Dawes et al., 2020; Keogh & Pearson, 2018; Milton et al., 2021; Wicken et al., 2021; Zeman et al., 2015, 2020). But one of the most fundamental features of aphantasia remains unclear: its prevalence within the general population. Previous prevalence estimates have varied widely (e.g., from 0.7-15.3%; Betts, 1909; Faw, 2009; Zeman et al., 2020) but so too have the robustness of their methodologies. Exploring these differences below leads us to believe that the true prevalence of aphantasia remains unknown. We therefore aim to provide a prevalence estimate for aphantasia, replicating our finding across two different samples from the general population, and using careful unbiased methodology. We begin with a brief overview of aphantasia and the key concepts motivating our research question.

Although most people use visual imagery habitually in everyday life, people with aphantasia are often outwardly unaffected, living their lives fully without realising they are in any way different (Zeman et al., 2015, 2016). They often interpret 'visual imagery' as merely metaphor, and are surprised to learn others might 're-see' (i.e., visually image) in order to achieve the same visual knowledge they themselves possess without this. Aphantasics are equally accurate in tasks measuring visual and number working memory when compared to imaging controls (although they use different strategies; see Keogh et al., 2021). Aphantasics are also just as accurate at other short-term memory tasks such as remembering details of a story (Milton et al., 2021). Despite their fairly typical lives (Zeman et al., 2015, 2016), people with aphantasia do, however, perform differently in some psychometric and cognitive tests. For example, people with aphantasia self-report the autistic traits of significantly poorer imagination and social skills (Dance, Jaquiery, et

al., 2021) within the *Autism Spectrum Quotient* questionnaire (*AQ*; Baron-Cohen et al., 2001). Aphantasics also self-report problems with face recognition (Milton et al., 2021; Zeman et al., 2020), experience less detailed autobiographical memories and imagined events, and have less sensorily rich night time dream-worlds than imaging controls (Dawes et al., 2020; Milton et al., 2021). However, they perform better in some spatial tasks (Keogh et al., 2021), experience less physiological fear-responses when asked to read and imagine fear evoking passages (Wicken et al., 2021), and are less susceptible to artificially-induced pseudo-hallucinations (see Königsmark et al., 2021). Hence in some tasks people with aphantasia are no different to imagers, while in other tasks they are significantly worse, or significantly better.

Some have suggested that impairments in visual imagery are nothing more than deficits in metacognition (i.e., the ability to introspect accurately about our thoughts; Flavell, 1979). In other words, rather than lacking visual imagery, aphantasics might perhaps lack the ability to accurately report it (de Vito & Bartolomeo, 2016). This hypothesis relies on the comparison problem: since one individual cannot experience the mental state of another, two people could *rate* their visual imagery very differently but yet *experience* comparable visual imagery. However, this hypothesis fails to take into account that aphantasia can also be acquired (as opposed to congenital), and that people with acquired cases are still able to remember what it was once like to experience visual imagery. For example, individual MX experienced the onset of aphantasia late in life following a coronary angioplasty (Zeman et al., 2010). This late onset allows him to compare his mental imagery before and after aphantasia, showing that reports of aphantasia are not simply a failure in comparing one's own imagery accurately to others.

Other research too suggests aphantasia is a genuine impairment in visual imagery. People are – in fact – rather good at evaluating the vividness of their own visual imagery, because self-reports correlate well with performance in objective tasks. Keogh and Pearson (2018) demonstrated this using a binocular rivalry paradigm in which participants are first asked to visualise a colour, and are then presented with a colour to each eye – one of which is the previously-imaged colour (see also Pearson et al., 2008). Prior imaging of a colour means it tends to become dominant (i.e., the colour seen when presented in the binocular display), and the extent of this dominance is indicative of visual imagery strength (Pearson, 2014; Pearson et al., 2008). Indeed, the dominance in binocular rivalry

correlates with the spectrum of scores in the *VVIQ* (Keogh & Pearson, 2018; Pearson et al., 2011), and imagery vividness on a trial-by-trial basis (Pearson et al., 2011). Importantly, this dominance effect is absent for aphantasic individuals (Keogh & Pearson, 2018). Taken together, this suggests that people - including those with aphantasia - generally do have good metacognition about their own visual imagery abilities.

Other research, too, provides further support for aphantasia being driven by a deficit in visual imagery. A study by Bainbridge et al. (2021) aimed to characterise the nature of aphantasia using an online drawing task. In this task, a group of aphantasics and non-aphantasic controls were presented with a series of pictures and were subsequently asked to draw the pictures from memory, draw them whilst visible on screen, and also to recognise them. Aphantasic drawings from memory contained fewer objects, and less colour, but they were equally accurate in the size and spatial location of objects. There was no difference between groups in the ability to recognise the pictures, or in drawings produced when the picture was visible (i.e., not from memory). These results suggest that aphantasia is linked to a selective deficit in visual imagery or object memory, while spatial abilities remain intact (see also Dawes et al., 2020; Keogh & Pearson, 2018). Taken together, these studies provide behavioural evidence for the existence of aphantasia, and validate aphantasia as a neurodevelopmental trait defined by an absence of visual object imagery.

Despite an increase in our understanding of the phenomenology and cognitive profile of aphantasia, we still know little about how commonly the trait occurs in the general population. Here, we intend to answer this question. First, it is important to clarify how aphantasia is recognised: within the existing literature, the most widely used approach is to administer the *Vividness of Visual Imagery Questionnaire* (*VVIQ*; Marks, 1973, 1995). This 16-item measure requires participants to generate a series of visual images and then rate their imagery on a 5-point scale, as either (5) *perfectly clear and as vivid as normal vision*, (4) *clear and reasonably vivid*, (3) *moderately clear and vivid*, (2) *vague and dim*, or (1) *no image at all, you only 'know' that you are thinking of the object*. Importantly for our purposes, the VVIQ is a robust way to measure visual imagery. Scores on the VVIQ map well onto behavioural measures of imagery (e.g., binocular rivalry paradigms, see above; Keogh & Pearson, 2018; Pearson et al., 2011), resting state fMRI data (aphantasics validated by the VVIQ have lower connectivity between frontal and visual

areas of the brain; Milton et al., 2021), and also onto convergent measures such as questionnaire and behavioural (pattern glare) tasks for sensory sensitivity (Dance, Ward, & Simner, 2021). An overall test score of 16-32 (i.e., participants tending to score 1 or 2 on each item) captures the phenomenology of aphantasia, where imagery is either entirely absent or vague/dim (Zeman et al., 2015). Indeed, aphantasics often err between these judgements, sometimes reporting dim 'flashes' of visual imagery which come and go (Zeman et al., 2016, p.337), or a dim dark outline that almost immediately dissolves. While the exact range of phenomenologies is yet to be charted, a test-score of 16-32 captures the absent or highly impoverished imagery of aphantasics, so we adopt this here. Although a small number of studies have used other thresholds (e.g., 23 or 25; Bainbridge et al., 2021; Zeman et al., 2020), or simply allowed aphantasics to self-identify (Dawes et al., 2020; Keogh & Pearson, 2018), we suggest that this meaningful and motivated score of 16-32 would be useful for future studies, and we use this here. At the same time, we will *additionally* identify a subtype of the most extreme cases of aphantasia, where visual imagery is reported to be *completely* absent (i.e., a VVIQ score of 16).

Previous attempts to estimate the prevalence of aphantasia have varied widely, and have included certain methodological limitations. The most commonly cited prevalence for aphantasia, published first by Betts (1909) and then by Faw (2009) a century later, is approximately 2% (as cited in Fulford et al., 2018; Watkins, 2018; Zeman et al., 2015, 2020 etc.). Faw (2009) describes surveying 2,500 people about their sensory imagery ability, including 750 people surveyed for visual imagery. Faw found that 2.1% of people reported a *total absence* of visual imagery. Hence, one problem with this estimate is that the trait Faw reported is not, by definition, the same trait we now recognise as aphantasia today. Indeed, when including people who rated their imagery as absent *or* vague/dim, the figure was as high as 10-11%. A second problem is that Faw's estimates come from a single question about visual imagery, rather than a multi-item questionnaire of the type used for diagnosis today (e.g., *VVIQ*) raising questions over the robustness of classifications.

In far earlier research, Betts (1909) provides a prevalence estimate for aphantasia using a questionnaire developed to assess imagery within different sensory modalities (*Bett's Questionnaire Upon Mental Imagery*). Out of a total of 143 people who completed his questionnaire across four experiments, 15.3% reported what we would recognise today

as aphantasia (visual imagery is absent or vague/dim). However, his estimates for visual imagery varied dramatically across experiments, from 6% to 40%. Equally variable were his estimates for cases where visual imagery was entirely absent (i.e., 6.7%, but ranging from 2% to 19%). Although Betts (1909) measured imagery with a multi-item tool, his wide range of prevalence estimates leave us with the conclusion that there was some bias or inconsistency, either in his recruitment, in his cohorts, or in his testing. Either way, it is unclear from these studies where the true prevalence of aphantasia lies.

Finally, one recent study by Zeman et al. (2020) estimated the prevalence of aphantasia in a large community sample using the aphantasia questionnaire typical of contemporary aphantasia research (*VVIQ*; Marks, 1973, 1995). This study provided an estimate for aphantasia where imagery is completely absent (i.e., 0.7% of 1288 individuals surveyed), and an estimate for 'moderate aphantasia' (p.428) which was categorised in this study as a VVIQ score of 16-23 (2.6%; see General Discussion). However, this study did not provide an estimate for aphantasia where visual imagery is *absent or vague/dim* (VVIQ score 16-32). Moreover, visual imagery was explicitly mentioned in their recruitment materials (personal communication) which may have encouraged individuals with particularly low (or indeed high) imagery to take part. We point out that their 'recruitment bias' here was in fact a careful strategy to encourage individuals from a local participant Biobank (the EXTEND study; http://exeter.crf.nihr.ac.uk/extend/) to take part. As such, their study was not designed to produce a prevalence estimate *per se*. Hence, although their study is very informative in many ways, it does not have the methodology for a targeted prevalence estimate.

In summary, it remains unclear from previous research where the true prevalence of aphantasia lies. In the present study, we therefore aim to fill this gap. Experiment 1 tests a general population sample taken from the student body at the University of Sussex. Experiment 2 aims to replicate our prevalence finding using a large online general population sample. Together, our studies screened over a thousand people. Both studies were designed to minimise recruitment biases by not mentioning visual imagery or aphantasia in any recruitment materials. To diagnose aphantasia, we used the index of visual imagery from contemporary aphantasia research (*VVIQ*; Marks, 1973), categorising people as aphantasic if they rated their imagery overall as absent or

vague/dim (a test score of 16-32). In doing this, our intention was to provide a useful population prevalence for aphantasia for future researchers.

Experiment 1: What is the prevalence of aphantasia in a student population?

Methods

Participants

Our participants were 502 undergraduate students registered at the University of Sussex (*M* age = 19.93, *SD* = 2.42; 415 females; 85 males; 2 another gender). Participants were enrolled as part of teaching or for participation credits within the Psychology undergraduate degree course.

Materials and procedure

Participants completed the *VVIQ* (see below; Marks, 1973) among a set of other tasks to be reported elsewhere. The questionnaire was completed via our online in-house testing platform (www.syntoolkit.org). Participants were provided with a URL to the testing site and completed the study from a classroom (if taking part for teaching credits) or from their own homes (if taking part for participation credits). Importantly, our methodology was designed to ensure that we obtained a reliable prevalence estimate for aphantasia. First, there was no mention of visual imagery or aphantasia in any recruitment materials, and participants were entirely unaware of our research goals when agreeing to take part. This ensured that we did not encourage people with aphantasia – or indeed any unusual imagery phenomenology – to take part in the study. Second, participants completed the VVIQ before any other measures in our study, to ensure that VVIQ performance was not influenced by other tasks. Finally, participants were told that it was important to pay attention to the instructions, and respond as accurately as possible. All participants provided informed consent prior to testing, and ethical approval for both experiments came from the *University of Sussex Sciences and Technology Cross-Schools Research Ethics Board*.

Vividness of Visual Imagery Questionnaire (VVIQ; Marks, 1973). In this questionnaire participants were asked to think of a series of four scenarios (e.g., “A country scene which involves trees, mountains and a lake”). For each scenario,

participants were asked to imagine four aspects of the scene (e.g., "The colour and shape of the trees") and to "consider carefully the picture that comes before your mind's eye". Participants rated the vividness of their visual imagery on scale from 1 ("No image at all, you only 'know' that you are thinking of the object") to 5 ("Perfectly clear and as vivid as normal vision"; see Introduction). The questionnaire was scored by summing responses to all 16 questions, giving possible total scores ranging from 16-80. In line with previous research showing that aphantasic individuals tend to rate their visual imagery as absent or vague/dim (Keogh & Pearson, 2018; Zeman et al., 2015), individuals were classified as aphantasic if they scored between 16 and 32 (i.e., ≤ 32).

Results

Using scores from the VVIQ, we divided participants into those with and without aphantasia. Within our total cohort (N = 502), there were 21 participants (18 female, 2 male and 1 another gender; *M* age = 19.33, *SD* = 1.07) scoring in the aphantasia range of 16-32 (*M* = 23.43, *SD* = 5.64), giving an overall aphantasia prevalence of 4.2% (95% CI [2.75, 6.31]; CIs calculated using R version 3.5.1, R Core Team, 2018; Hmisc version 4.5, Harrell, 2021). This means there were 481 people (397 female, 83 male, 1 another gender; *M* age = 19.95, *SD* = 2.45) without aphantasia scoring above 32 on the VVIQ (*M* = 57.65, *SD* = 10.02).

To understand the phenomenology of aphantasia in more detail, we next divided our aphantasic cohort into those reporting a *complete absence* of visual imagery (scoring 16 on the VVIQ), and those reporting *vague/dim* visual imagery (scoring 17-32 on the VVIQ). There were 5 individuals reporting a complete absence of visual imagery (4 females, 1 male; *M* age = 19.40, *SD* = 1.14), and 16 people reporting imagery that was vague/dim (14 females, 1 male, 1 another gender; *M* age = 19.31, *SD* = 1.08). This provides a prevalence of 1.0% (95% CI [0.43, 2.31]) for the most severe cases of aphantasia where no visual imagery is experienced at all, and 3.2% (95% CI [1.97, 5.11]) for those reporting vague/dim visual imagery.

Discussion

We provide a new prevalence estimate for aphantasia using the *VVIQ* (Marks, 1973), and a threshold for aphantasia where imagery tends to be absent or vague/dim (score of 16-

32). Our prevalence of aphantasia was 4.2%, with a prevalence of 1.0% for the most severe cases (no visual imagery whatsoever). Importantly, our methodology ensured that we did not especially encourage people with aphantasia to take part in the study, and in doing so we ensured as much as was possible a random screening for aphantasia. In the following study we seek to replicate our finding in a new cohort of individuals.

Experiment 2: What is the prevalence of aphantasia? A replication

To establish the reliability of the prevalence estimates provided in Experiment 1, our second study sought to replicate and further investigate the prevalence of aphantasia within another general population sample. Our methods again ensured that we screened for aphantasia randomly (i.e., in a way that minimised recruitment biases), and we again utilised the most widely used aphantasia diagnostic (*VVIQ*; Marks, 1973).

Methods

Participants

We tested 502 participants (208 female, 292 male, 2 another gender; *M* age = 36.55, *SD* = 11.53) recruited through the online crowdsourcing market place known as Amazon's Mechanical Turk (MTurk; a platform for collecting quality data; Casler, Bickel, & Hackett, 2013). Our participants came from two English speaking countries (UK and USA) with the following levels of education: 8.6% were schooled up to 16 years (only), 19.5% were schooled to 17-19 years (only), 40.2% of participants had an undergraduate degree, 25.5% had a Masters degree, 3.6% had a PhD, and 2.6% did not state their education level. Participants were paid $2 (for our 15 minute test) and again provided informed consent prior to testing.

Materials and procedure

As before, we used the VVIQ to classify individuals as aphantasics, and details of the VVIQ are described in Experiment 1. Again, individuals were classified as aphantasic if they scored between 16-32 on the VVIQ (i.e., tending to rate their visual imagery as absent, or vague/dim). Participants were provided with a URL to the testing site (www.syntoolkit.com) and completed the study from their own homes. Steps were again taken to ensure that we obtained a reliable prevalence of aphantasia. We not only kept

participants naïve to our focus and emphasised accuracy in responding (as in Experiment 1) but we also administered two attention-check questions, one before the VVIQ ('If you are paying attention please select the 'agree' option'), and one afterwards ('If you are paying attention please select the 'disagree' option'). All participants responded correctly, showing that they paid good attention and responded to the questions accurately.

Results

To examine the prevalence of aphantasia in our sample, we again divided participants into those with and without aphantasia. Within our total cohort (N = 502), there were 18 participants (8 female, 10 male; *M* age = 36.89, *SD* = 15.30) scoring in the aphantasia range of 16-32 (*M* = 24.44, *SD* = 5.48), giving an overall aphantasia prevalence of 3.6% (95% CI [2.28, 5.60]). In addition to these individuals classified as aphantasic there were 484 people (200 female, 282 male, 2 another gender; *M* age = 36.54, *SD* = 11.39) without aphantasia (VVIQ >32; *M* = 60.43, *SD* = 10.72). As before, we also divided our aphantasic cohort into those reporting a *complete absence* of visual imagery (scoring 16 on the VVIQ), and those reporting *vague/dim* visual imagery (scoring 17-32 on the VVIQ). There were 3 aphantasic individuals reporting a complete absence of visual imagery (2 females, 1 male; *M* age = 42.67, *SD* = 15.14), and 15 further people reporting imagery that was vague/dim (6 females, 9 males, *M* age = 35.73, *SD* = 15.59). This provides a prevalence of 0.6% (95% CI [0.20, 1.74]) for the most severe cases of aphantasia where no visual imagery is experienced at all, and 3.0% (95% CI [1.82, 4.87]) for those reporting vague/dim visual imagery.

To establish the reliability of our aphantasia prevalence estimates, we then compared our online general population sample to the estimates elicited from our student population in Experiment 1. The prevalence estimates for aphantasia (visual imagery that is absent/vague/dim) were largely aligned across our student population (4.2%) and online general population (3.6%; see Table 1), and there was no statistical difference in these estimates ([χ^2 (1, *N* = 1004) = .107, *p* = .744; chi-square test with Yates continuity correction]). To investigate this null effect further, we calculated a Bayes factor (BF) of .07 (using R version 3.5.1, R Core Team, 2018; Bayes Factor version 4.2, Morey & Rouder, 2018). On the assumption that BF < .33 is taken as evidence for the null hypothesis, while BF > 3 as evidence for the alternative hypothesis (Dienes, 2014), our

finding (BF = .07) provides strong evidence in favour of the null hypothesis (no difference in prevalence of aphantasia across our two samples). When looking within aphantasics themselves, we again find no difference in the prevalence rate of those reporting the strongest form of aphantasia (i.e., a *complete* absence of visual imagery) when comparing our student population (1.0%) and our online general population sample (0.6%; [χ^2 (1, N = 1004) = .126, p = .723; chi-square test with Yates continuity correction]; BF = .04).

Experiments 1 and 2 provide comparable prevalence estimates for aphantasia. Next, we collapsed across both samples, to provide a single robust prevalence estimate for aphantasia from the 1004 people sampled (see Figure 1 for the distribution of VVIQ scores). We also examine whether the prevalence of aphantasia is similar in males and females. Within this sample (N = 1004), there were 39 participants (26 female, 12 male, 1 another gender; M age = 27.44, SD = 13.56) scoring in the aphantasia range (VVIQ 16-32; M = 23.90, SD = 5.52), giving an overall aphantasia prevalence of 3.9% (95% CI [2.85, 5.27]). Within these individuals, there were 8 aphantasics (0.8%; 95% CI [0.40, 1.56]; 6 female, 2 male; M age = 28.13, SD = 14.54) reporting a *complete* absence of visual imagery (scoring 16 on the VVIQ), and 31 individuals (3.1%; 95% CI [2.18, 4.35]; 20 female, 10 male, 1 another gender; M age = 27.26, SD = 13.55) reporting dim/vague imagery (scoring 17-32 on the VVIQ). In addition to these individuals classified as aphantasics, there were 965 people (597 female, 365 male, 3 another gender; M age = 28.27, SD = 11.70) without aphantasia (VVIQ >32; M = 59.04, SD = 10.47).

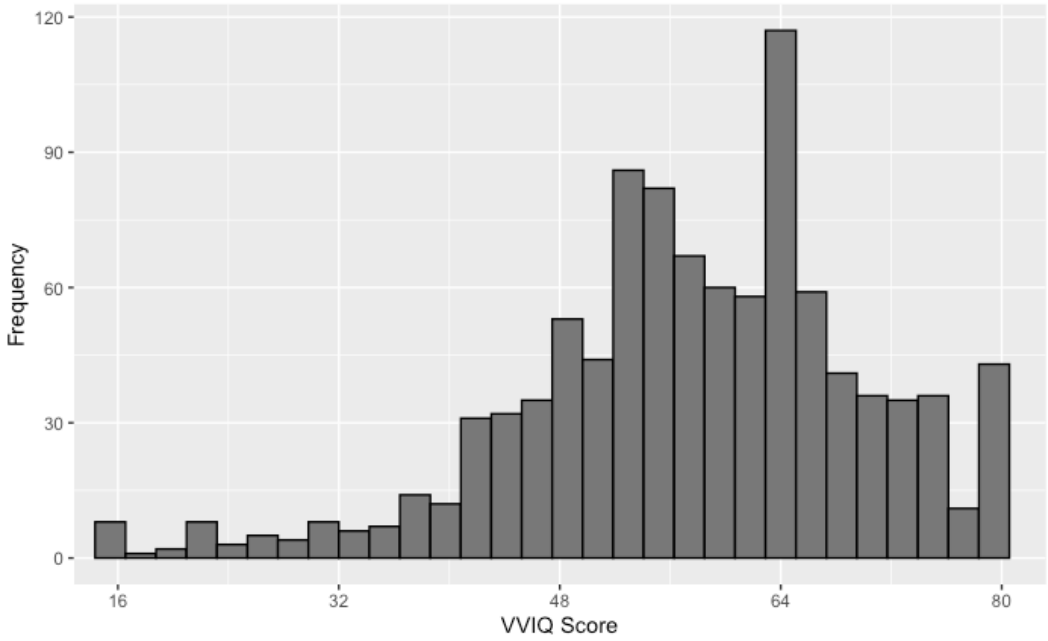

Figure 1. A histogram to show the distribution of VVIQ scores across both our student and online general population samples.

Looking next within gender, of the 623 females in the sample, there were 26 aphantasics (scoring 16-32 on the VVIQ; M = 23.23, SD = 5.49), of whom 6 reported a *complete* absence of visual imagery (i.e., 1.0% of females). Of the 377 males in the sample, there were 12 aphantasics (scoring 16-32 on the VVIQ; M = 25.17, SD = 5.78), of whom 2 reported a *complete* absence of visual imagery (i.e., 0.5%). The prevalence estimates for aphantasia in females (4.2%; 95% CI [2.86, 6.04]) and males (3.2%; 95% CI [1.83, 5.48]) were not statistically different, ([χ^2 (1, N = 1000) = .388, p = .533; chi-square test with Yates continuity correction]; BF =.08), and neither were estimates for the most extreme subtype of aphantasia where visual imagery was completely absent (females 1.0%, 95% CI [0.44, 2.09]; males 0.5%, 95% CI [0.15, 1.91]; [χ^2 (1, N = 1000) = .143, p = .706; chi-square test with Yates continuity correction]; BF = .03). This suggests that there are no gender differences between men and women in the occurrence of aphantasia (see Table 1). There was one aphantasic among the four participants who self-identified as another gender. These four participants were excluded from the analyses of gender simply given the small number of data points, but we flag this for future researchers.

Table 1

Frequency of aphantasia occurrence (i.e., visual imagery absent or vague/dim; scoring 16-32 on the VVIQ) within the student general population sample (Experiment 1) and online general population sample (Experiment 2), as well as overall i.e., both samples are collapsed together (displayed as a percentage of each sample in brackets). Numbers are broken down by gender where there are sufficient data points for statistical comparison (i.e., females vs males).

	Experiment 1 Student population (*n*=502)	**Experiment 2** Online population (*n*=502)	**Overall** Student and Online populations (*n*=1004)		
			Overall	**Female (n=623)**	**Male (n=377)**
Aphantasic	21 **(4.2%)**	18 **(3.6%)**	39 **(3.9%)**	26 **(4.2%)**	12 **(3.2%)**
95% CI	[2.75, 6.31]	[2.28, 5.60]	[2.85, 5.27]	[2.86, 6.04]	[1.83, 5.48]

Discussion

In Experiment 2 we replicated the methodology from Experiment 1 in a novel population. The prevalence estimates for aphantasia were similar across studies (Experiment 1 = 4.2%; Experiment 2 = 3.6%). Our estimates were also similar for the subtype of aphantasics where visual imagery is entirely absent (Experiment 1 = 1.0%; Experiment 2 = 0.6%). When collapsing across both samples for a cohort size of 1004, our final prevalence estimate for aphantasia was 3.9% (visual imagery is absent or dim/vague) and 0.8% for cases where visual imagery is completely absent. The prevalence of aphantasia was similar in males (absent/vague/dim: 3.2%; completely absent: 0.5%) and females (absent/vague/dim: 4.2%; completely absent: 1.0%). Our methods sought to avoid recruitment biases, and we used the measure for aphantasia used in contemporary literature (VVIQ; Marks, 1973). As a result of our methods, we suggest that our studies provide useful prevalence estimates for aphantasia.

General Discussion

In our study we investigated how commonly aphantasia occurs in the general population. We screened for aphantasia using the VVIQ in two separate population samples, and minimised the risk of recruitment biases by avoiding mention of visual imagery or aphantasia in our recruitment materials. Previous research has included varying or ad hoc VVIQ thresholds, including 16-32 (Dance, Jaquiery, et al., 2021; Dance, Ward, et al., 2021; Wicken et al., 2021) but also 16-23 (Milton et al., 2021; Zeman et al., 2020), 16-25 (Bainbridge et al., 2021), or even self-declaration regardless of VVIQ score (Dawes

et al., 2020; Keogh & Pearson, 2018, 2021). Here, we applied the widely used aphantasia diagnostic (VVIQ; Marks, 1973) and set a threshold (VVIQ 16-32) in line with the original definition of aphantasia (Zeman et al., 2015) as imagery which is absent *or* vague/dim (i.e., aphantasics report a complete absence of visual imagery, sometimes interspersed by flashes of imagery that tend to be dim and fleeting; Keogh & Pearson, 2018; Zeman et al., 2015, 2016).

The prevalence of aphantasia in our combined cohort of over a thousand participants was 3.9% (Experiment 1: 4.2%; Experiment 2: 3.6%). We also found that 1 in 5 aphantasics (0.8%) had no visual imagery whatsoever (Experiment 1: 1.0%; Experiment 2: 0.6%). We found additionally that the prevalence of aphantasia appears to be similar in males (3.2%) and females (4.2%). Our analyses of gender may have lacked statistical power given small sample sizes (26 females, 12 males), but we point out that our results are consistent with more general findings showing men and women experience visual imagery that is equally vivid (Campos, 2014; Richardson, 1995), and with previous studies showing no gender bias in aphantasia (Dawes et al., 2020; Zeman et al., 2020; though we add here the important contribution of (a) random sampling, and (b) naïve recruits). We also had too few datapoints to statistically explore rates of aphantasia in our four participants who self-identified as another gender, although we flag for future researchers that one in four were aphantasic.

How do our results compare with other estimates of aphantasia described in earlier literature? The most frequently cited “aphantasia” prevalence had been 2% (Betts, 1909; Faw, 2009; as cited in Fulford et al., 2018; Watkins, 2018; Zeman et al., 2015, 2020 etc.). However, this figure was for a *complete absence of visual imagery*, and rose as high as 10-11% (Faw, 2009) or 15.3% (Betts, 1909) when all aphantasics were included (i.e., whose imagery was absent *or* vague/dim; see Introduction). These rates appear inflated compared to our own, but in Faw’s case, this could be attributed to the use of a lone question (rather than full questionnaire, as used here). In Bett’s research, however, his estimates varied so widely across four experiments (6-40%) it is likely his study engendered some type of recruitment bias, or other methodological problem.

More recently a study by Zeman et al. (2020) provided a prevalence estimate for aphantasia by testing a large community sample, using the same standard diagnostic tool

used here (*VVIQ*; Marks, 1973). Although their study is very informative in many ways, Zeman et al. (2020) described their interest in imagery/aphantasia during recruitment, making their study unsuitable for prevalence per se. Despite this concern, we have found their data to be highly similar to our own. Where our results are comparable (i.e., for the subtype within aphantasia where visual imagery is completely absent; i.e., VVIQ score = 16) we find our values to be similar (Zeman et al. (2020) [N=1288] = 0.7%; here [N=1004] = 0.8%; χ^2 = .0007, p = .979 with Yates correction; BF = 0.02). Although Zeman et al. (2020) did not provide a prevalence estimate for aphantasia as defined here (i.e., VVIQ 16-32), we can nonetheless re-inspect our own data for their category of 'moderate aphantasia' (VVIQ 16-23). Here, we find our estimates again converge (Zeman et al. (2020) = 2.6%; here =1.9%; χ^2 = .859, p = .354 with Yates correction; BF = .05). These comparisons suggest that the sample used by Zeman et al. (2020) were not unduly biased by the recruitment materials. On close inspection this may be because their cohort had volunteered to engage in psychology experiments irrespective of topic (i.e., individuals from a biobank called the EXTEND study; http://exeter.crf.nihr.ac.uk/extend/). With a cohort already motivated for testing, any novel recruitment materials (e.g., with information about imagery) may not have swayed their interest further. We therefore offer our findings as largely equivalent to those of Zeman et al. (2020), but designed specifically with prevalence in mind. Importantly, for the first time we also provide a prevalence for the general class of aphantasia – where visual imagery is absent *or* vague/dim (VVIQ 16-32).

One potential limitation to our study is that we identified aphantasia using a self-report questionnaire (*VVIQ*; Marks, 1973). The ability of questionnaires to accurately measure subjective experience has been questioned. For example, people may respond in a socially desirable manner (e.g., if one deems having good visual imagery as desirable, their scores may be inflated). In addition, although we tested the reliability of our estimates by comparing across two different general population samples, we did not test for reliability *within* each sample (i.e., we tested participants at one time point, so cannot ascertain whether VVIQ scores remained consistent over time). However, previous research has shown that scores on visual imagery questionnaires (such as the VVIQ) are not heavily influenced by social desirability (Durndell & Wetherick, 1975; Hiscock, 1978; but see also Allbutt, Ling, Rowley, & Shafiullah, 2011), and importantly, self-reported visual imagery maps well onto objective visual imagery tasks. As we saw in the Introduction,

the degree of colour dominance in imagery based binocular rivalry tests correlates positively with VVIQ scores (Pearson et al., 2011; Rademaker & Pearson, 2012), suggesting that people generally have good metacognition about their own visual imagery abilities. In addition to this, scores on the VVIQ also map onto resting state fMRI data (see below; aphantasics have less connectivity between frontal and visual areas of the brain; Milton et al., 2021), and onto convergent behavioural tasks (e.g., aphantasics score lower in questionnaire and behavioural (pattern glare) measures of sensory sensitivity; Dance, Ward, & Simner, 2021). Taken together, it appears that the VVIQ is a highly useful measure for classifying aphantasia. Moreover, the *VVIQ* is the most widely accepted test for diagnosing aphantasia in contemporary literature, suggesting that our prevalence figure fits within this body of research (e.g., Bainbridge et al., 2021; Dance, Jaquiery, et al., 2021; Dance, Ward, et al., 2021; Dawes et al., 2020; Jacobs et al., 2018; Keogh & Pearson, 2018, 2021; Milton et al., 2021; Wicken et al., 2021; Zeman et al., 2020, 2015). Future researchers seeking to replicate our findings may wish to also administer a behavioural measure of aphantasia (e.g., the drawing task of Bainbridge et al., 2021) or test visual imagery at multiple time points to ensure reliability. For now, we suggest that our use of the *VVIQ* is a helpful way of diagnosing aphantasia consistent with contemporary literature, and has provided meaningful prevalence estimates for contemporary researchers.

Another potential limitation to our study relates to the representativeness of our samples. To ensure the reliability of our prevalence estimates, it is important to ask whether our samples are representative against the population as a whole. One possibility is that we may have under-estimated prevalence since both our samples were highly educated: our first sample came from a university undergraduate population, and our second had relatively high levels of education (see *Experiment 2, Participants*) compared to population-wide rates (OECD, 2021; Office for National Statistics, 2011). This may be important, since visual imagery is linked to skills that can improve attainment, such as reading comprehension (e.g., Borduin, Borduin, & Manley, 1994; Center, Freeman, Robertson, & Outhred, 1999; Commodari, Guarnera, Di Stefano, & Di Nuovo, 2020; Joffe, Cain, & Marić, 2007), and working memory (e.g., Baddeley & Andrade, 2000; Keogh & Pearson, 2011, 2014). We may therefore have tested samples with higher visual imagery than usual, and thereby under-estimated the prevalence of aphantasia in the population at large. Alternatively, we may have *over*-estimated aphantasia since rates of

aphantasia may be higher in populations with degrees, given aphantasics tend to migrate towards degree-linked jobs in science and maths (Zeman et al., 2020; UK Comission for Employment and Skills, 2016), although we point out that our sample in Experiment 1 were psychology students. Equally, we may have *over*-estimated the prevalence of aphantasia because people with aphantasia score highly on certain autism traits (Dance, Jaquiery, et al., 2021), as do participants from our online recruitment platform MTurk in Experiment 2 (Chandler & Shapiro, 2016; Eriksson, 2013; Palmer, Paton, Enticott, & Hohwy, 2015). It seems unlikely, however, that levels of autistic traits might have influenced our aphantasia prevalence rates in Experiment 2, since we found a similar prevalence rate in Experiment 1, albeit with a different profile of participants entirely (i.e., psychology students are not likely to have elevated autistic traits; Stewart & Austin, 2009). Taken together with our careful methodology designed to minimise recruitment biases, we suggest that our prevalence estimates are therefore indicative of the occurrence of aphantasia in the general population. Finally, our results also converge with a prevalence estimate from the community sample of Zeman et al. (2020) (see above). These converging results suggest that we are likely to have produced a reliable prevalence for aphantasia typical of the population at large.

We now move on to discussing potential avenues for future research. While the current study shows the prevalence of aphantasia within two general population samples, further research might explore the prevalence of aphantasia in special populations. This is particularly important for populations where features of aphantasia might explain symptoms characteristic of the special population. One population that is worth exploring is people with *autism spectrum conditions* (henceforth *autism*). People with autism experience sensory, developmental, and communicative differences, but also weaknesses in social skills and imagination (American Psychiatric Association, 2013; Baron-Cohen et al., 2001). A recent study by Dance, Jaquiery, et al. (2021) found that people with aphantasia also share these latter two non-clinical autistic traits (*AQ*; Baron-Cohen et al., 2001), having specific weaknesses in imagination, and social skills. Visual imagery may aid imagination by enhancing the ability to create and construct imagined events or scenarios etc. and may assist social skills by helping people to perspective take, or understand the mental state of others (see Dance, Jaquiery, et al., 2021). It is therefore possible that features of aphantasia may play a part in driving the weaknesses in imagination and social skills seen in autism. On the other hand, aphantasic people report

lower levels (compared to imaging controls) of other symptoms commonly characteristic of autism, such as sensory sensitivity (over- or under-responsivity to incoming sensory information; Dance, Ward, et al., 2021). As such, it would be useful to screen for aphantasia within populations who have been diagnosed with autism to see whether the two conditions often co-occur, and to investigate further whether particular features of aphantasia may contribute to autism symptomatology.

Another important avenue for future research is investigating the neural mechanisms that underpin aphantasia. Visual imagery activates similar brain areas as visual perception (Dijkstra, Ambrogioni, Vidaurre, & van Gerven, 2020; Dijkstra et al., 2017, 2019; Pearson et al., 2015), including visual cortex (Cui et al., 2007; Dijkstra et al., 2017; McGeown et al., 2012; Winlove et al., 2018), and areas involved in visual attention (e.g., parietal cortex) and information retrieval/integration (e.g., medial frontal cortex) (see Dijkstra et al., 2017, 2019; Pearson et al., 2015). People with aphantasia have reduced (resting state fMRI) connectivity between top-down control (prefrontal cortex) and visual-occipital areas relative to very vivid imagers (Milton et al., 2021). They also show less BOLD activity in the left anterior parietal cortex when asked to imagine (as compared to perceive) famous faces/places, relative to both average and high imagers (Milton et al., 2021). In addition to this, low imagers show less activation in visual cortex during imagery-linked tasks compared to vivid imagers (mental rotation; Logie, Pernet, Buonocore, & Della Sala, 2011). Taken together, this suggests that aphantasics may rely less on visual areas when completing imagery-based tasks, and – importantly – that a reduction in connectivity between frontal and visual areas may be a potential neural mechanism underpinning aphantasia (Milton et al., 2021).

Future research might also use paradigms to measure cortical excitability in the visual cortex (e.g., transcranial magnetic stimulation), which we have hypothesised as being low in aphantasia, given their lower rates of sensory sensitivity (see Dance, Ward, et al., 2021; but see also Keogh, Bergmann, & Pearson, 2020; who did not test aphantasics, but found imagery correlates *negatively* with visual cortex excitability, at least in those with an ability to image). Finally, topological analysis using EEG has been used to characterise brain activation during perception and imagery: vivid imagers have more similarity (overlap) in global topological brain activation patterns during imagery (visualising their head turned to one side) and perception (physically turning their head to one side) relative

to people with less vivid imagery (Ibáñez-Marcelo, Campioni, Phinyomark, Petri, & Santarcangelo, 2019). Future research may therefore also use electrophysiological methods such as this to examine differences in brain activity in people with and without aphantasia. In sum, examining neural mechanisms in aphantasia will be a fruitful and important avenue for future research.

In conclusion, our study provides prevalence estimates for aphantasia within the general population. We have shown that aphantasia – where visual imagery is *absent* or *dim/vague* – is experienced by 3.9% of the population, with men and women equally affected (but a lack of statistical information in people of another gender). For the most extreme subtype of aphantasia in which individuals experience a *complete* absence of visual imagery, the prevalence is 0.8%. We suggest that our prevalence estimates can be used as comparison rates for future research endeavouring to characterise the nature of aphantasia, and further understand its relationship with other psychological and neurodevelopmental traits.

Chapter 3

What is the relationship between aphantasia, synaesthesia and autism?

Chapter Summary

In Chapter 2, I provided a prevalence estimate for aphantasia, establishing how often this neurodevelopmental trait occurs within the general population. In Chapter 3 I ask a second epidemiological question, examining the link between aphantasia and synaesthesia on the one hand, and aphantasia and autism on the other. Integral to the motivation for this chapter, all three conditions have been previously characterised by differences in sensation and imagery. Here, I test previous theories that have linked synaesthesia to strong imagery (Price, 2009; Ward, 2019b), by assessing whether aphantasia (weak imagery) and synaesthesia can co-occur. I also ask whether aphantasia affects the *type* of synaesthesia experienced. Specifically, I assess whether aphantasia gives rise to a less vivid form of synaesthesia (characterised as having more 'associator' than 'projector' traits; see below). In this chapter, I also examine whether aphantasics tend to experience high levels of autistic traits given that both conditions are linked to weak visual imagery (or at least weak imagination; Baron-Cohen et al., 2001). Overall, Chapter 3 therefore aims to show how aphantasia intersects with other clinical or neuropsychological traits linked to sensory differences.

Full citation

Dance, C. J., Jaquiery, M., Eagleman, D. M., Porteous, D., Zeman, A., & Simner, J. (2021). What is the relationship between Aphantasia, Synaesthesia and Autism? *Consciousness and Cognition*, *89*, 103087.

Abstract

For people with aphantasia, visual imagery is absent or markedly impaired. Here, we investigated the relationship between aphantasia and two other neurodevelopmental conditions also linked to imagery differences: synaesthesia, and autism. In Experiment 1a and 1b, we asked whether aphantasia and synaesthesia can co-occur, an important question given that synaesthesia is linked to strong imagery. Taking grapheme-colour synaesthesia as a test case, we found that synaesthesia can be objectively diagnosed in aphantasics, suggesting visual imagery is not necessary for synaesthesia to occur. However, aphantasia influenced the type of synaesthesia experienced (favouring 'associator' over 'projector' synaesthesia - a distinction tied to the phenomenology of the synaesthetic experience). In Experiment 2, we asked whether aphantasics have traits associated with autism, an important question given that autism – like aphantasia – is linked to weak imagery. We found that aphantasics reported more autistic traits than controls, with weaknesses in imagination and social skills.

Introduction

When asked to imagine a sunset, most people can create a picture of the scene within their mind's eye, as a visual mental image. People with *aphantasia* lack this ability (no clear 'mental picture' in the mind) even though they can describe what a sunset looks like (Zeman et al., 2015). Scientists have known for at least 100 years that some individuals do not experience visual imagery (Galton, 1880) but aphantasia has only recently entered mainstream research (e.g., Jacobs et al., 2018; Keogh & Pearson, 2018; Zeman et al., 2015, 2020). The present study adds to this body of literature by investigating the relationship between aphantasia and two other neurodevelopmental conditions, *synaesthesia* and *autism spectrum conditions* (henceforth, autism). As we shall see below, synaesthesia has been linked with higher-than-average mental imagery (e.g., Barnett & Newell, 2008; Price, 2009; Spiller et al., 2015), raising the question of whether people with aphantasia can experience synaesthesia at all. In contrast, autism has been linked with *poor* imagery (or at the very least, with poor *imagination* - a distinction we will elaborate on below) raising the question of whether aphantasia and autism might overlap in some way. We present two studies testing these ideas, but begin with a brief overview of key concepts.

Aphantasia is most often a congenital or life-long condition, in which individuals experience an absence of visual imagery, or imagery that is only vague or dim (Zeman et al., 2015). Previous prevalence estimates for aphantasia range from 0.7% (Zeman et al., 2020) to 2.1% (Faw, 2009) and 6.7% (Betts, 1909) for individuals with no mental imagery at all, but are as high as 10-11% or 15.3% (Faw and Betts respectively) for imagery that is either absent or dim/vague. Since visual imagery has been linked to a number of aspects important in everyday life (e.g., autobiographical memory recall; short term memory recall; task-oriented motivation; Keogh & Pearson, 2014; Schacter & Addis, 2007; Vasquez & Buehler, 2007), it may seem surprising that many aphantasics live their lives without knowing they are different (Watkins, 2018; Zeman et al., 2016). However, some aphantasics describe problems with autobiographical memory and face recognition (Zeman et al., 2020), and the condition may have implications for visual processing strategies and even career choices (e.g., aphantasics are less likely to enter the arts, and more likely to work in science and maths; Zeman et al., 2020). In the current study we will further investigate the implications of aphantasia, asking whether it is related to two other conditions: autism and synaesthesia. We explore these conditions in particular

because they too have potentially atypical imagery phenomenology, and this raises questions about how they might intersect with aphantasia.

Synaesthesia is a neurodevelopmental trait in which the senses intermingle (e.g., Simner, 2019; Simner & Hubbard, 2013). For example, listening to music can trigger the experience of colours for *sound-colour synaesthetes* (Ward, Huckstep, et al., 2006) or tastes in the mouth for *sound-taste synaesthetes* (Beeli, Esslen, & Jäncke, 2005), while *sequence-space synaesthetes* think about time units and other sequences (letters, numbers) in spatial patterns (e.g., the synaesthete might feel that days unfold in a zigzag line across the visual field, or that calendar months wrap around the body; e.g., Havlik, Carmichael, & Simner, 2015). Here, we focus on a variant of synaesthesia known as *grapheme-colour synaesthesia*, in which colours are triggered by numbers or letters (Meier & Rothen, 2013; Simner, Glover, et al., 2006; Simner, Mulvenna, et al., 2006; Ward, Simner, & Auyeung, 2005). This variant is relatively common (affecting 1.1-1.5% of people; Carmichael, Down, Shillcock, Eagleman, & Simner, 2015; Simner & Carmichael, 2015; Simner, Mulvenna, et al., 2006), easily diagnosed (e.g., Eagleman, Kagan, Nelson, Sagaram, & Sarma, 2007), and the best understood synaesthesia to date (Simner, 2019). Important to our purposes here, researchers have suggested that heightened visual imagery may be required for experiencing synaesthesia generally (Barnett & Newell, 2008; Price, 2009). This means that aphantasics, due to their absence of visual imagery, may be less likely to experience synaesthesia – or indeed, unable to experience it at all. In our study we therefore ask whether synaesthesia (associated with high imagery) is precluded in those with aphantasia (associated with low imagery), or if not, whether it influences the type of synaesthesia experienced (e.g., does a less vivid synaesthesia result?). We briefly review the literature linking synaesthesia with imagery below.

Links between synaesthesia and imagery have been found in multiple domains. Grapheme-colour synaesthetes, for example, score significantly higher than controls on visual imagery measures such as the *Vividness of Visual Imagery Questionnaire* (*VVIQ*; Marks, 1973) (Barnett & Newell, 2008; Chiou et al., 2018), *Verbalizer-Visualizer Questionnaire* (Richardson, 1978) (Kirby, Moore, & Schofield, 1988; see Meier & Rothen, 2013), *Sussex Cognitive Styles Questionnaire* (*Imagery sub-scale*; Mealor et al., 2016), and the *French Questionnaire on Mental Imagery-51* (*FQMI-51*; Chun & Hupé,

2016). Indeed, synaesthetes report higher imagery than controls across multiple sensory modalities, and especially for modalities involved in their synaesthesia (e.g., higher taste imagery for people with synaesthetic tastes; Spiller et al., 2015). Chiou et al. (2018) have recently backed up these self-reports with a binocular rivalry test. Grapheme-colour synaesthetes showed stronger priming than controls when asked to imagine a colour that subsequently appeared in the display. The degree to which the imaged colour becomes dominant in the rivalry reflects strength of visual imagery (Pearson, 2014; Pearson et al., 2008, 2011) and indeed, aphantasic individuals do not show this priming effect (Keogh & Pearson, 2018). Grapheme-colour synaesthetes are also faster than controls when making decisions about letters held in mind as a visual image (Spiller & Jansari, 2008). Some have even gone as far as to suggest that synaesthesia may be nothing more than imagery itself: Price (2009) proposed that heightened imagery may simply allow certain individuals to become aware of naturally occurring cross-modal associations held by all people (e.g., associations between space and time, triggering 'sequence-space synaesthesia'; Price, 2009). Either way, this body of research claims that vivid visual imagery plays a pivotal role in the development of synaesthesia (see Ward, 2019b for a discussion). If true, this would suggest aphantasic individuals may be less likely to experience synaesthesia – or may not be able to experience synaesthesia at all.

An alternative view is that heightened imagery in synaesthesia may be simply a referral bias. Simner (2013) pointed out that most studies have tended to test self-referred synaesthetes (who have made some effort to reach out to researchers), and it may be precisely those synaesthetes with the strongest imagery (i.e., most intense synaesthesia) who self-refer. Support comes from studies that have *failed* to show superior self-reported visual imagery in some cohorts of synaesthetes (Seron et al., 1992; Spiller & Jansari, 2008; Ward, Ipser, et al., 2018) and that heightened imagery seems to emerge only in synaesthetes who are most *aware* of their synaesthesia (Ward, Ipser, et al., 2018). Simner (2013) therefore suggests that enhanced visual imagery may be a characteristic of certain synaesthetes (e.g., those who self-refer) but not others. A recent study by Brang and Ahn (2019) sought to eliminate the bias of self-referred synaesthetes, by screening a general population sample for synaesthesia, and using a double-blind recruitment in which neither participant nor researcher knew who were the target group. As a consequence, their grapheme-colour synaesthetes were no different to controls in their imagery vividness (measured by the *VVIQ*) although they did still differ in how often they used imagery

(measured by the *Subjective Use of Imagery Scale, SUIS;* Reisberg, Pearson, & Kosslyn, 2003). Another recent study also sought to eliminate the self-referral bias, again by screening for synaesthesia in the general population (rather than asking synaesthetes to come forward). Spiller, Harkry, McCullagh, Thoma, and Jonas (2019) found that scores in a test assumed to identify grapheme-colour synaesthetes (or synaesthesia-like behaviour) did correlate with scores in a test assumed to tap mental imagery ('*Animal Tails Test*'; Farah, Levine, & Calvanio, 1988). However, we suggest here that neither test may have met its intended goal: the test for grapheme-colour synaesthesia was successfully passed by 24% of their sample (instead of the known synaesthesia prevalence of 1-2%), and their test for mental imagery – although widely accepted as such – can be performed just as well by aphantasics (who have no imagery) as by controls (Milton et al., 2021; Zeman et al., 2010)[2]. From this brief literature review it seems evident that the relationship between visual imagery and synaesthesia is somewhat complex and remains unclear. We note here that if synaesthesia were *not* causally linked to heightened imagery, we would anticipate finding cases of grapheme-colour synaesthesia in aphantasic individuals (and vice versa). Alternatively, if high imagery is necessary for synaesthetes, we would expect synaesthesia in aphantasic individuals to be absent.

A related hypothesis is that aphantasia might influence the *type* of synaesthesia experienced. This hypothesis is built on the distinction between 'projector synaesthetes' and 'associator synaesthetes' (Dixon et al., 2004). For projectors, synaesthetic colours feel like part of the outside world (e.g., projected onto the written typeface in grapheme-colour synaesthesia). For associator synaesthetes, colours are less 'veridical', often feeling internal to the body (e.g., appearing in the 'mind's eye') or are simply "known" in some intrinsic way (Ward, Li, Salih, & Sagiv, 2006). The projector-associator distinction has been supported by neuroscientific measures (projectors have greater white matter coherence in inferior temporal cortex than associators; Rouw & Scholte, 2007), and behavioural measures (i.e., projectors are faster to name synaesthetic colours when viewing coloured graphemes, while associators are faster to name text-colour; Dixon et

[2] In this task (*Animal Tails Test*; Farah, Levine, & Calvanio, 1988), participants imagine an animal and state whether its tail is 'short' or 'long'. This was devised as a test of mental imagery although we now know it is entirely possible to perform this task without any imagery at all (aphantasics perform well, and use instead their intact visual memory and visual semantic knowledge). We direct the reader to debates on how scientists use their own imagery abilities to direct their science (Reisberg et al., 2003), which may have played a role in devising this test.

al., 2004; Ward, Li, et al., 2006). Simner (2013) has suggested that this projector-associator distinction may rely on imagery differences, in that projectors may simply be synaesthetes who – aside from synaesthesia – happen to have high mental imagery. Simner (2013) suggests their high imagery may allow synaesthesia to become 'scene-like' to an extreme extent. Supporting this hypothesis, Amsel et al. (2017) showed that self-reported visual imagery is higher in projectors than associators (using the *Object-Spatial Imagery and Verbal Questionnaire*; Blazhenkova & Kozhevnikov, 2009) and that projectors show neurological markers of heightened imagery (i.e., larger lateral occipital N170 responses and smaller P1 event-related potentials to visual stimuli compared to associators; see Ganis & Schendan, 2008; Hirschfeld, Feldker, & Zwitserlood, 2012). In summary, this body of literature proposes two ideas in parallel: that synaesthesia requires high mental imagery, and that projector synaesthetes have higher imagery than associators. Here we directly test these ideas using people with aphantasia. We hypothesise that having aphantasia may *preclude* having synaesthesia, or if it does not, it may make synaesthetes more likely to be associators than projectors.

Our final hypothesis relates to the relationship between aphantasia and autism, another neurodevelopmental condition with links to imagery. People with autism show a range of developmental differences, for example, in social processing, communication, sensory sensitivity, and – importantly for us – deficits in imagination (American Psychiatric Association, 2013). For example, children with autism engage less in imaginative behaviour such as pretend play (Baron-Cohen, 1987; Davis, Simon, Meins, & Robins, 2018; Jarrold, Boucher, & Smith, 1996) and have deficits in imaginative drawing (Low, Goddard, & Melser, 2009; Scott & Baron-Cohen, 1996; Ten Eycke & Müller, 2014). Although visual imagery and imagination are in many ways distinct concepts, they are often confused or seen as inter-changeable (Arcangeli, 2020; Faw, 2009; Thomas, 1999) perhaps especially by people with high visual imagery themselves. Nonetheless, people with aphantasia can use their imagination without having any visual imagery at all[3], as

[3] The final author of this paper has aphantasia, but reports no trouble whatsoever imagining any event, hypothetical or real. (For her, imagination is a sense of 'knowing' rather than 'seeing'). She scores highly on traits related to imagination (e.g., she is high on Openness to Experience; McCrae & Costa, 1987) and scored highly throughout schooling on imaginative activities such as creative writing. However, when imagination is conflated with visual imagery (e.g., in questionnaires asking about the vividness of mental images) she scores poorly. In this sense, imagery and imagination are unrelated. However, there may yet be 'hidden' links between the two (e.g., it is theoretically possible that some elements of a scene might be more difficult to imagine in the absence of mental imagery). We explore these hidden links here.

evidenced by writers and artists with aphantasia (Zeman et al., 2019). Nonetheless, there may yet be a relationship between visual imagery and imagination more broadly. For example, González, Campos, and Pérez (1997) found a positive correlation between scores on *The Torrance Tests of Creative Thinking* (in verbal and figural creative thinking; Torrance & Ball, 1992) and conventional mental imagery tasks (*The Spatial Test of Primary Mental Abilities* and *The Gordon Test of Visual Imagery Control*; Richardson, 1969; Thurstone & Thurstone, 1989). Similarly, when participants are asked to imagine future events, their visual imagery (in the *VVIQ*; Marks, 1973) predicts the amount of sensory, spatial, and emotional information described, as well as the personal importance and significance of the event (D'Argembeau & Van der Linden, 2006). Important for our purposes, the highly influential questionnaire for traits associated with autism (the *Autism Spectrum Quotient; AQ*; Baron-Cohen et al., 2001), contains a question asking about the strength of visual imagery within its 'imagination' subscale. This shows that deficits in visual imagery are directly captured by assessments of autism (albeit under the ambiguous heading of 'imagination'). Here, we therefore investigate whether aphantasia and autism may be linked, through some weak visual imagery ability.

In contrast to the above research, other studies have questioned whether there is any link between autism and low visual imagery at all. Hughes et al. (2018) found no difference between autistic participants and controls in the '*Imagery Ability*' subscale of the *Sussex Cognitive Styles Questionnaire* (*SCSQ*; Mealor et al., 2016), and indeed other research shows that people with higher AQ scores find it easier (not harder) to judge the 'imageability' of words (e.g., does 'blush' lend itself to a visual picture more easily than 'hypothesis'?; Esposito, Dellantonio, Mulatti, & Job, 2016). Other research too shows the cognitive style in autism of 'thinking in pictures', which may point to *elevated* imagery in autistic individuals (e.g., Kana, Keller, Cherkassky, Minshew, & Just, 2006; Kunda & Goel, 2008, 2011; Soulières, Zeffiro, Girard, & Mottron, 2011). Overall, the relationship between visual imagery and autism therefore remains unclear. But if autism is driven, even in part, by a deficit in visual imagery – or if visual imagery deficits are a consequence of autism – we might predict higher levels of autistic traits in people with aphantasia. This might be especially evident in the imagination subscale of the AQ (even when its question about visual imagery is removed).

We therefore present two studies investigating the relationship between aphantasia, synaesthesia and autism. Experiments 1a and 1b explore the interplay between aphantasia and synaesthesia within two separate large-scale samples: (a) members of the general population recruited from the testing cohort *Generation Scotland* (www.ed.ac.uk/generation-scotland) (Smith et al., 2013), and (b) a group of synaesthetes tested via the online platform known as the *Synesthesia Battery* (www.synesthete.org; Eagleman et al., 2007). These cohorts are described further below, but we hypothesise that if high visual imagery is related to (or necessary for) synaesthesia, aphantasia may not co-occur with synaesthesia at all. However, if synaesthesia can exist in the absence of imagery, we may find people with synaesthesia among those with aphantasia (and vice versa), although synaesthesia might manifest differently (more associator-traits than projector-traits within aphantasia). In a final Experiment, we investigate the relationship between aphantasia and traits associated with autism given the possibility of low visual imagery across both aphantasia and autism. We hypothesise that autism traits may be higher in aphantasics than controls from the general population, and perhaps particularly in the AQ subscale of imagination.

Experiment 1a: Can synaesthesia and aphantasia co-occur?

In this study we screened over a thousand members of the general population for both synaesthesia and aphantasia. Both conditions are relatively rare, so we designed our recruitment materials to attract as many cases as possible; i.e., we explicitly described both conditions within our invitation to participate. This was designed to maximise the number of cases for us to examine, but also means that our prevalences of synaesthesia or aphantasia cannot be taken as population-wide estimates (and indeed our study was not designed to be a prevalence count). Importantly, however, our samples can be compared to each other, to ask whether we can find cases of synaesthesia in a cohort of aphantasics (and vice versa).

Methods

Participants

Our participants were 1285 people from the G*eneration Scotland Scottish Family Health Study cohort* (770 female, 515 male) (Smith et al., 2013). Generation Scotland is a large-

scale resource of data available for research purposes (www.ed.ac.uk/generation-scotland), which includes adult volunteers recruited through primary care services across Scotland (e.g., via General Practices, for more information on the recruitment process see Smith et al., 2006). As part of the Generation Scotland project, the *Scottish Family Health Study* (GS:SFHS; Smith et al., 2013) collected genetic and health data for a cohort of over 20,000 of these volunteers. Some of the participants consented to re-contact in the event of further studies, and approximately 6,000 of these also provided email addresses. This allowed us to recruit a sample of participants from these volunteers via email to take part in the present study. Participants were asked to complete a series of questions and tasks online to investigate 'health, minds and bodies of the people of Scotland'. Both imagery and synaesthesia were mentioned explicitly in recruitment.

Four additional participants (2 female, 1 male, 1 other) were excluded because they did not complete the imagery component of our test. In our final sample of 1285 participants, age was not specified but all were 18 years or older. As compensation for taking part, participants were entered into a prize draw for £100.

Materials and procedure

Participants were contacted via email and provided with a URL link to our online study. The study contained the following two measures in the order shown below. These were embedded among other measures to be reported elsewhere. Individuals provided informed consent prior to testing, and ethical approval was provided by the *University of Sussex Cross-Schools Science and Technology Ethics Board*, and ethical approval for the GS:SFHS study was obtained from the *Tayside Committee on Medical Research Ethics (on behalf of the National Health Service)*.

Visual Imagery Questionnaire. We measured visual imagery using an in-house task (see Figure 1), in which we asked participants to think about the building where they lived. Participants were asked to rate 'how much your memory of it is like a picture'. Participants responded with one of five responses (ranging from 0 – 'I have no image or picture at all and I don't know what it looks like', to 5 – 'I have an image like a picture that is perfectly clear and vivid – just as if I were really standing there'; see Figure 1 for all response options). In line with previous research showing that aphantasic individuals rate their visual imagery as absent or vague/dim (Keogh & Pearson, 2018; Zeman et al.,

2015), participants were classified as aphantasic if they responded in the range of 0-2 (i.e., rating their imagery as 'no image at all', or 'vague or fleeting or dim')[4].

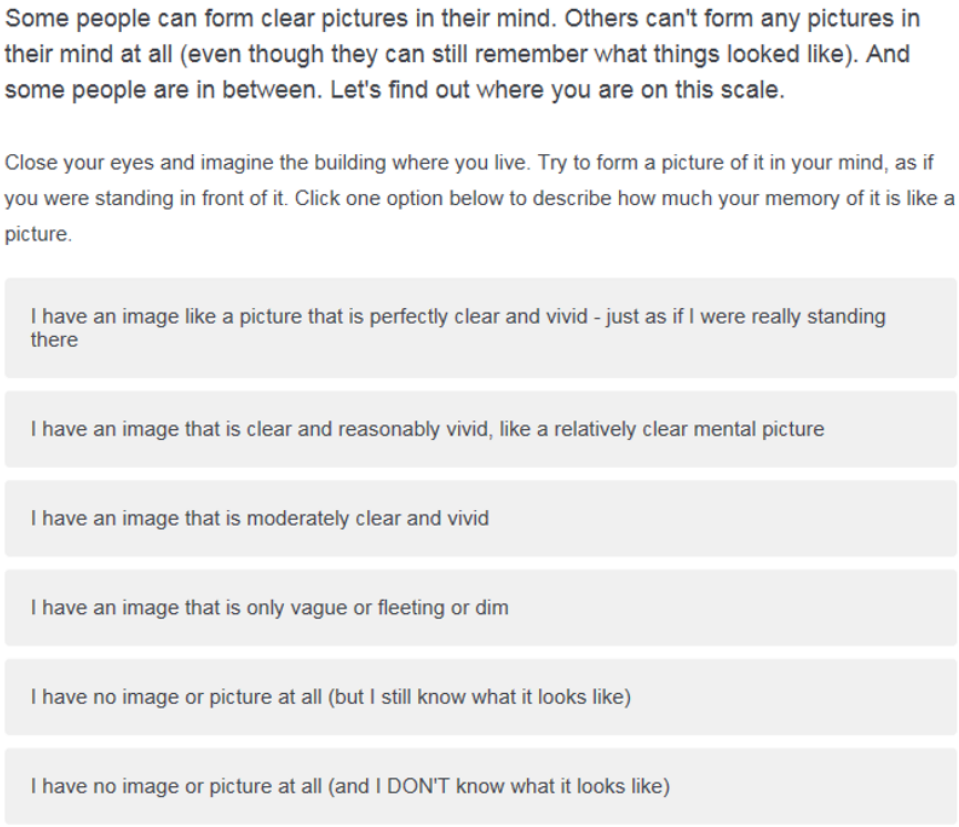

Figure 1. The visual imagery question used to classify aphantasia within the *Generation Scotland* sample.

Grapheme-colour synaesthesia diagnostic. We tested only one variant of synaesthesia to limit the burden on our participants (and we selected grapheme-colour synaesthesia in particular for the reasons given in the Introduction). We diagnosed grapheme-colour synaesthesia by using the 'gold standard' diagnostic (Eagleman et al., 2007), which we replicated to host on our local server. This diagnostic is the most widely used scientific assessment for grapheme-colour synaesthesia worldwide. In the first part of the test, participants are asked to indicate whether they believe they experience grapheme-colour synaesthesia ("Do numbers or letters cause you to have a colour experience?"). Participants *not* reporting grapheme-colour synaesthesia are shown to an exit page (and can then complete the other parts of the test). Participants reporting grapheme-colour

[4] We did not use longer imagery questionnaires (e.g., VVIQ) because our participants were time-restricted, given other measures in the same battery. We return to this issue in Experiment 1b.

synaesthesia are subsequently given an objective diagnostic test to verify their self-report. In this test, participants are identified as synaesthetes if they demonstrate the known characteristic of *consistency-over-time* (i.e., for genuine synaesthetes, associations tend to stay the same over time; e.g., if A is red for any given synaesthete, it is always red when the synaesthete is repeatedly asked). Hence, in the objective test, participants were presented with the letters A-Z and the numbers 0-9, each displayed three times in a randomised order. Participants selected their associated colour for each grapheme using an on-screen colour palette. The test computes the variation in the colours provided for the same grapheme (e.g., summing the distances between the colours chosen for the three 'A' trials). The mean colour variation is then calculated across all graphemes (for information about the testing interface and other measurement details, see Eagleman et al., 2007). If this mean colour distance is small, the individual's synaesthetic colours are consistent (i.e., close in RGB colour space; Eagleman et al., 2007), with scores lower than 1.43 indicating a diagnosis of synaesthesia (Rothen, Seth, Witzel, & Ward, 2013).

Results

Can aphantasia and synaesthesia co-occur?

We first divided our sample into those with and without aphantasia. Within our total cohort (N = 1285), there were 212 participants (115 female, 97 male) on the visual imagery question who scored within the aphantasic range (i.e., they scored within the range of 0-2; M score = 1.33, SD = .48). This gives an overall aphantasia prevalence of 16.5%. Here our classification follows prior literature in defining aphantasics with imagery either vague/dim/fleeting (n71), or completely absent (n141).

We then classified all our participants according to the synaesthesia diagnostic, and found 14 people with grapheme-colour synaesthesia (6 female, 8 male), giving an overall synaesthesia prevalence of 1.1%. There were 2 people (both male) with grapheme-colour synaesthesia within our 212 aphantasic individuals, and 12 people (6 female, 6 male) with grapheme-colour synaesthesia within our 1073 non-aphantasics. This provides a prevalence of grapheme-colour synaesthesia of 0.9% in aphantasics, and 1.1% in non-aphantasics (see table 1), which was a non-significant difference (p = 1.00, Fishers Exact Test). Bayes factors (BF) were calculated (using R version 3.5.1, R Core Team, 2018; BayesFactor version 4.2, Morey & Rouder, 2018) to better understand this null effect. A BF of <.33 is taken as evidence for the null hypothesis (Dienes, 2014), while >3 is taken

as evidence for the alternative hypothesis. We calculated a BF of .022 which provides good evidence in favour of the null hypothesis, suggesting no difference in rates of synaesthesia between aphantasics and non-aphantasics.

We next looked at the prevalence of aphantasia in our synaesthete and non-synaesthete groups. There were 2 aphantasics among our 14 synaesthetes (2 male), and 210 aphantasics among our 1271 non-synaesthetes (115 female, 95 male). This gives an aphantasia prevalence of 14.3% in the synaesthete group, and 16.5% in the non-synaesthete group – again a non-significant difference (p = 1.00, Fishers Exact Test) with a BF of .022, again supporting the null hypothesis. Overall, our results suggest that high imagery is therefore not a pre-requisite of synaesthesia, and that is possible to have synaesthesia with little or no imagery at all. Finally we found that mean imagery scores for synaesthetes (M = 3.86; SD = 1.46) and non-synaesthetes (M = 3.84; SD = 1.30) were statistically equivalent, (t(13.23) = -0.047, p = .963, d = 0.01, 95% CI [-.86, .83], BF = .27), following others (e.g., Seron et al., 1992; Spiller & Jansari, 2008; but see also Barnett & Newell, 2008; Chiou et al., 2018; Ward, Ipser, et al., 2018).

Table 1
Frequency of grapheme-colour synaesthetes in aphantasic and non-aphantasic groups (displayed as a percentage of each group in brackets) within the Generation Scotland sample.

	Aphantasic (n=212)	**Non-aphantasic (n=1073)**	**Total (n=1285)**
Synaesthete	2 (0.9%)	12 (1.1%)	14 (1.1%)
Non-synaesthete	210 (99.1%)	1061 (98.9%)	1271 (98.9%)

Discussion

Our two key results were that grapheme-colour synaesthesia can indeed exist within people with aphantasia, and it is no less prevalent in aphantasics versus non-aphantasics. Rates of grapheme-colour synaesthesia in aphantasics (0.9%) and non-aphantasics (1.1%) were statistically equivalent, and reflective of the prevalence of grapheme-colour synaesthesia in the general population (1.1-1.5%; Carmichael et al., 2015; Simner & Carmichael, 2015; Simner, Mulvenna, et al., 2006). Rates of aphantasia were also statistically equivalent across synaesthetes (14.3%) and non-synaesthetes (16.5%), and

again largely reflective of other estimates (e.g., Betts, 1909). Importantly, our overall group-sizes were small. In other words, despite screening over a thousand participants, our groups of synaesthetes with and without aphantasia were 2 and 12 respectively. We therefore sought to replicate our finding in Experiment 1b, with a larger sample of synaesthetes. We also sought to improve on our methodology: in Experiment 1a we evaluated aphantasia using a single question, while in Experiment 1b we will use the full *Vividness of Visual Imagery Questionnaire 2* (*VVIQ-2*; Marks, 1995). This change can allow us to have more confidence in our assessments of aphantasia. In testing larger samples in Experiment 1b, we also asked an additional question: does aphantasia influence the *type* of synaesthesia experienced (projector or associator synaesthesia; see below)?

Experiment 1b: Can synaesthesia and aphantasia co-occur (a replication)? Does aphantasia influence the type of synaesthesia experienced?

In this study we recruited far larger numbers of synaesthetes by examining data from the largest international online destination for synaesthetes: *The Synesthesia Battery* is a standardised collection of tests and questionnaires for assessing synaesthesia and related phenomenology (Eagleman et al., 2007). The battery includes the gold-standard consistency test for objectively diagnosing grapheme-colour synaesthesia cloned in Experiment 1a, as well as questionnaires to assess the type of synaesthesia experienced (e.g., the *Projector-Associator (PA) Questionnaire*; Rouw & Scholte, 2007). The battery also includes a questionnaire to measure visual imagery (*VVIQ-2*; Marks, 1995). Using these three elements of the *Synesthesia Battery*, we aim to replicate our results from Experiment 1a (synaesthesia occurring equally in aphantasics, and non-aphantasics, and vice versa), but also predict that people with aphantasia may tend to have a different variant of synaesthesia (i.e., more associator-like than people without aphantasia).

Methods

Participants

Our participants were 16,246 individuals who had taken the Synesthesia Battery between 2007 to 2018. Participants had a mean age of 29.11 years (*SD* = 11.76), and were predominantly female (77.2%; i.e., 12,539 female; 3482 male; 225 ‘other’). This female

bias does not reflect an underlying epidemiological fact (see Simner & Carmichael, 2015; Simner, Mulvenna, et al., 2006) but likely the well-known finding that women participate in online surveys more often than men (Smith, 2008). Participants navigated to the website (www.synesthete.org) via various means: some located it by internet searches about synaesthesia, and others were directed by a range of synaesthesia researchers, since this testing portal is the most widely used diagnostic tool in the field of synaesthesia.

Materials and procedure

Participants completed three measures: (a) the objective assessment for grapheme-colour synaesthesia, (b) the *Projector-Associator (PA) Questionnaire* (Rouw & Scholte, 2007), designating any synaesthete as either an associator or projector, and (c) a visual imagery questionnaire (*VVIQ-2*; Marks, 1995), which allows us to assess participants for aphantasia. Details of these tests are given below. All participants provided informed consent prior to taking part, and ethical approval for this study came from the *University of Sussex Cross-Schools Science and Technology Ethics Committee*, and the ethics board at *Baylor College of Medicine*.

Grapheme-colour synaesthesia diagnostic. Details of the diagnostic for grapheme-colour synaesthesia are given in Experiment 1a. Unlike Experiment 1a however, which replicated the diagnostic tool and hosted it on a server at the University of Sussex, in the current study the site was hosted at the online international meeting place for synaesthetes: www.synesthete.org.

Projector-Associator Questionnaire (PA questionnaire; Rouw & Scholte, 2007). This questionnaire distinguishes projector synaesthetes from associator synaesthetes. All participants were given this test, irrespective of how they performed in the diagnostic above (but our results will focus on scores from those ultimately diagnosed as synaesthetic). Participants indicated the degree to which they agreed with 10 statements about their synesthetic experiences (on a scale of 1 – 'Strongly disagree', to 5 – 'Strongly agree'). Half the statements captured the experiences of projectors (e.g., "When I look at a certain letter/number, the synesthetic color appears somewhere outside my head (such as on the paper)"), and half captured the experiences of associators (e.g., "When I look at a certain letter/number, the accompanying color appears only in my thoughts and not somewhere outside my head (such as on the paper)"). In line with standard procedure

(Eagleman et al., 2007; Rouw & Scholte, 2007), an overall PA score was generated by subtracting the mean score for the associator questions from the mean score for the projector questions. Associators were classified by scoring below 0, and projectors by scoring above 0. The original questionnaire also contained two additional items which we removed because they did not appear to adequately distinguish projectors from associators ("When I look at a certain letter or number, I see a particular color", and "The color has the same shape as the letter/number").

Vividness of Visual Imagery Questionnaire-2 (VVIQ-2; Marks, 1995). Within this questionnaire, participants were asked to think of a series of eight scenarios (e.g., "A country scene …"). Participants had to rate "the picture that comes before your mind's eye" for four aspects per scenario (e.g., "The contours of the landscape"). Ratings were given on a scale of 1-5 as follows: 1 – 'No image at all, you only "know" that you are thinking of the object'; 2 – 'Vague and dim'; 3 – 'Moderately clear and vivid'; 4 – 'Clear and reasonably vivid'; and 5 – 'Perfectly clear and as vivid as normal vision'. The questionnaire was scored by summing responses to all 32 questions, giving possible scores in the range of 32-160. In line with previous research showing that aphantasic individuals rate their visual imagery as absent or vague/dim (Keogh & Pearson, 2018; Zeman et al., 2015), aphantasics were classified by scoring between 32-64 on the VVIQ-2 (where a score of 32 indicates no imagery at all, and a score of 64 represents rating all items as 'vague and dim').

Results

Can aphantasia and synaesthesia co-occur?

We first divided our sample into those with and without aphantasia so we could compare rates of synaesthesia across each group. Here we classified aphantasics as those scoring ≤64 on the VVIQ-2, again in line with aphantasics rating their imagery is absent, vague or dim within prior literature. Within the total sample (16,246), there were 196 people (*M* age = 28.03, *SD* = 11.49; 145 female, 49 male, 2 other) with a VVIQ-2 ≤64 (*M* score = 52.76, *SD* = 10.36), giving an overall aphantasia prevalence of 1.2%. We next considered the synaesthesia diagnostic, and found 12,589 people with grapheme-colour synaesthesia overall (*M* age = 28.62, *SD* = 11.36; 9844 female, 2572 male, 173 other). There were 144 synaesthetes among our 196 aphantasics (*M* age = 27.24, *SD* = 11.42; 104 females, 38 males, 2 other), and there were 12,445 synaesthetes among our 16,050 non-aphantasics

(*M* age = 28.63, *SD* = 11.36; 9740 females, 2534 males, 171 other). This provides a prevalence for grapheme-colour synaesthesia of 73.5% in aphantasics, and 77.5% in non-aphantasics (see table 2), a difference that was non-significant [χ^2 (1, *N* = 16246) = 1.61, *p* = .204; chi-square test with Yates continuity correction] (NB. the percentage of synaesthetes is high because synaesthetes are a priori attracted to this website, but our values can be meaningfully compared to each other[5]). We calculated a BF of .018 which provides good support for the null hypothesis, suggesting there is no difference in rates of synaesthesia in aphantasics and non-aphantasics.

We next looked at the reverse case: i.e., the prevalence of aphantasia in our synaesthete and non-synaesthete groups. There were 144 aphantasic individuals among our 12,589 synaesthetes (*M* age = 27.24, *SD* = 11.42; 104 females, 38 males, 2 other), and 52 aphantasics among our 3657 non-synaesthetes (*M* age = 30.21, *SD* = 11.52; 41 females, 11 males). This gives an aphantasia prevalence of 1.1% in the synaesthete group, and 1.4% in the non-synaesthete group, with no significant difference [χ^2 (1, *N* = 16246) = 1.61, *p* = .204; chi-square test with Yates continuity correction]. Again, a BF of .018 provides good support for the null hypotheses, providing evidence for there being no difference in rates of aphantasia in synaesthetes and non-synaesthetes.

Our results replicate our finding from Experiment 1a, showing that high imagery is not a pre-requisite of synaesthesia, and that it is perfectly possible to have synaesthesia with little or no imagery at all. As an added measure of conservativeness, we can also look for synaesthesia by splitting aphantasics into those with a small amount of imagery (dim and vague; scoring 64 on the VVIQ-2) versus those with no imagery whatsoever (scoring 32

[5] Both groups also had high imagery, which may be unsurprising since both were self-motivated to navigate to a website to explore the possibility they had synaesthesia (and awareness of synaesthesia grows with imagery; Ward et al., 2018). Control means were yet higher (*M* = 124.44; *SD* = 23.19) than synaesthetes (*M* = 121.83; *SD* = 22.49; *t*(5799.91) = 6.03, *p* <.001), which may be the reason they (falsely) believed they were synaesthetic. We can confirm both are high against a more typical control group, who were given the VVIQ-2 at the online workplace Amazon Turk (n502; *M* age = 36.55, *SD* = 11.53; n208 female, n292 male, n2 other). A one-way ANOVA comparing the three groups revealed a significant effect of group (*F*(2,16745) = 38.98, *p* <.001), with Amazon Turk participants (*M* = 115.84; *SD* = 25.80) reporting significantly lower imagery than both synaesthetes (*t*(531.78) = -5.12, *p* <.001) and our original controls (*t*(617.21) = -7.08, *p* <.001). Hence, while the recruitment methods in our study are appropriate for our own aims (i.e., finding synaesthetes with aphantasia), they are likely unsuitable for the type of groupwise comparisons which have been the focus of other papers (e.g., Barnett & Newell, 2008; Brang & Ahn, 2019; Chiou et al., 2018; Seron et al., 1992; Spiller & Jansari, 2008; Spiller et al., 2015).

on the VVIQ-2). Out of the 21 individuals with 'dim and vague' imagery (scoring 64 on the VVIQ-2), 15 were classified as synaesthetes, and out of the 20 individuals with no imagery at all (scoring 32 on the VVIQ-2), 12 were classified as synaesthetes. There was no difference between these proportions [χ^2 (1, N = 41) = .195, p = .659; chi-square test with Yates continuity correction] showing that synaesthesia occurs to the same degree whether aphantasics have little imagery or none at all. We calculated a BF of .90 which provides anecdotal support for there being no difference in the rates of synaesthesia in aphantasics with weak imagery, compared to absent visual imagery. Importantly, our results show that it is indeed possible to have synaesthesia with no visual imagery whatsoever.

Table 2
Frequency of grapheme-colour synaesthetes in aphantasic and non-aphantasic groups (displayed as a percentage of each group in brackets) within the Synesthesia Battery sample.

	Aphantasic (n=196)	**Non-aphantasic (n=16050)**	**Total (n=16246)**
Synaesthete	144 (73.5%)	12445 (77.5%)	12589 (77.5%)
Non-synaesthete	52 (26.5%)	3,605 (22.5%)	3,657 (22.5%)

Does aphantasia influence the type of synaesthesia experienced? (i.e., are aphantasics more likely to be associators than projectors)?

Next, we examined the PA questionnaire data to categorise synaesthetes as associators or projectors. From the original dataset of 12,589 confirmed synaesthetes, we excluded 476 individuals with incomplete PA data (i.e., 346 individuals who had not attempted the questionnaire, and 130 who had not finished it), and a further 244 individuals with an overall PA score of 0 (i.e., whose status as projector vs. associator was unclear). Out of the remaining 11,869 (n = 130 aphantasics, n = 11,739 non-aphantasics), 1073 were projectors (n = 14 aphantasics, n = 1059 non-aphantasics) and 10,796 were associators (n = 116 aphantasics, n = 10,680 non-aphantasics). As such, 89.2% of aphantasic synaesthetes were associators, and 10.8% were projectors. A similar result was found in non-aphantasics: 91.0% of synaesthetes were associators, and 9.0% were projectors. Indeed, there was no significant difference in the occurrence of associator or projector

synaesthesia across groups [χ^2 (1, N = 11869) = .289, p = .591; chi-square test with Yates continuity correction]. We then calculated a BF to quantify evidence for this null effect. Our calculated BF of .007 provides strong evidence for there being no difference in the occurrence of projector or associator synaesthesia across aphantasic and non-aphantasic synaesthetes.

This analysis initially suggests that people with aphantasia have the same proportions of associator/projector synaesthesia as non-aphantasics. However, this finding required us to make categorical divisions of the PA data to form two groups (associators, projectors). We therefore ask whether a link between aphantasia and associator synaesthesia might emerge by treating scores continuously. This is likely to be a more appropriate treatment of the data since Skelton, Ludwig, and Mohr (2009) have suggested that projector/associator status resides on a continuum, rather than as discrete categorisations. Hence, we investigated whether aphantasic synaesthetes have lower/more negative PA scores (indicating 'stronger' associator traits) than non-aphantasics. As hypothesised, we found that overall PA scores were indeed significantly lower in aphantasic synaesthetes (i.e., more associator-like; M = -2.58, SD = 1.57) than in non-aphantasic synaesthetes (M = -2.20, SD = 1.54; see Figure 2). This difference was significant in an independent samples t-test with Welch correction (t(131.79) = 2.78, p = .006, d = 0.25, 95% CI [.11, .66]).

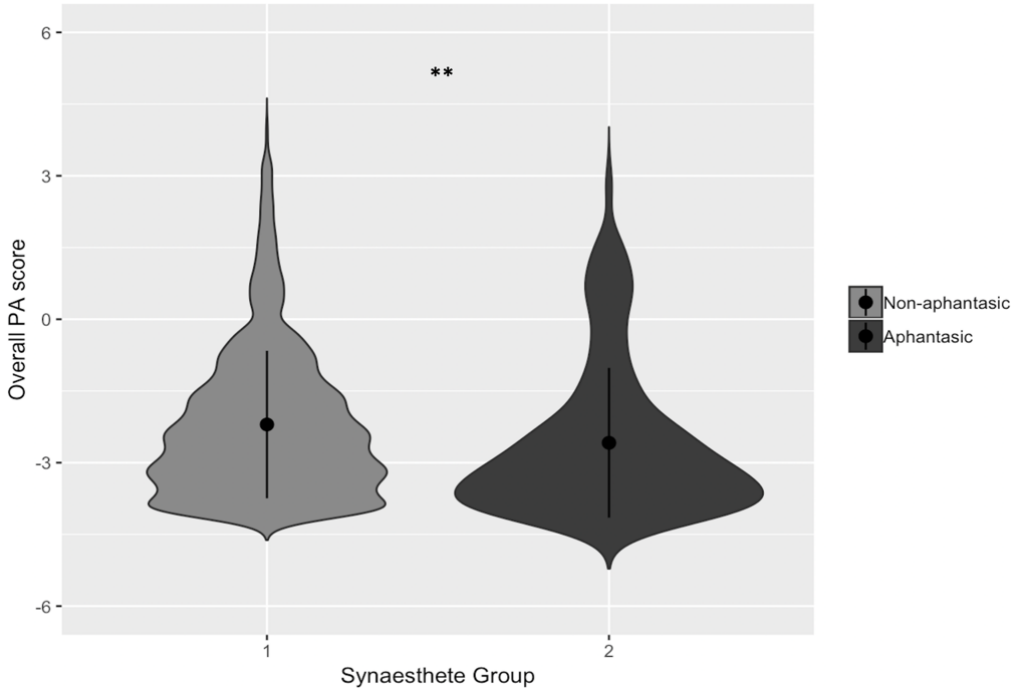

Figure 2. Overall PA questionnaire score distribution and means (with standard deviations) as a function of group (aphantasic synaesthetes vs non-aphantasic synaesthetes) within the *Synesthesia Battery* sample. Lower scores represent more associator-like traits. *Note.* * p <.05, ** p <.01, *** p <.001.

Discussion

In Experiment 1b we have investigated visual imagery using the largest sample of synaesthetes to date. We have replicated our finding from Experiment 1a that synaesthesia can co-occur with aphantasia, thereby showing that visual imagery is not a pre-requisite for synaesthesia to arise. The prevalence of synaesthesia was again equivalent for both aphantasics and non-aphantasics (and equally, rates of aphantasia were equivalent in synaesthetes and non-synaesthetes). As expected, rates of synaesthesia were very high in this sample (73.5% in aphantasics, and 77.5% in non-aphantasics) and far higher than in the general population (i.e., around 1.1-1.5%; Carmichael et al., 2015; Simner & Carmichael, 2015; Simner, Mulvenna, et al., 2006). This is simply because our testing took place at an online destination for synaesthetes (i.e., people who visit the Synesthete Battery are already likely to be synaesthetes, by virtue of their visit). Importantly, however, we can compare the prevalences within our aphantasic and non-aphantasic groups, and in doing so, we replicate our finding from Experiment 1a. Our results also show that even individuals who report the most extreme experience of aphantasia (i.e., no visual imagery whatsoever) can still have synaesthesia, indicating that high imagery (or indeed any imagery) is not a pre-requisite. Nonetheless, we found that

low imagery influences the *type* of synaesthesia experienced: aphantasic individuals reported a synaesthesia that was more associator-like (i.e., more negative scores on the PA questionnaire) compared to non-aphantasics. We return to these issues in our General Discussion.

Experiment 2: What is the relationship between aphantasia and autism?

In Experiment 1 a and b, we looked at two conditions traditionally thought to differ in imagery (synaesthesia as high imagery; aphantasia as low/absent imagery). We now turn our attention to a third neurodevelopmental condition – autism – which has been linked to low visual imagery (related to weak imagination symptomatology). We investigate here whether aphantasia and autism may be linked in some way, given their possible shared deficits in visual imagery.

Methods

Participants

We recruited 118 aphantasics (*M* age = 38.47, *SD* = 14.14) and 118 matched controls (*M* = 37.87, *SD* = 15.22). Groups were matched for age ($t(234) = -0.31$, $p = .757$) and gender (aphantasic group: 69 females, 49 males; controls: 66 females, 51 males, 1 other). Aphantasic participants were recruited from two sources: 102 were recruited from the University of Sussex's *Imagery Lab - Aphantasia Cohort* (*M* age = 39.94, *SD* = 14.27; 57 females, 45 males), and an additional 16 aphantasics were recruited from the University of Exeter *Eye's mind* database (*M* age = 29.06, *SD* = 8.88; 12 females, 4 males). Aphantasics were classified by their scores on the 16-item *VVIQ* (Marks, 1973), a shorter version of the *VVIQ-2* (Marks, 1995). In line with previous research showing that aphantasic individuals rate their visual imagery as absent or vague/dim (Keogh & Pearson, 2018; Zeman et al., 2015), aphantasics (*M* =18.50, *SD* = 3.65) were classified by scoring between 16-32 on the VVIQ. The score range for classifying aphantasia is lower in the VVIQ (16-32) compared to the VVIQ-2 (32-64) simply because the former contains fewer items.

Control participants were recruited from social media, word-of-mouth, Amazon's MTurk, and the participant recruitment system at the University of Sussex. Controls were

confirmed as non-aphantasic again using the VVIQ (scores >32, $M = 58.49$; $SD = 13.07$). As a compensation for taking part, MTurk participants were compensated $2 for our 15 minute test, undergraduate students received course credits and non-students were entered for a prize draw of £25.

Materials and procedure

Our aphantasic participants had already completed the VVIQ prior to participating in the present study, as part of their entering our participant databases. We presented the AQ (and VVIQ for controls) using our online in-house testing platform (www.syntoolkit.org), and participants completed the study from their own homes. All participants provided informed consent prior to taking part, and ethical approval was provided by the University of Sussex *Cross-Schools Science and Technology Ethics Board*.

Autism Spectrum Quotient (AQ; Baron-Cohen et al., 2001). The AQ measures traits associated with autism. Participants rate how much they agree with a series of 50 statements on a scale of 1 (‘Definitely agree’) to 4 (‘Definitely disagree’). The questions are divided equally into five subscales measuring different aspects of autism symptomology: communication (e.g. “I frequently find that I don’t know how to keep a conversation going”), imagination (e.g. “When I’m reading a story, I can easily imagine what the characters might look like”; reversed scored), social skills (e.g. “I find it hard to make new friends”), attention switching (e.g. “I prefer to do things the same way over and over again”) and attention-to-detail (e.g. “I tend to notice details that others do not”). Responses are scored as 0 or 1, where 1 is allocated to responses of “definitely” or “slightly” for behaviours associated with autism (good attention-to-detail, but poor communication, imagination, social skills, and attention switching). Approximately half of the items are reversed scored. In line with standardised scoring (Baron-Cohen et al., 2001), scores equal or greater than 32 indicate the possible presence of autism.

Results

Firstly, we removed an item from the imagination subscale of the AQ which asks directly about visual imagery ability (question 3: ‘If I try to imagine something, I find it very easy to create a picture in my mind’). Since aphantasics would, by definition, score low on this question due to their lack of visual imagery, it was necessary to remove this question to

avoid circularity[6]. We analysed overall AQ scores for aphantasics and controls using an independent samples t-test with Welch correction. Aphantasics had higher overall AQ scores (M = 23.84; SD = 8.57) compared to controls (M = 20.51, SD = 5.98), and this difference was significant (t(209.07) = -3.46, p = .001, d = 0.45, 95% CI [-5.23, -1.44]). Hence, aphantasic individuals reported significantly more traits associated with autism than controls (see Figure 3).

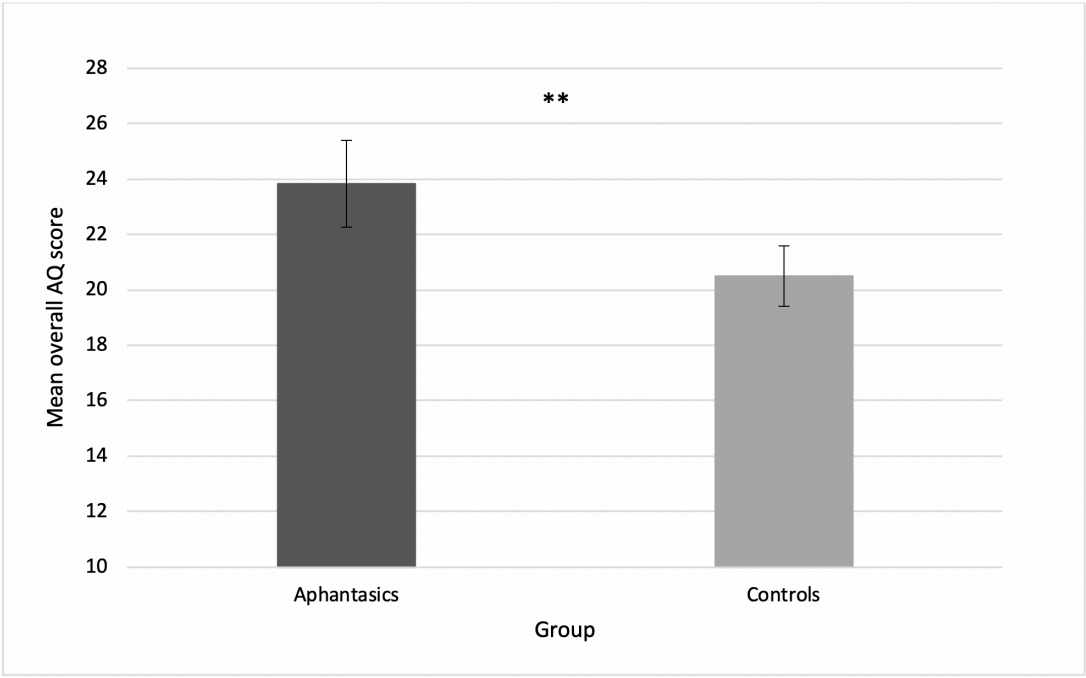

Figure 3. Mean overall AQ scores (with 95% confidence intervals) as a function of group (aphantasics, controls). *Note.* * p <.05, ** p <.01, *** p <.001.

We then conducted a 5 x 2 ANOVA broken down by subscale (imagination, social skills, communication, attention-to-detail, attention switching) and group (aphantasic, control). There was a main effect of group (F(1, 234) = 11.95, p = .001, η_p^2 = .049), and subscale (F(3.34, 782.54) = 37.15, p <.001, η_p^2 = .137; with Greenhouse-Geisser correction), and an interaction between them (F(3.34, 782.54) = 8.56, p <.001, η_p^2 = .035; with Greenhouse-Geisser correction). To investigate our results further, we conducted a series of planned independent samples t-tests with Welch correction for the AQ sub-scales, and adjusted our p values for (n5) multiple comparisons using the Bonferroni method. After correction, aphantasics scored significantly higher on the imagination subscale

[6] We note that the overall pattern of results remained the same overall when including question 3 of the AQ.

(indicating 'less imaginative'; $M = 4.55$, $SD = 2.05$) compared to controls ($M = 2.87$, $SD = 1.71$; $t(226.63) = -6.83$, $p_{corrected} < .001$, $d = 0.89$, 95% CI [-2.16, -1.19]). Aphantasics also scored significantly higher on the social skill subscale (indicating poorer social skills; $M = 4.96$, $SD = 2.97$) compared to controls ($M = 3.85$, $SD = 2.56$), ($t(228.93) = -3.08$, $p_{corrected} = .01$, $d = 0.40$, 95% CI [-1.82, -.40]). There were no significant differences on the remaining three subscales: for communication (aphantasics: $M = 3.90$, $SD = 2.53$; controls: $M = 3.31$, $SD = 1.98$; $t(221.43) = -1.98$, $p_{corrected} = .245$, $d = 0.26$, 95% CI [-1.17, -.002]), attention-to-detail (aphantasics: $M = 4.85$, $SD = 2.34$; controls: $M = 5.14$, $SD = 2.32$), ($t(233.98) = 0.98$, $p_{corrected} = 1.65$, $d = 0.12$, 95% CI [-.30, .89]), and attention switching (aphantasics: $M = 5.56$, $SD = 2.47$; controls: $M = 5.32$, $SD = 2.05$), ($t(226.47) = -0.80$, $p_{corrected} = 2.11$, $d = 0.11$, 95% CI [-.82, .34]). Our data are illustrated in Figure 4.

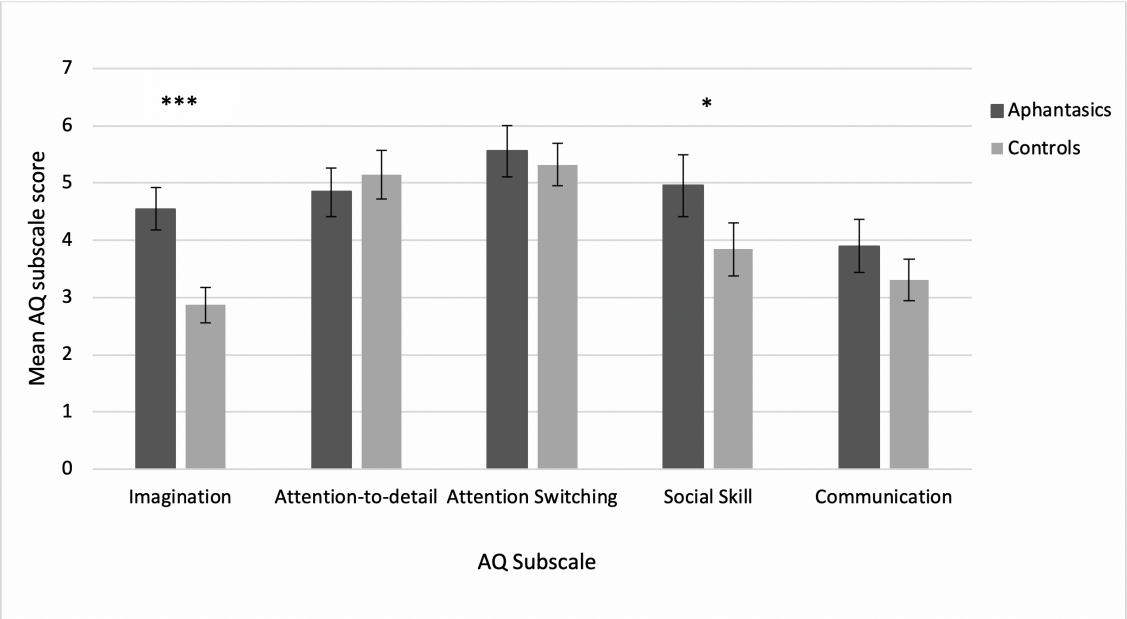

Figure 4. Mean AQ subscale scores (with 95% confidence intervals) as a function of group (aphantasics, controls). Higher scores indicate more autistic-like responses (e.g., poorer imagination and social skills).
Note. Corrected *p* values shown at * $p < .05$, ** $p < .01$, *** $p < .001$.

Finally, since AQ scores ≥32 indicate the possible presence of autism (Baron-Cohen et al., 2001), we also investigated whether there were more people with scores in this range in the aphantasic group, compared to our control group. There were 24 people in our aphantasic group with an overall AQ score ≥32, and 4 people in our control group (again, overall AQ scores excluded question 3 for reasons stated above). This difference was

highly significant [χ^2 (1, N = 236) = 14.63, p <.001; chi-square test with Yates continuity correction]. A bayes factor of 802.43 provides very strong evidence for the alternative hypotheses, i.e., a difference between groups in the number categorised as ≥32 AQ. Although we cannot draw conclusions about clinical diagnoses of autism in our sample, our results show that aphantasic people were far more likely to report levels of autism traits suggestive of an autism spectrum condition compared to controls.

Discussion

In Experiment 2 we have shown that people with aphantasia report higher AQ scores (more traits associated with autism than controls), and fall more often within the range suggestive of autism (≥32). When examining subscales, we found AQ differences within the social skills subscale, and also within the imagination subscale (even when removing a confounding question that asks directly about aphantasia phenomenology). This shows that weak imagination symptomatology associated with autism may also be characteristic of aphantasia. Overall then, our results from Experiment 2 demonstrate a link between aphantasia and autism traits, not only in the expected imagination subscale, but also more broadly.

General Discussion

We investigated the relationship between aphantasia, and two other neurodevelopmental conditions: synaesthesia, and autism. In Experiment 1a we asked whether aphantasia and synaesthesia can co-occur, an important question given the conflicting visual imagery phenomenology assumed in aphantasia (low imagery; Zeman et al., 2015) and synaesthesia (high imagery; e.g., Spiller et al., 2015; but see Seron et al., 1992; Ward, Ipser, et al., 2018). We found that rates of synaesthesia were equal across aphantasics and non-aphantasics, showing that synaesthesia can indeed occur within aphantasic individuals. In Experiment 1b we replicated this with a larger sample, and a more robust measure of aphantasia. Here we also found that synaesthesia can arise even in the most extreme aphantasia cases where there is a total absence of visual imagery (at least in terms of self-reported imagery on the VVIQ). These results reinforce the fact that visual imagery is not necessary for the trait of synaesthesia to arise.

Our findings conflict with previous research suggesting that synaesthesia involves heightened visual imagery compared to controls (e.g., Barnett & Newell, 2008; Chiou et al., 2018; Chun & Hupé, 2016; Meier & Rothen, 2013; Price, 2009; Spiller et al., 2015) and we point to a hypothesis offered by Simner (2013). Simner has suggested that synaesthesia studies may sometimes exaggerate imagery differences in synaesthetes because they often test self-referred synaesthetes (and perhaps those with the most vivid phenomenological experience are more likely to self-refer). To our knowledge, our sample size is the largest ever used to examine imagery in synaesthetes and this has allowed us to capture (and count) instances of low imagery synaesthetes who might otherwise be less obvious within smaller samples. It is interesting to note that we found aphantasic synaesthetes not only when screening a general population sample (Experiment 1a), but even when examining rather self-aware synaesthetes, who had made a self-motivated effort to navigate to an online testing site for synaesthesia (Experiment 1b). Hence, while higher imagery might make synaesthetes more self-aware of their synaesthesia (Ward, Ipser, et al., 2018), a complete absence of imagery does not preclude this.

In the present study we did, however, find that aphantasia influenced the *way* synaesthesia was experienced: aphantasic synaesthetes had stronger associator traits than synaesthetes without aphantasia. Importantly, associator synaesthesia not only encompasses colours in the 'mind's eye' (potentially linked to imagery, but a phraseology used by aphantasics too) but also includes simply *knowing* what colours must be. This latter seems particularly compatible with the notion of aphantasia and may explain how aphantasics are just as likely as the general population to experience synaesthesia, albeit with associator-like traits. Our behavioural results mirror neuropsychological evidence from synaesthetes and aphantasics, who both appear to rely less on sensory cortices: low imagers show lower activation in visual cortex during a mental imagery task (compared to high imagers; Logie et al., 2011), while associator synaesthetes have lower grey matter volume (compared to projector synaesthetes; Rouw & Scholte, 2010). It therefore appears that aphantasic individuals experience a synaesthesia characterised less by 'perceiving' synaesthetic colours, and more by an intrinsic awareness or 'knowing' of colours.

Although we found more associator-traits in aphantasic synaesthetes (vs. non-aphantasic synaesthetes) when treating scores on a continuum, we found no categorical effects: in

both groups there were approximately 90% associators and 10% projectors. The projectors with aphantasia are initially puzzling, because projectors experience synaesthetic colours like real-world percepts outside their body, and this has often been interpreted as very high imagery. One reason may simply be measurement error, but an alternative is that aphantasics might report projected synaesthesia in same way they report other visual knowledge – as a metaphor. For example, aphantasics could easily imagine someone standing opposite them, including the colour of their hair. Aphantasics would be likely to report that the hair–colour is 'out there in space' even though they have no visual imagery to form an iconic representation. Likewise, aphantasics may interpret projector items in the *PA Questionnaire* in a similar way. Put differently, a core feature of projector synaesthesia is the spatial location of the colour, and spatial relations are intact in aphantasics (Bainbridge et al., 2021), and may even be stronger than usual (Keogh & Pearson, 2018). Intact spatial imagery could give aphantasics the ability to describe synaesthetic associations 'in space', despite being unable to image colours iconically. This opens a new and interesting debate about whether projectors should be defined by their percept-like experiences (possibly linked to strong imagery) or their external reference frames (possibly without imagery at all). We leave this debate open to explore in future work (see also Ward, Li, et al., 2006).

In Experiment 2 we used a standardised measure of autism traits (*Autism Spectrum Quotient; AQ*; Baron-Cohen et al., 2001) and found that aphantasics scored higher than controls from the general population. We also found significantly more aphantasics (vs controls) over the threshold of 32 – suggestive of clinical levels of autism (Baron-Cohen et al., 2001). Previous research had conflicted on whether autism speaks to visual imagery. Whereas some have suggested low imagery in some autistics (i.e. 'word-fact' thinking styles; Grandin, 2009), others found no difference (e.g., Hughes et al., 2018; using the '*imagery ability*' subscale of the *SCSQ*), while others still suggested imagery may be stronger (Esposito et al., 2016; Kana et al., 2006; Kunda & Goel, 2008, 2011; Soulières et al., 2011). Our own study shows clearly that there is a link between poor imagery (aphantasia) and higher levels of autism traits. When looking within the subscales, we found a difference firstly for 'imagination'. Poor imagination in people with low imagery suggests these concepts may be related – but we stress that they are *not identical* (something high imagers may find counter-intuitive). Our results show that even people with an *entire absence* of imagery can still report imagination, and some

aphantasics reported very strong imagination indeed. To be clear, our finding shows that – as a group – aphantasics have a slight deficit in imagination, but this is not true for all, and imagination can nonetheless operate at a functional level even when imagery is absent. However, our group difference suggests that mental imagery may aid in the process of imagination, perhaps in the construction and maintenance of imagined scenarios. This suggestion has neurological parallels since areas involved in visual imagery (e.g., precuneus; Fletcher et al., 1995; Fulford et al., 2018) are activated when members of the general population are asked to imagine past and future events (Addis, Wong, & Schacter, 2007; Hassabis, Kumaran, & Maguire, 2007; Okuda et al., 2003; Schacter & Addis, 2007), and especially future events which arguably require more imagination (Szpunar, Watson, & McDermott, 2007). But similar studies have yet to be conducted on those without imagery, leaving it unclear whether imagining and imagery necessarily overlap by definition, or whether they simply overlap in most people. Nonetheless, our data show that low imagery and poor imagination are linked, and suggest, perhaps, that lacking visual imagery may drive – at least in part – some of the impairments in imagination that are widely seen in autism (e.g., Baron-Cohen, 1987; Davis et al., 2018; Jarrold et al., 1996; Ten Eycke & Müller, 2014).

We also found that aphantasics scored higher in the autistic subscale of social skills (i.e., had poorer social skills). This may rest on notions of theory of mind and perspective taking. Problems in 'theory of mind' (i.e., the ability to understand another's mental state) have long been a defining feature of autism (Baron-Cohen, 2000; Baron-Cohen, Leslie, & Frith, 1985; Frith, 2001) and correlate with autistic deficits in social skills (Dawson & Fernald, 1987; Mazza et al., 2017; Perner, Frith, Leslie, & Leekam, 1989). According to a dominant view, theory of mind relies on perspective taking, which involves being able to mentally transform a representation of one's body from a first-person (ego-centric) point of view, to a third-person (hetero-centric) perspective (Decety & Grèzes, 2006; Vogeley & Fink, 2003). People with autism or those high in autistic traits often have impairments in this type of mental body transformation as well as in "embodiment" more broadly (simulating the mental and physical state of another within one's own body) (Conson et al., 2015; Gauthier et al., 2018; Kessler & Wang, 2012; Pearson, Marsh, Hamilton, & Ropar, 2014). Here, we suggest that visual imagery may be involved in this process of mental body transformation, helping individuals to take on the perspective of others. Low visual imagery may therefore give rise to difficulties in social skills by

influencing how well an individual can perspective-take, and thereby understand the mental state of others.

One interesting consideration is how people high in autism traits are more likely to have aphantasia (as suggested here), but also synaesthesia (Baron-Cohen et al., 2013; Neufeld et al., 2013). Prior to our study this may have been a confusing finding, since autism and synaesthesia were assumed to have polar imagery requirements (low and high respectively). We have now shown, however, that having aphantasia does not preclude synaesthesia at all – and the slightly elevated rates of synaesthesia in autism may themselves be further evidence for the fact that high imagery is not a pre-requisite for synaesthesia to arise. Importantly, however, we note that the traits linking autism and aphantasia on the one hand (subscales of imagination and social skills), and autism and synaesthesia on the other (the subscale of attention to detail; Ward et al., 2017) seem to be very different.

One possible limitation of our study is that control participants were recruited differently to aphantasics in Experiment 2. Our recruitment of controls involved, at times, what might be considered a more 'social' approach (i.e., 5% of controls were recruited via word of mouth), which may have lowered the mean AQ score for social skills in our control group (indicating 'better' social skills). However, most controls (81%) were recruited via online methods, which were either similar to methods for recruiting aphantasics (e.g., online groups/forums), or targeted a forum where rates of autism are known to be higher than the general population, not lower (i.e., 49% came from MTurk, which has been shown to have twice the rates of autism compared to real-world community samples; Chandler & Shapiro, 2016). The remaining controls (14%) were recruited via the participant recruitment system at the University of Sussex. As such, we suggest that our recruitment likely did not differentially influence AQ scores across aphantasics and controls. A second limitation is that we did not investigate clinical cases of autism, but looked at AQ trait profiles. This was an important step, given limited previous research investigating the relationship between autism and aphantasia, and the novelty of our research question. Our results provide – for the first time – evidence for aphantasia and autism being linked in some way, and future research may continue to investigate this association further.

A third limitation is that we diagnosed aphantasia using self-report measures. Although we used just a single question to identify aphantasia in Experiment 1a, we replicated our results using a multi-item questionnaire in Experiment 1b (*VVIQ-2*; Marks, 1995). The VVIQ (Marks, 1973, 1995) is the current gold standard measure of aphantasia, used widely in contemporary aphantasia research (e.g., Bainbridge et al., 2021; Dawes et al., 2020; Jacobs et al., 2018; Keogh & Pearson, 2018; Zeman et al., 2015, 2020). Indeed, VVIQ scores map consistently onto behavioural measures of visual imagery (Pearson, 2014). In their imagery binocular rivalry task (see Introduction), Pearson and colleagues show that scores on the VVIQ correlate positively with the degree of colour dominance (Pearson et al., 2008, 2011; Rademaker & Pearson, 2012), suggesting that people generally have good metacognition about their visual imagery abilities. Taken together, this shows that the VVIQ is a robust measure for classifying aphantasia.

One final limitation is our focus on one variant of synaesthesia only (grapheme-colour synaesthesia). However, we imagine a similar pattern of results for other visual synaesthesias (e.g., *taste-to-colour*, *sound-to-colour*; Downey, 1911; Ward, Huckstep, et al., 2006). And since people with aphantasia often experience imagery deficits in other sense modalities (Dance, Ward, et al., 2021; Dawes et al., 2020), we may yet find that people without taste imagery, for example, can nonetheless experience taste synaesthesias (i.e., lexical-gustatory synaesthesia, where words trigger tastes; Ipser, Ward, & Simner, 2020; Ward & Simner, 2003). Future research asking questions such as these will further our understanding of the relationship between aphantasia and synaesthesia.

In conclusion, our results from the present study begin to characterise the relationship between aphantasia, synaesthesia and autism. We have shown that synaesthesia occurs at equal rates in aphantasia as in the general population, but that aphantasic synaesthetes show stronger associator (than projector) traits compared to 'phantasic' synaesthetes. This suggests that aphantasics experience synaesthesia phenomenologically differently, but that visual imagery is not necessary for synaesthesia to develop. We also found that it was possible for aphantasics to experience projector synaesthesia – a variant usually thought to require strong visual imagery. We have assumed this means that aphantasic projectors have an external spatial reference frame for their experiences, even if those experiences are not percept-like. Finally, we showed that people with aphantasia are higher in traits associated with autism, especially within the imagination and social skills

subscales. This suggests that absent or weak visual imagery may play a part in driving weaknesses in these domains within autism spectrum conditions. In sum, our study serves to enhance our understanding of aphantasia, in showing its relationship to synaesthesia, and to autism traits.

Chapter 4

What is the link between mental imagery and sensory sensitivity? Insights from Aphantasia

Chapter Summary

In the previous chapter, I examined how aphantasia intersects with synaesthesia and autism, two other neurodevelopmental conditions linked to sensory/imagery differences. In Chapter 4, I examine the broader sensory profile of aphantasia in more detail. I assess whether the visual imagery weakness seen in aphantasia extends to other (non-visual) sense domains, and question for the very first time whether levels of mental imagery predict levels of sensory sensitivity (i.e., a clinically relevant trait linked to over- or under-responsivity to incoming sensory stimulation; Robertson & Simmons, 2013, 2015). I also present a novel model linking imagery and sensory sensitivity via levels of cortical excitability in the brain. Chapter 4 therefore seeks to enhance our understanding of the sensory trait profile of aphantasia, and the link between mental imagery and sensory sensitivity more broadly.

Full citation

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.

Abstract

People with *aphantasia* have impoverished visual imagery so struggle to form mental pictures in the mind's eye. By testing people with and without aphantasia, we investigate the relationship between sensory imagery and sensory sensitivity (i.e., hyper- or hypo-reactivity to incoming signals through the sense organs). In Experiment 1 we first show that people with aphantasia report impaired imagery across *multiple* domains (e.g., olfactory, gustatory etc.) rather than simply vision. Importantly, we also show that imagery is related to sensory sensitivity: aphantasics reported not only lower imagery, but also lower sensory sensitivity. In Experiment 2, we showed a similar relationship between imagery and sensitivity in the general population. Finally, in Experiment 3 we found behavioural corroboration in a Pattern Glare Task, in which aphantasics experienced less visual discomfort and fewer visual distortions typically associated with sensory sensitivity. Our results suggest for the very first time that sensory imagery and sensory sensitivity are related, and that aphantasics are characterised by both lower imagery, and lower sensitivity. Our results also suggest that aphantasia (absence of *visual* imagery) may be more accurately defined as a subtype of a broader imagery deficit we name *dysikonesia*, in which weak or absent imagery occurs across multiple senses.

Introduction

Mental imagery is the mechanism by which we mentally simulate perceptual experiences - from visualising a friend's face in the mind's eye, to hearing our favourite song in the 'mind's ear' (and similarly for our other senses). Mental images themselves are "iconic" in that visual images in some way resemble internal pictures, and auditory images in some way resemble internal sounds. Being able to form mental images is an essential part of life for many people but this capability varies from person to person. For some people, visual mental imagery is exceptionally strong and nearly as vivid as real-life perception, while for others it is virtually or completely absent, a condition known as *aphantasia* (Zeman et al., 2015, 2016, 2020). One question we ask in this paper is whether the imagery deficit that characterises aphantasia in the visual domain is apparent too in other sense domains. For example, we ask whether people with aphantasia also have impaired imagery for taste, smell, touch and so on. But as well as differing on how sensory information is imaged, people also vary on how sensory information makes them *feel*. Some people have a comfortable tolerance for incoming sensory stimuli from the outside world, while others have *sensory sensitivities* (i.e., an under- or over- responsiveness to sounds, smells, tastes etc., see below; e.g., Robertson & Simmons, 2013). Here, we also ask for the very first time whether sensory sensitivities and sensory imagery are related. Using self-report and behavioural methods we will show that aphantasic individuals experience not only lower imagery, but also lower sensory sensitivity. We present three experiments testing these ideas but begin with brief overviews of aphantasia, imagery and sensory sensitivity.

People with aphantasia report either a complete lack of visual mental imagery, or imagery that is only dim, vague or fleeting (e.g., Keogh & Pearson, 2018; Zeman et al., 2015). Aphantasia has until now been defined exclusively as an absence of *visual* imagery (Milton et al., 2021; Zeman et al., 2015, 2016, 2020) but of course mental imagery itself can encompass other modalities too, including auditory, olfactory, gustatory, tactile, motor, and bodily imagery. For us to investigate the nature of aphantasia (and subsequently, to investigate how imagery and sensation are themselves linked) an important initial question is whether the imagery impairment seen in aphantasia also extends to other sense domains. There are a number of reasons to think this might be the case. 'Aphantasia-like' deficits in non-visual imagery do indeed exist: there are vast individual differences in the vividness of auditory imagery (Berger & Ehrsson, 2017),

olfactory imagery (Leclerc et al., 2019), gustatory imagery (Tiggemann & Kemps, 2005), tactile imagery (Belardinelli et al., 2009), motor imagery (Isaac & Marks, 1994), and bodily imagery (Andrade, May, Deeprose, Baugh, & Ganis, 2014), with some individuals reporting imagery almost as vivid as real-life perception, while others report a total absence of imagery. Although poor imagery in one domain does not *necessarily* preclude high imagery in another (Andrade et al., 2014), positive correlations in imagery strength do exist across different sensory modalities (Andrade et al., 2014; Leclerc et al., 2019; Lima et al., 2015; White, Ashton, & Law, 1974). This suggests that people with aphantasia may indeed have weak/absent imagery in other non-visual domains.

Other suggestions come from first-hand reports from writers with aphantasia (e.g., Watkins, 2018) and from one valuable recent survey of 2000 aphantasic people, of whom 54% suggested they might have weak/absent imagery in *all* sensory modalities (Zeman et al., 2020). However, this was elicited by a single question ('*Are all types of imagery affected, or can you imagine sounds (including music), textures (by imagined touch), tastes or smells?'*; Zeman et al., 2020), which required participants to self-diagnose without other means to determine whether their imagery (e.g., gustatory imagery) was better or worse than the average person. A recent study, too, by Dawes et al. (2020) found that aphantasics reported significantly weaker mental imagery than controls in the auditory, tactile, kinaesthetic, taste, olfactory and bodily subscales of Sheehan's adaption of *Bett's Questionnaire Upon Mental Imagery* (Sheehan, 1967). Although their findings represent an important step in understanding the phenomenology of aphantasia, their questionnaire has been criticised for having a small number of items per modality, and for including items that are unclear or out-dated (e.g., imagining 'the whistle of a locomotive'; see e.g., Pecher, van Dantzig, & Schifferstien, 2009; Spiller et al., 2015). Here, we therefore address this same question, measuring aphantasics' mental imagery across multiple sense domains, but using a series of contemporary standardised questionnaires. These measures may indeed show that the phenomenology of aphantasia extends across multiple senses. Such a finding would be important because it would indicate that the current literature-wide focus on aphantasia as a visual deficit may be hiding a wider phenomenology.

A final reason to suspect that aphantasics may have low or absent imagery in non-visual domains comes from considering its possible neural basis. Visual imagery has been linked

to levels of excitability in the visual cortex, meaning that aphantasics may potentially have excitability differences (Cattaneo, Pisoni, Papagno, & Silvanto, 2011; Sparing et al., 2002, but see Keogh et al., 2020). If this is true, then these differences in excitability might potentially be found in other sensory regions within the same brain. One way to understand the link between imagery and excitability is to look at studies of the visual cortex using transcranial magnetic stimulation (TMS). Using TMS can induce visual effects known as phosphenes. Importantly, engaging in visual imagery lowers phosphene thresholds (Cattaneo et al., 2011; Sparing et al., 2002), where low thresholds represent high excitability (i.e., less stimulation to generate a phosphene; e.g., Stewart, Walsh, & Rothwell, 2001). This suggests that since aphantasics have poor imagery, they may therefore have low excitability. Although phosphene thresholds have not been measured in aphantasia, supporting evidence is that *high* imagers have *high* visual excitability. For example, people with the sensory condition *synaesthesia* tend to score highly on mental imagery questionnaires (e.g., Barnett & Newell, 2008; Price, 2009; Spiller et al., 2015, but see Dance, Jaquiery, et al., 2021; Simner, 2013), while also showing high visual cortex excitability compared to controls (Terhune, Murray, et al., 2015; Terhune, Song, & Cohen Kadosh, 2015; Terhune, Tai, Cowey, Popescu, & Cohen Kadosh, 2011). Important to our purposes here, cortical excitability has also been implicated in other imagery domains too (e.g., engaging in motor imagery reduces motor thresholds; Fadiga et al., 1999; Hashimoto & Rothwell, 1999; Stinear, Byblow, Steyvers, Levin, & Swinnen, 2006). We propose, therefore, that people with aphantasia may have lower excitability in the visual cortex, and that this reduced excitability may extend to other sensory regions. If so, people with aphantasia may typically have poor imagery in more than one sense.

The role of cortical excitability in imagery also raises a second important question. Visual cortex excitability has also been linked to a second phenomenon, *sensory sensitivity* (e.g., Green et al., 2013; Wilkins et al., 1984). Sensory sensitivity is a disturbance in the way individuals react to incoming sensory stimuli from the outside world. For example, someone who is visually hyper-sensitive might find lights too glaring, and avoid bright environments (e.g., supermarkets). They might also experience greater discomfort and visual distortions (e.g., shimmering, and flashes) in the *Pattern Glare Ta*sk, in which participants are shown gratings (parallel lines) at particular spatial frequencies that trigger visual sensitivities (Braithwaite, Broglia, Bagshaw, & Wilkins, 2013; Evans & Stevenson, 2008; Ward et al., 2017; Wilkins et al., 1984). Visual discomfort and

distortions themselves are a form of 'visual stress' ('*pattern glare*'), indicative of visual sensitivity (Braithwaite, Broglia, Bagshaw, et al., 2013; Braithwaite, Broglia, Brincat, et al., 2013; Ward et al., 2017; Wilkins et al., 1984). Important to our purposes here is that sensory sensitivities – like imagery – appear to relate to cortical excitability (e.g., Green et al., 2013; Wilkins et al., 1984). Hyper-excitability in the cortex (i.e., lower phosphene thresholds; e.g., Stewart et al., 2001) has been found in sensory-sensitive populations (e.g., people with migraine) and is associated with heightened pattern glare in these individuals (Aurora, Cao, Bowyer, & Welch, 1999; Aurora, Welch, & Al-Sayed, 2003; Aurora & Wilkinson, 2007; Brigo et al., 2012; Chadaide et al., 2007; Coutts, Cooper, Elwell, & Wilkins, 2012; Huang, Cooper, Satana, Kaufman, & Cao, 2003; Huang et al., 2011; Mulleners, Chronicle, Palmer, Koehler, & Vredeveld, 2001; Wilkins et al., 1984; Young et al., 2004). Other research, too, shows that special populations who tend to be high in imagery (synaesthetes, noted above for having high visual cortex excitability; Terhune et al., 2015, 2011) report high levels of sensory sensitivity (Van Leeuwen, Van Petersen, Burghoorn, Dingemanse, & Van Lier, 2019; Ward, Brown, Sherwood, & Simner, 2018; Ward et al., 2017), and increased susceptibility to pattern glare (Ward et al., 2017). Taken together, this research raises the possibility that excitation within the visual cortex may give rise to both imagery differences and sensitivity differences within the same individuals. Here, we draw an explicit link between these findings, hypothesising for the first time that imagery and sensitivity are linked. Specifically, we predict that individuals with *low/absent* imagery (aphantasics) might report lower sensory sensitivity than the average person.

So far, our predictions have linked imagery and sensory sensitivity, but we have implicitly focussed on *hyper*-sensitivity (e.g., high/low excitability in high/low imagery linked to high/low *hyper*-sensitivity). But properly speaking, sensory sensitivity encompasses both hyper- and hypo- sensitivities (Robertson & Simmons, 2013). A person who is hyper-sensitive (in the visual domain for example) may find bright lights too glaring and seek to avoid them, while a person who is hypo-sensitive may experience low responsivity ('sensory dampening') and actively *seek* visual stimulation (e.g., flick their fingers in front of the eyes; Ben-Sasson et al., 2009; Bogdashina, 2003; Simmons et al., 2009). Paradoxically, hyper- and hypo-sensitivities are often found within the same individual, either across sense modalities (e.g., avoiding bright lights, but seeking odours) or within modalities (e.g., disliking loud noises, but playing the same song over and over;

Robertson & Simmons, 2013; Smees, Rinaldi, Simmons, & Simner, 2020; Ward et al., 2017). Above, we hypothesised a link between imagery and hyper-sensitivity, but are there links to both hyper- *and* hypo-sensitivity? In fact, we predict lower sensory sensitivities in aphantasics in both hyper- and hypo-domains, because these domains have themselves been linked through mechanisms of *adaptation* (Takarae & Sweeney, 2017; Ward, 2019a).

Adaptation is when we stop noticing the smell of someone's perfume, for example, after having been in their company for a while (Dalton, 2000; Takarae & Sweeney, 2017; Ward, 2019a). Normal adaptation is driven by a reduction in neural activity in response to continuous or repeated sensory input (Takarae & Sweeney, 2017; Thompson & Spencer, 1966). Failures in adaptation are linked to high levels of cortical excitation and therefore hyper-sensory sensitivity (Green et al., 2015). Specifically, high cortical excitation may prevent neural adaptation, meaning a stimulus remains 'prominent' in attention and becomes overbearing (Takarae & Sweeney, 2017; Ward, 2019a). However, this same failure in adaptation may also give rise to hypo-sensitivity through the prioritising of old stimuli (i.e., the stimuli we are not adapting to). Prioritising old stimuli means that new stimuli are not easily recognised, leading to sensory dampening (i.e., hypo-sensitivity; Takarae & Sweeney, 2017). This fact may explain why individuals who report high levels of hyper-sensitivity often report high levels of hypo-sensitivity (Robertson & Simmons, 2013; Ward et al., 2017), and taken together, all these facts also make predictions about people with aphantasia. If strong imagery is linked to heightened cortical excitability, and heightened cortical excitability is linked to hyper-sensitivity (but also hypo-sensitivity via problems with sensory adaptation), we would expect people with very *low* imagery (aphantasia) to have *lower* levels of hyper-/hypo-sensitivity, when compared to imaging controls.

We summarise our hypothesised link between imagery, sensory sensitivity, and cortical excitation in Figure 1. Our model suggests that aphantasia may be characterised by lower levels of visual cortex hyperexcitability, with this mediating both low imagery *and* low sensitivities to incoming sensory information. Our model makes several predictions. First, people with aphantasia (impoverished visual imagery) may report fewer sensory sensitivities within the visual sense, relative to imaging controls (Experiment 1). Second, we predict that people with aphantasia may in fact have poor imagery across *multiple*

domains (visual, gustatory, olfactory etc; Experiment 1). If so, we predict, thirdly, that people with aphantasia may also therefore self-report lower *sensory sensitivities* across multiple domains (Experiment 1). Thirdly, we predict that the sensory sensitivities reported in Experiment 1 by people with aphantasia should be corroborated by behavioural findings in a subsequent study (Experiment 3): aphantasics should experience fewer visual distortions and less visual discomfort than controls in response to irritable visual gratings in a *Pattern Glare Task*. Finally, if our model holds true in the population at large, then we predict visual imagery and sensory sensitivity to be positively correlated in a general population sample (Experiment 2). In summary, by investigating, for the first time, the relationship between imagery and sensory sensitivity, the present research aims to enhance our understanding of imagery, sensory sensitivity, cortical excitation, and aphantasia.

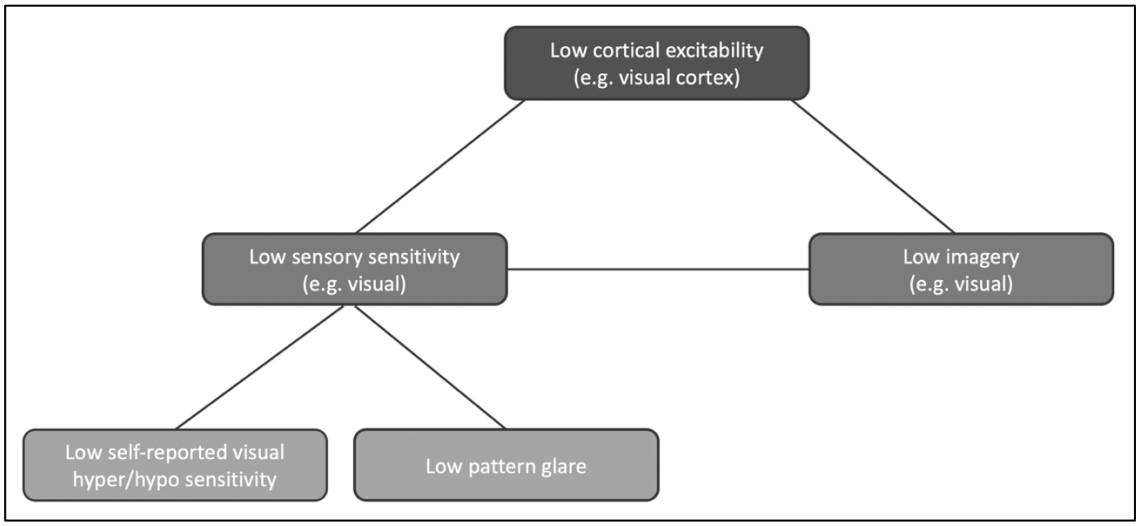

Figure 1. Excitability, Imagery, and its Measurements. Our proposed model links imagery, cortical excitation, and sensory sensitivity. Our model posits that low cortical excitability may be tied to low imagery (e.g., in people with aphantasia) and lower levels of sensory sensitivity. Our model is ambivalent about directionality (i.e., causality) and may be multi-directional. Indeed, levels of imagery (a 'top-down' process) and incoming sensory signals (a 'bottom-up' process) may influence levels of cortical excitability (or vice versa) on a moment-to-moment basis. We have presented our model in terms of *low* excitability to capture the experiences of aphantasics but note that the links proposed are correlational (i.e., it also predicts a link between *high* excitability/ imagery/ sensitivity). Here we have applied our model to *visual* imagery, but suggest it extends to the other sense domains also.

Experiment 1

In this experiment we compared the profile of mental imagery and sensory sensitivity in people with and without aphantasia. We asked whether the imagery deficit reported by people with aphantasia extends across multiple sense domains, and whether this maps on to fewer sensory sensitivities. In our study we were mindful to factor out unwanted influences. Sensory sensitivities are not only variable within the general population (Horder, Wilson, Mendez, & Murphy, 2014; Robertson & Simmons, 2013) but are particularly characteristic of individuals with autism spectrum conditions (henceforth autism) (American Psychiatric Association, 2013) or non-clinical autistic traits (Ben-Sasson et al., 2009; Bogdashina, 2003; Horder et al., 2014; Robertson & Simmons, 2013, 2015; Simmons et al., 2009). Importantly, higher autistic traits are also found in people with aphantasia (Dance, Jaquiery, et al., 2021; Milton et al., 2021). These facts could potentially link imagery, aphantasia, and sensory sensitivities via the medium of autism. Participants in Experiment 1 were therefore screened using the *Autism Spectrum Quotient* (AQ; Baron-Cohen et al., 2001) and AQ scores were added into our analyses of sensory sensitivity as covariates. Note, however, that our predictions stand in the opposite direction to any confound. If people with aphantasia had any predisposition with regards to sensory sensitivities, their higher traits of autism would make this *more* likely. Instead, we are predicting *fewer* sensitivities for people with aphantasia, as an outcome of our model above (Figure 1). Our General Discussion explores this opposition more fully.

Methods

Participants

We recruited 164 aphantasics (101 female, 62 male, 1 other; *M* age = 42.35, *SD* = 15.95) and 138 controls (67 female, 70 male, 1 other; *M* age = 37.39, *SD* = 13.83)[7]. The majority of our aphantasic participants (n=158) were recruited from the University of Sussex's *Imagery Lab - Aphantasia Cohort* while the remaining aphantasic participants (n=6) were recruited from the student body of the University of Sussex. Control participants were recruited from multiple sources including Amazon's Mechanical Turk (MTurk; a

[7] This age difference of ~5 years across groups was significant ($t(299.72) = -2.89$, $p = .004$). However, all our analyses (see Results) remain the same when repeated with age as a covariate (see Appendix A).

platform for collecting quality data; Casler et al., 2013), and via word-of-mouth and social media.

Participants were separated into their two groups using the 'gold standard' questionnaire for aphantasia (*Vividness of Visual Imagery Questionnaire*; *VVIQ*; Marks, 1973, 1995) whose scores range from 16-80 (see below) and where a score between 16-32 is indicative of aphantasia (i.e., imagery is either absent or vague/dim; Keogh & Pearson, 2018; Zeman et al., 2015). The protocol for the VVIQ is given below, but we point out here that people show good metacognition about their own imagery abilities, and self-report measures such as this correlate well with behavioural validations (e.g., Keogh & Pearson, 2018; Pearson et al., 2011). Our aphantasics scored the required 16-32 on the VVIQ (M = 17.49, SD = 3.44), while our non-aphantasic controls scored above 32 (M = 59.67, SD = 11.93). As compensation for taking part, non-students were entered into a prize draw for a shopping voucher, and MTurk participants were given a monetary payment of $4 (for our 25 minute test). For all experiments reported in this paper, participants provided informed consent prior to taking part, and ethical approval came from the *University of Sussex Sciences and Technology Cross-Schools Research Ethics Board*.

Materials and Procedure

Participants completed eight questionnaires online in a random order. (Our *Imagery Lab - Aphantasia Cohort* had already completed the VVIQ, so completed the remaining seven questionnaires.) Six questionnaires measured imagery (summarised in Table 1), while the final two questionnaires measured sensory sensitivities, and autistic traits. These eight measures are described below. Our study was hosted on our in-house testing platform (www.syntoolkit.org).

Table 1

Names and Domains of Imagery Questionnaires, with Abbreviations (Abbr.). Citations for these scales are given in the text below.

Domain	Questionnaire	Abbr.
Visual	Vividness of Visual Imagery Questionnaire	VVIQ
	Plymouth Sensory Imagery Questionnaire	Psi-Q
Auditory	Clarity of Auditory Imagery Scale	CAIS
	Plymouth Sensory Imagery Questionnaire	Psi-Q
Olfactory	Vividness of Olfactory Imagery Questionnaire	VOIQ
	Plymouth Sensory Imagery Questionnaire	Psi-Q
Tactile	Adapted Shortened Betts' Questionnaire Upon Mental Imagery	Betts-ad
Gustatory	Plymouth Sensory Imagery Questionnaire	Psi-Q
Bodily		
Movement	Vividness of Movement Imagery Questionnaire Kinaesthetic Scale	VMIQ
Feeling	Plymouth Sensory Imagery Questionnaire	Psi-Q

Vividness of Visual Imagery Questionnaire (VVIQ; Marks, 1973). The VVIQ was our aphantasia diagnostic. In this questionnaire, participants were asked to imagine a series of four scenarios (e.g., "A beach by the ocean on a warm summer's day") and to rate the strength of their visual imagery for four aspects of each scene (e.g., "The appearance and colour of the water"). Imagery was rated on a five-point scale, comprising 1 ("No image at all, you only "know" that you are thinking of the object "), 2 ("Vague and dim"), 3 ("Moderately clear and vivid"), 4 ("Clear and reasonably vivid"), and 5 ("Perfectly clear and as vivid as normal vision"). Responses were summed to give scores ranging from 16-80. As noted above, responses between 16-32 (no imagery or vague/dim) are indicative of aphantasia.

Clarity of Auditory Imagery Scale (CAIS; Willander & Baraldi, 2010). Here, participants were asked to imagine a series of 16 sounds (e.g., "A dog barking") and to rate how clearly they could 'hear' the sounds on a scale of 1 ("No sound at all, you only "know" that you are thinking of the sound") to 5 ("Perfectly realistic and as vivid as the actual sound"). All scale-points here (and in the other imagery questionnaires with the exception of the Psi-Q; see below) mirror the wording of the VVIQ (for this wording, see above). To achieve this, we slightly re-worded the original response-scale (which had been: 1-'Not at all', to 5-'Very clear') to ensure consistency with other imagery questionnaires in our study (e.g., VVIQ). Total scores range from 16-80.

Vividness of Olfactory Imagery Questionnaire (VOIQ; Gilbert, Voss, & Kroll, 1997). Participants were asked to imagine four different odorous situations (e.g., "An

outdoor cookout or barbeque"). Participants rated the vividness of their olfactory imagery for four aspects of each scenario (e.g., "The charcoal or wood has just been lit and is beginning to burn") on a scale from 1 ("No odour at all, you only "know" that you are thinking of the odour") to 5 ("Perfectly realistic and as vivid as the actual odour"). Total scores range from 16-80.

Adapted Shortened Betts' Questionnaire Upon Mental Imagery (Sheehan, 1967; Spiller et al., 2015). We used this questionnaire to measure imagery for taste, tactile, and bodily sensations. Participants were asked to imagine tasting 12 items (e.g., "Coffee"), touching 12 items (e.g., "Sand"), and the experience of 12 bodily sensations (e.g., "Hunger"). Given the methodological issues noted in our Introduction, we used the version of the Betts' questionnaire updated by Spiller et al. (2015). This version differs from the original (used by Dawes et al. 2020) in a number of ways. First, our version replaces out-dated items (e.g., fur muff → fur), and makes items sensorily clearer where necessary (e.g., jelly → strawberry jelly). Also, in place of just 5 items in the original, it uses 12 items per sense - and the novel items (n7 per modality) are again contemporary (e.g., "Clingfilm (plastic wrap, Saran wrap)"). Since Spiller et al.'s updates had been made to only three scales (taste, tactile, and bodily sensations), we used just these three scales in our own work. Participants rated their imagery on a scale from 1 ("No tactile sensation/taste/bodily sensation at all, you only "know" that you are thinking of the tactile sensation/taste/bodily sensation") to 5 ("Perfectly realistic and as vivid as the actual tactile sensation/taste/bodily sensation"). Responses for each imagery scale (taste, tactile, bodily sensations) were summed separately, with possible scores ranging 12-60.

Vividness of Movement Imagery Questionnaire 2 (VMIQ-2; Roberts, Callow, Hardy, Markland, & Bringer, 2008). Using the kinaesthetic subscale of this questionnaire, participants were asked to imagine the feeling of performing 12 movements (e.g., "jumping sideways"). The direction of the scoring scale was reversed to match our other imagery questionnaires, meaning participants rated vividness from 1 ("No image at all, you only "know" that you are thinking of the skill") to 5 ("Perfectly clear and vivid as normal feel of movement"). Responses were summed, with possible scores ranging from 12-60.

Plymouth Sensory Imagery Questionnaire (Psi-Q; Andrade et al., 2014). This imagery questionnaire covers multiple domains, and was used in addition to our separate measures to ensure reliability of our results. Participants were asked to form a mental image of five items for each of the seven domains: visual (e.g., "a cat climbing up a tree"), auditory (e.g., "an ambulance siren"), olfactory (e.g., "a rose"), gustatory (e.g., "toothpaste"), tactile (e.g., "a soft towel"), bodily sensation (e.g., "having a sore throat"), feeling (e.g., "excited"). Participants rated each item on a scale from 0 ("No image at all") to 10 ("Image as clear and vivid as real life"). This questionnaire provided imagery scores for each sense domain by averaging the items for each modality separately, with possible scores 0-10.

Glasgow Sensory Questionnaire (GSQ; Robertson & Simmons, 2013). This 42 item questionnaire measures sensory sensitivity across seven sense domains (visual, auditory, olfactory, tactile, gustatory, vestibular, and proprioceptive) with six items per sense. Within each sense, half of the items (n=3) measure hyper-sensitivity and half measure hypo-sensitivity. Examples for hyper-sensitivity include "Do bright lights ever hurt your eyes/cause a headache?" (visual) and items for hypo- sensitivity include "Do you really like listening to certain sounds (for example, the sound of paper rustling)?" (auditory). Each question is rated on a scale of 0 ("Never"), 1 ("Rarely"), 2 ("Sometimes"), 3 ("Often"), and 4 ("Always"). The measure outputs an overall sensitivity score summed across all items (ranging 0 to 168), as well as one score for each of the seven senses (e.g., auditory; ranging 0 to 24). It also produces two scores collapsed over senses for hypo- and hyper-sensitivity respectively (ranging from 0 to 84 each).

The Autism Spectrum Quotient (Baron-Cohen et al., 2001). This 50 item questionnaire asks about five types of autism symptomatology: imagination (e.g. "I find making up stories easy"; reversed scored), social skills (e.g. "I find it hard to make new friends"), communication (e.g. "I frequently find that I don't know how to keep a conversation going"), attention switching (e.g. "I prefer to do things the same way over and over again"), and attention-to-detail (e.g. "I tend to notice details that others do not"). Participants rated their agreement on a 4-point scale: "Definitely agree", "Slightly agree", "Slightly disagree", "Definitely disagree". Responses are scored 1 for slightly/definitely agreeing with autism traits (poor imagination, communication, social skills, attention switching, but good attention-to-detail), or 0 otherwise. Scores of ≥32 are taken as the

usual suggestive threshold for autism (Baron-Cohen et al., 2001). However, following Dance, Jaquiery, et al. (2021), we excluded one item that directly taps aphantasia phenomenology ('I find it very easy to create a picture in my mind'), leaving AQ scores ranging from 0 to 49.

Results

Do Aphantasics have Poor Imagery in Multiple Domains?

Using the Psi-Q (i.e., our measure of imagery across multiple sense domains), we first conducted a 2x7 ANOVA crossing group (aphantasia, controls) with sense (visual, auditory, tactile, olfactory, gustatory, bodily, feeling). Aphantasics reported significantly weaker imagery than controls overall ($F(1, 300) = 858.17$, $p < .001$, $\eta_p^2 = .741$; aphantasics: $M = 1.29$; $SD = 1.94$; controls: $M = 7.62$; $SD = 1.78$) and there was also a main effect of sense domain ($F(3.91, 1173.01) = 22.72$, $p < .001$, $\eta_p^2 = .070$; with Greenhouse-Geisser correction; e.g., olfaction was a weaker imagery domain, see Figure 2). There was also a significant interaction ($F(3.91, 1173.01) = 20.17$, $p < .001$, $\eta_p^2 = .063$; with Greenhouse-Geisser correction) because group differences were more pronounced for some senses over others (e.g., more pronounced for vision, less for feeling; see Figure 2). However, group differences were significant (i.e., aphantasics had weaker imagery) for every sense modality, using a series of bootstrapped independent samples t-tests with Welch correction (all $p < .001$, with large hedges' g effect sizes = 1.77-4.48; see table S1 in Appendix A for further details). Here, and throughout our manuscript, we report bias-corrected and accelerated (BCa) bootstrapped confidence intervals (with bootstrapping performed 1000 times) for group comparisons where deviations from normality are present in our data. After vision, tactile imagery was most affected, followed by olfactory, gustatory, bodily, auditory, and feeling imagery (see Figure 2). All differences survive correction for the multiple pairwise comparisons performed using the Benjamini-Hochberg False Discovery Rate method (Benjamini & Hochberg, 1995, 2000).

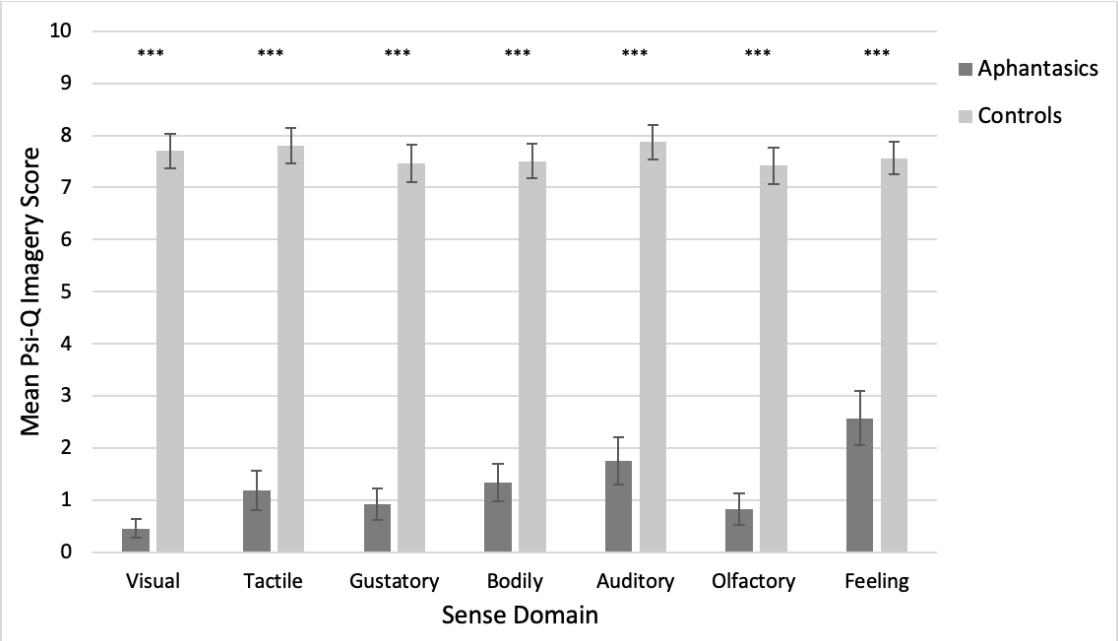

Figure 2. Mean imagery scores (with 95% confidence intervals) as a function of group (aphantasia, control) and sense domain using the Psi-Q. Higher scores indicate stronger imagery (scores for each sense domain are on the same scale ranging from 0-10). *Note.* * uncorrected $p < .05$, ** $p < .01$, *** $p < .001$.

Next, we sought to replicate our findings using our other sensory imagery questionnaires (auditory *CAIS*; olfactory *VOIQ*; gustatory *Betts-ad*; tactile *Betts-ad*; body sensation *Betts-ad*; movement *VMIQ*). Given that our imagery questionnaires were independent measures and had different total possible scores (see methods), we ran a series of bootstrapped independent samples t-tests with Welch correction comparing aphantasics to controls. Again, group differences were significant for every sense (i.e., aphantasics had weaker imagery; all $p < .001$, with large hedges' g effect sizes 2.33-3.37; see table S1 in Appendix A for further details). These differences are depicted visually in Figure 3, and all differences again survive correction for multiple pairwise comparisons.

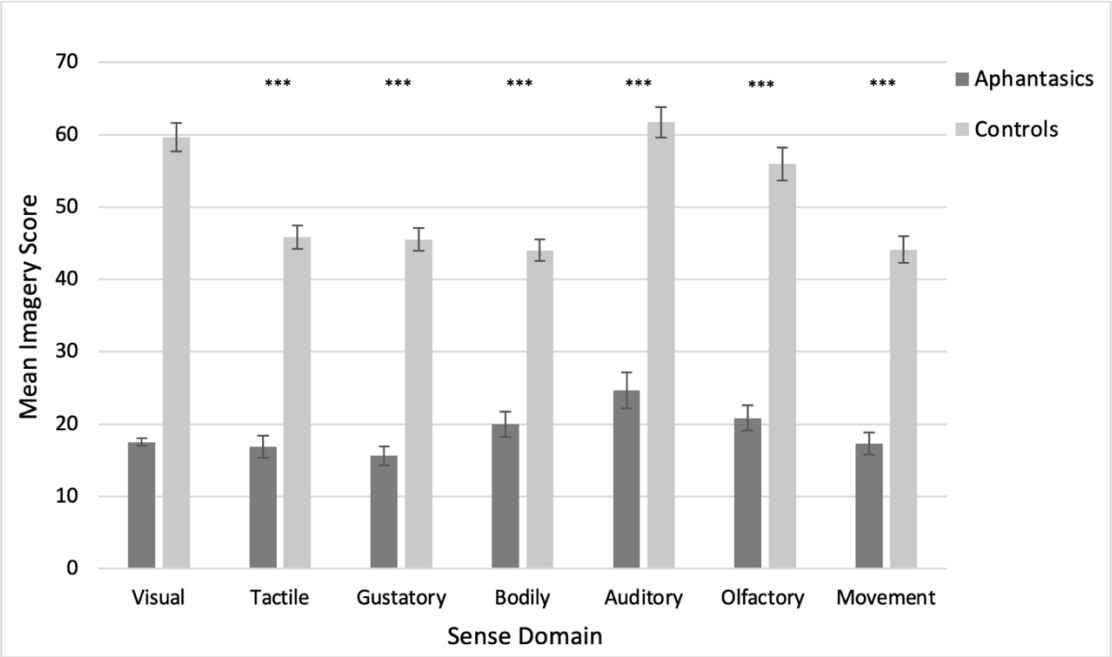

Figure 3. Mean imagery scores (with 95% confidence intervals) as a function of group (aphantasia, control) and sense domain using the independent sensory imagery questionnaires (visual; *VVIQ*, auditory; *CAIS*, olfactory; *VOIQ*, gustatory; *Betts-ad*, tactile; *Betts-ad*, body sensation; *Betts-ad*, movement; *VMIQ*). Higher scores indicate stronger imagery (maximum possible scores for each sense domain vary depending on the questionnaire; see methods). *Note.* * uncorrected $p < .05$, ** $p < .01$, *** $p < .001$ (we did not conduct a group comparison of VVIQ scores because we used this measure to categorise people as aphantasic; see methods).

Next, we asked whether the imagery deficits seen in our aphantasics were severe enough to be considered 'aphantasia-like' (i.e., imagery is entirely absent or vague/dim), even outside the visual domain. We therefore looked within each independent sensory imagery questionnaire (visual; *VVIQ*, auditory; *CAIS*, olfactory; *VOIQ*, gustatory; *Betts-ad*, tactile; *Betts-ad*, body sensation; *Betts-ad*, movement; *VMIQ*) and identified participants who scored within the standard aphantasia range (imagery is absent or vague/dim: scoring 16-32 on the *VVIQ*, *CAIS*, *VOIQ*; and 12-24 on *Betts-ad*, *VMIQ*), albeit for non-visual imagery. We found that 159 (97%) aphantasics showed aphantasia-like imagery weakness in at least one other (non-visual) domain, compared to 15 controls (11%), which was highly significant (χ^2 (1, N = 302) = 223.89, $p < .001$; using Yates correction). Moreover, 101 aphantasic individuals (62%) had aphantasia-like imagery (absent or vague/dim) across *all* other sense domains, compared to 0 controls (χ^2 (1, N = 302) = 124.94, $p < .001$), and 56 aphantasics (34%) reported *no imagery at all* (not vague/dim) in *any domain at all* (compared to 0 controls; χ^2 (1, N = 302) = 55.61, $p < .001$). In

summary, aphantasics not only had weaknesses in *visual* imagery, but almost always (97%) had another imagery deficit (e.g., olfactory), and sometimes (62%) had deficits across *all* imagery modalities.

Does Imagery Predict Sensory Sensitivity?

Next, we asked whether aphantasics reported fewer sensory sensitivities than controls by looking at the GSQ. We conducted a 2x2x7 ANCOVA crossing group (aphantasic, control) with sensitivity type (hyper-, hypo-sensitivity) and sense domain (GSQ; visual, auditory, olfactory, tactile, proprioception, vestibular, gustatory). We included AQ scores as a covariate to control for the influence of autism traits. The ANCOVA revealed a significant main effect of group ($F(1, 299) = 27.08$, $p < .001$, $\eta_p^2 = .083$), with aphantasics reporting fewer sensitivities overall ($M = 50.52$, $SD = 22.88$) relative to controls ($M = 63.62$, $SD = 32.89$).

Although there was no significant main effect of sensitivity type (hyper/hypo), ($F(1, 299) = 1.81$, $p = .277$, $\eta_p^2 = .004$; with Greenhouse-Geisser correction), there was a significant interaction between sensitivity-type and group ($F(1, 299) = 9.16$, $p = .003$, $\eta_p^2 = .030$; with Greenhouse-Geisser correction). Figure 4 suggests that although aphantasics are less sensitive than controls for both hyper-sensitivities and hypo-sensitivities, the effect is smaller for the former. We conducted two comparisons using bootstrapped independent samples t-tests with Welch correction to confirm that aphantasics reported significantly fewer hyper-sensitivities ($M = 27.68$; $SD = 13.60$), ($t(263.81) = 2.84$, $p = .005$, $g = 0.33$, BCa 95% CI [1.78, 8.37]), *and* hypo-sensitivities ($M = 22.84$, $SD = 10.80$), ($t(224.09) = 4.81$, $p < .001$, $g = 0.58$, BCa 95% CI [5.05, 11.17]), than controls (hyper: $M = 32.72$, $SD = 16.68$; hypo: $M = 30.90$, $SD = 17.00$; see Figure 4). Both differences survive correction for multiple pairwise comparisons using the Benjamini-Hochberg False Discovery Rate method (Benjamini & Hochberg, 1995, 2000).

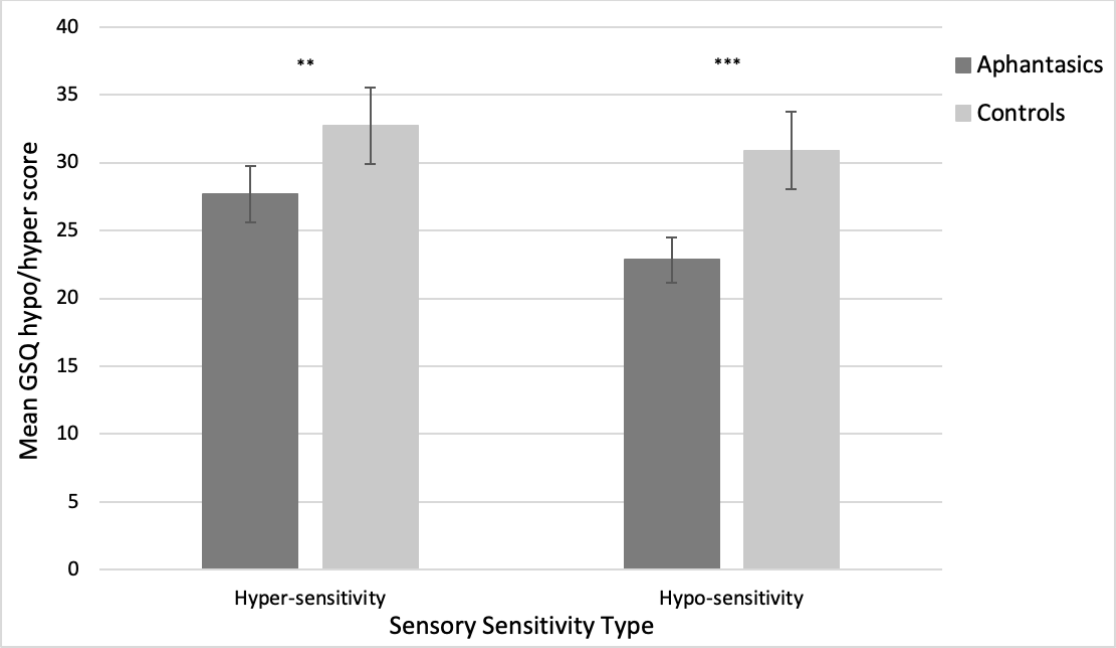

Figure 4. Mean hypo- and hyper-sensitivity scores (with 95% confidence intervals) as a function of group (aphantasia, control) using the GSQ. Higher scores indicate higher levels of sensory sensitivity. *Note.* * uncorrected p <.05, ** p <.01, *** p <.001.

Our ANCOVA also revealed a significant interaction between group and sense domain, (F(5.14, 1538.13) = 7.41, p <.001, η_p^2 = .024; with Greenhouse-Geisser correction). To explore this interaction, we conducted a series of comparisons using bootstrapped independent samples t-tests with Welch correction, depicted visually in Figure 5 (see Appendix A, Table S2 for further information). Aphantasics reported significantly fewer sensitivities in the visual, olfactory, tactile, gustatory, vestibular, and proprioception domains, relative to controls (all p <.01, g = 0.41-0.54). All differences here survive correction for multiple comparisons. There was, however, no difference between aphantasics and controls in auditory sensory sensitivity (p = .791, g = 0.03). A Bayes Factor (using R version 3.5.1, R Core Team, 2018; Bayes Factor version 4.2, Morey & Rouder, 2018) of 0.13 suggested no group differences, with moderate confidence (assuming BF <.33 is taken as evidence for the null hypothesis; Dienes, 2014). Overall, this shows that with the exception of audition, aphantasics report lower levels of sensory sensitivity across all sensory modalities in comparison to imaging controls, while also controlling for the influence of autism traits.

Our ANCOVA also revealed other effects unrelated to our hypotheses (which we did not explore further, to reduce proliferation of multiple comparisons). For example, there was

a significant main effect of AQ score, ($F(1, 299) = 92.94$, $p < .001$, $\eta_p^2 = .237$) and sense domain, ($F(5.14, 1538.13) = 12.60$, $p < .001$, $\eta_p^2 = .040$; Greenhouse-Geisser correction). As expected from the autism literature (American Psychiatric Association, 2013) higher AQ scores linked with higher sensory sensitivity, and as expected from the sensitivity literature (Robertson & Simmons, 2013), some sense domains were more sensitive than others (e.g., auditory domain most sensitive; see Figure 5).

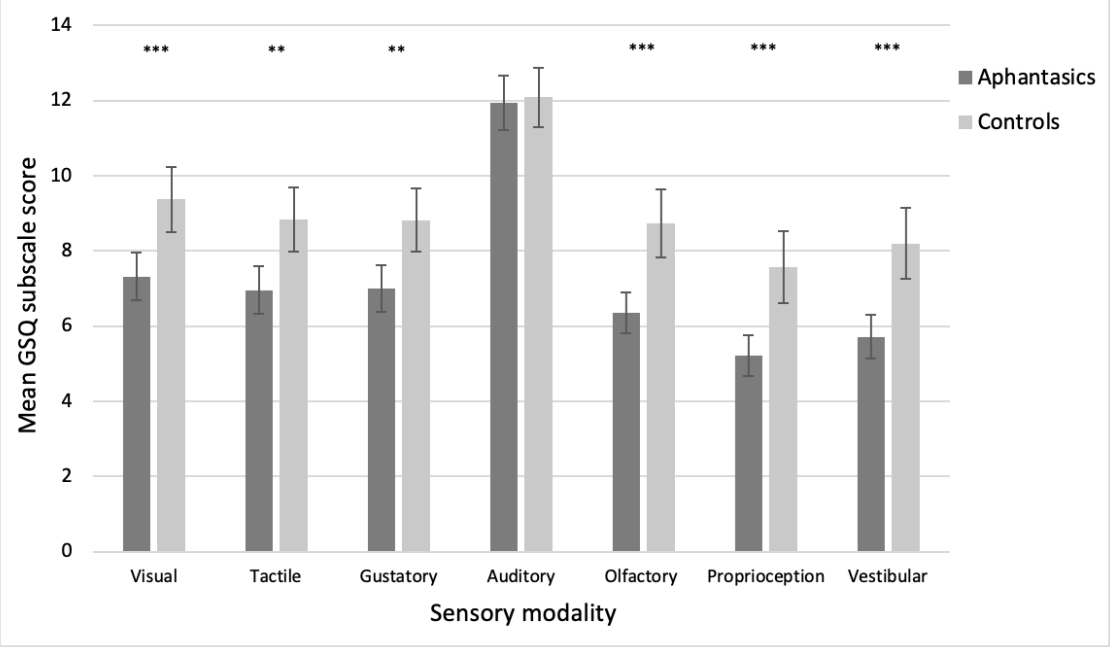

Figure 5. Mean overall sensory sensitivity scores (with 95% confidence intervals) as a function of group (aphantasia, control) and sense domain using the GSQ. Higher scores indicate higher levels of sensory sensitivity (encompassing hypo- and hyper-sensitivity). *Note.* * uncorrected $p < .05$, ** $p < .01$, *** $p < .001$

Discussion

Here, we found that aphantasics experience significantly weaker imagery than controls across multiple sense domains, and are significantly more likely to have aphantasia-like imagery-weakness across multiple (and even all) senses compared to controls. Indeed, 97% of aphantasics (i.e., almost every one) had imagery deficits not only in the visual domain but also in at least one other. Aphantasics also reported fewer sensory sensitivities overall (in both hyper- and hypo-sensitivity), and fewer sensitivities within each of the senses with the exception of audition. In sum, our results show that imagery and sensory sensitivity are related: people with aphantasia experience lower levels of imagery and sensory sensitivity across multiple sense domains.

Experiment 2

In Experiment 1 we demonstrated that imagery and sensory sensitivity are linked. But we are left with the question of whether the link between imagery and sensory sensitivity is seen specifically in aphantasic people, or whether it applies to the general population also. In Experiment 2 we addressed this question, by examining whether there is an association between visual imagery and sensory sensitivity in a student general population sample.

Methods

Participants

Our participants were 83 undergraduate students registered at the University of Sussex (63 females, 20 males; *M* age = 19.87, *SD* = 3.60). Participants took part in our study in return for research participation credits, and were sampled without mention of imagery or aphantasia in order to represent a random sampling of the population. Six of these participants had VVIQ scores in the aphantasia range (16-32), but they are included here because their scores represent part of the natural continuum of imagery within a general population sample.

Materials and Procedure

Participants completed the VVIQ (Marks, 1973), GSQ (sensory sensitivity measure; Robertson & Simmons, 2013), and the AQ (to factor out autistic traits; Baron-Cohen et al., 2001) in a random order. Details of these three measures are described in Experiment 1. Participants were provided with a URL to the testing site (www.syntoolkit.org) and they completed the study from their own homes.

Results

To examine whether there was an association between imagery and sensory sensitivity we conducted a linear regression (using the enter method) predicting sensory sensitivity (GSQ) from visual imagery (VVIQ) and autism traits (AQ score). We constructed two regression models, the first predicting GSQ score from AQ score, and the second adding VVIQ score as an additional predictor. Our data was normally distributed, and the residuals in our models met the required assumptions for parametric tests (so confidence intervals were not bootstrapped). Both model one ($F(1, 81) = 21.35$, $p < .001$) and model

two ($F(2, 80) = 13.18$, $p < .001$) significantly predicted sensory sensitivity scores. Model two was a significantly better model than model one, explaining 24.8% of the variance in sensory sensitivity scores (3.9% more than model one), ($F(1, 80) = 4.17$, $p = .045$). As expected from prior literature (Robertson & Simmons, 2013), model two revealed a significant positive relationship between autism traits (AQ) and sensory sensitivity (GSQ score; $b = 1.24$, SE(b) = .27, $t = 4.61$, $p < .001$, 95% CI [.706, 1.78]), showing that sensory sensitivity increased with levels of autism traits. Importantly, model two also showed a significant positive association between VVIQ score and GSQ score, ($b = .26$, SE(b) = .13, $t = 2.04$, $p = .045$, 95% CI [.007, .519]), indicating that as visual imagery scores increased, so did overall sensory sensitivity scores. Overall, our analysis shows that even when taking into account the influence of autistic traits (std. $b = .45$), visual imagery is a significant positive predictor (std. $b = .20$) of sensory sensitivity (see Figure 6 for the distribution of VVIQ scores).

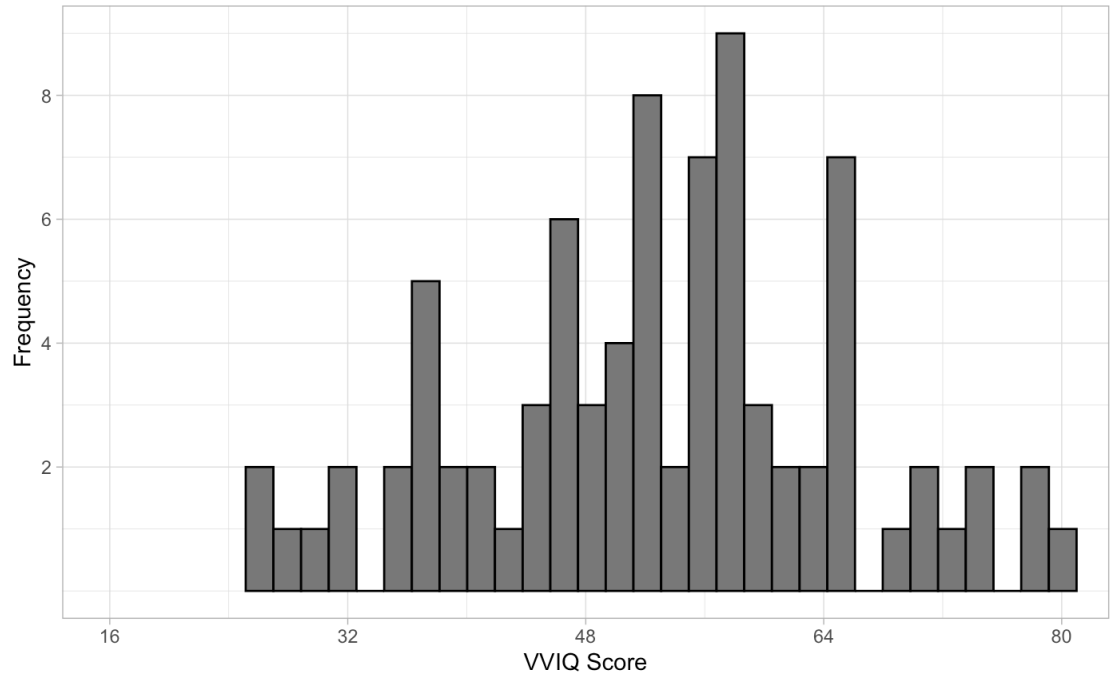

Figure 6. A histogram to show the distribution of VVIQ scores in our student general population sample.

Discussion

In Experiment 2 we investigated the relationship between visual imagery and sensory sensitivity in a general population sample taken from a student body. We found that imagery and sensory sensitivity are positively correlated, even when autistic traits are

factored out. Our results therefore show that imagery is not only associated with levels of sensory sensitivity in people with aphantasia (Experiment 1), but that this relationship is also evident in the general population (Experiment 2). In other words, while in Experiment 1 we found that people with aphantasia (who are low in imagery) tend to experience reduced levels of sensory sensitivity, here we found further support in the general population: as visual imagery level increases, so does overall sensory sensitivity.

Experiment 3

In Experiments 1 and 2 we demonstrated that people who report lower visual imagery also report lower sensory sensitivities – both in aphantasia, and in the general population. Experiment 3 aims to validate our findings using behavioural evidence from a *Pattern Glare Task*. In response to gratings designed to elicit pattern glare, we asked whether people with poor imagery would experience less visual sensory sensitivity (less visual discomfort, and fewer visual distortions) than we might otherwise expect. Here, we compared a sample of people with aphantasia (low imagery) to controls with typical visual imagery abilities. We predict that the former will show lower pattern glare effects, indicative of lower sensory sensitivity.

Methods

Participants

Our participants were 56 aphantasics (28 female, 26 male, 2 other; *M* age = 33.66; *SD* = 8.36) and 56 controls (39 female, 17 male; *M* age = 29.84, *SD* = 16.91). Participants were matched in age ($t(80.39) = -1.52$, $p = .133$). Controls were recruited from the same sources as Experiment 1, with the addition of the undergraduate student body at the University of Sussex, and excluding Amazon's Mechanical Turk. Most of our aphantasic participants were recruited from the University of Sussex's *Imagery Lab - Aphantasia Cohort* (n=54), with the remaining participants (n=2) recruited via the same sources as controls. Participants were verified as aphantasics in the same way as in Experiment 1 (scoring 16-32 on the VVIQ; *M* = 17.77, *SD* = 3.86; compared to controls who scored >32 on the VVIQ; *M* = 58.89, *SD* = 10.62). Of our participants, 24 also took part in Experiment 1 (19 aphantasics, 5 controls), and 5 also took part in Experiment 2 (5 controls). Participants were advised not to take part in the study if they had a history of epilepsy

given that uncomfortable visual patterns can trigger photosensitive epilepsy (Wilkins, Binnie, & Darby, 1980). Participants were compensated in the same way as Experiment 1.

Materials and Procedure

All participants completed the pattern glare task online using the testing platform Inquisit (*Inquisit 5*, 2016). After the pattern glare task, participants completed the VVIQ, with the exception of participants who had already completed it previously. Please see Experiment 1 for a description of these latter participants (from our *Imagery Lab - Aphantasia Cohort*), as well as the VVIQ protocol.

Pattern Glare Task. Participants viewed achromatic gratings, which were oval in shape and consisted of black and white horizontal stripes, presented on a grey background (RGB 128, 128, 128). Our version was based on Ward et al. (2017), who used gratings from Braithwaite, Broglia, Brincat, et al. (2013). There were three types of grating, each differing in spatial frequency: *low* (approx. 0.4 cycles per degree; cycles per degree is a measure of spatial frequency, see Braithwaite, Broglia, Bagshaw, et al., 2013; Ward et al., 2017; Wilkins et al., 1984), *medium* (approx. 3 cycles per degree), and *high* (approx. 10 cycles per degree). The low spatial frequency grating acted as a baseline stimulus, whilst the medium and high spatial frequency gratings were 'irritable' gratings expected to elicit 'pattern glare' (Braithwaite, Broglia, Bagshaw, et al., 2013; Braithwaite, Broglia, Brincat, et al., 2013; Conlon, Lovegrove, Barker, & Chekaluk, 2001; Evans & Stevenson, 2008; Wilkins et al., 1984). Gratings were presented at their actual size of 230.01mm wide x 176.39mm high in the centre of the screen, and participants were instructed to sit 80cm/32 inches from the monitor. A calibration procedure ensured that gratings were presented at the standardised size (approx. 12.58 degrees in height) on each computer monitor (participants adjusted the length of a line to match the length of a bank card, to calculate the number of pixels required for the gratings to reach the standardised size).

Participants were told they should fixate their gaze on the centre of each pattern presented. Each grating was presented twice (each for 5 seconds) in a random order (6 trials overall). Following each grating (i.e., each trial), participants were asked two questions: how uncomfortable the image was, and how many visual effects were experienced. Comfort was measured on a sliding scale from 1 ("Extremely uncomfortable") to 11 ("Extremely

comfortable"), and visual effects were reported by checking any that applied (colours, bending of the lines, blurring of the lines, shimmering of the lines, flickering, fading, shadowy shapes, other effects; total effects for each grating ranging from 0-8). See Figure 7 for an example of the trial sequence. For each grating type (low, medium, high) the visual discomfort ratings and the number of visual effects elicited were averaged separately. Participants began the task with a practise grating of black and white checkers.

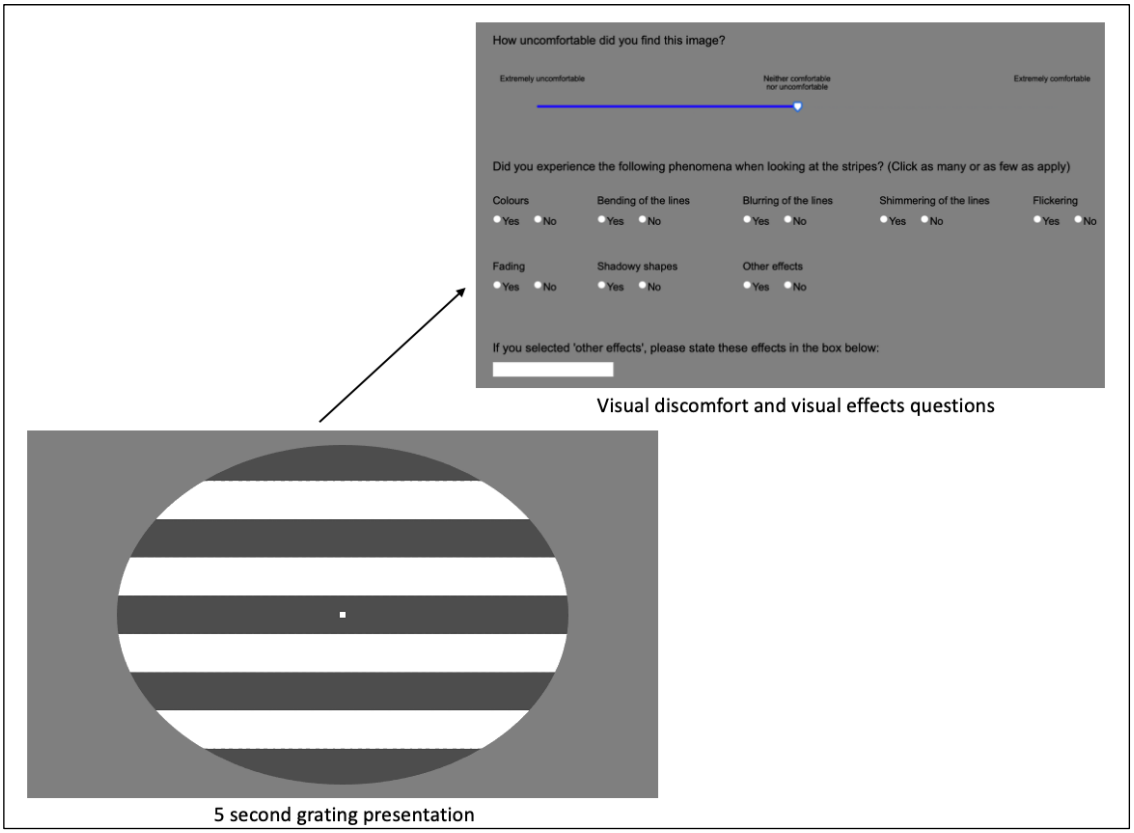

Figure 7. The trial sequence in the Pattern Glare Task (using an example of a low spatial frequency grating). In each trial the grating was presented for 5 seconds, followed by questions about visual discomfort and visual effects experienced.

Results

For each of our measures (see below), we conducted a 2x3 ANOVA crossing group (aphantasic, control) and grating spatial frequency (high, medium, and low). These were followed by bootstrapped independent samples t-tests with Welch correction to examine differences in response to the gratings designed to elicit visual sensory sensitivity (medium- and high-spatial frequency). We did not conduct pairwise comparisons for the low (baseline) grating to avoid proliferation of multiple comparisons, and because our hypotheses did not motivate this (i.e., it has been well established that low spatial

frequency gratings do not trigger visual sensory sensitivity; e.g., Braithwaite, Broglia, Bagshaw, et al., 2013; Evans & Stevenson, 2008; Ward et al., 2017; Wilkins et al., 1984).

The ANOVA for visual discomfort ratings revealed the expected main effect of spatial frequency ($F(1.78, 195.81) = 46.25$, $p < .001$, $\eta_p^2 = .269$; with Greenhouse-Geisser correction), with more visual discomfort from the high ($M = 5.31$, $SD = 1.67$) and medium gratings overall ($M = 5.81$, $SD = 1.71$), relative to the low (baseline) grating ($M = 6.70$, $SD = 1.75$). There was no significant main effect of group ($F(1, 110) = 2.21$, $p = .140$, $\eta_p^2 = .020$), but there was a significant interaction ($F(1.78, 195.81) = 4.25$, $p = .019$, $\eta_p^2 = .037$; with Greenhouse-Geisser correction). Although there was no significant difference in discomfort for the medium grating between aphantasics ($M = 6.04$, $SD = 1.84$) and controls ($M = 5.58$, $SD = 1.54$; $t(106.73) = -1.42$, $p = .159$, $d = 0.27$, BCa 95% CI [-1.09, .156]; Bayes factor = 0.49, i.e., anecdotal support for null hypothesis), there was a significant difference for the high spatial frequency grating. Here, aphantasics reported significantly less visual discomfort ($M = 5.71$, $SD = 1.70$) than controls ($M = 4.91$, $SD = 1.55$), ($t(109.03) = -2.62$, $p = .010$, $d = 0.49$, BCa 95% CI [-1.39, -.249]). This demonstrates that although both groups found the high grating uncomfortable (scores <6 indicate visual 'discomfort', and scores >6 indicate visual 'comfort'), aphantasic individuals reported less discomfort than controls (see Figure 8, where scores are recoded around zero for ease-of-display). This result survives when correcting for the multiple comparisons performed using the Benjamini-Hochberg False Discovery Rate method (Benjamini & Hochberg, 1995, 2000).

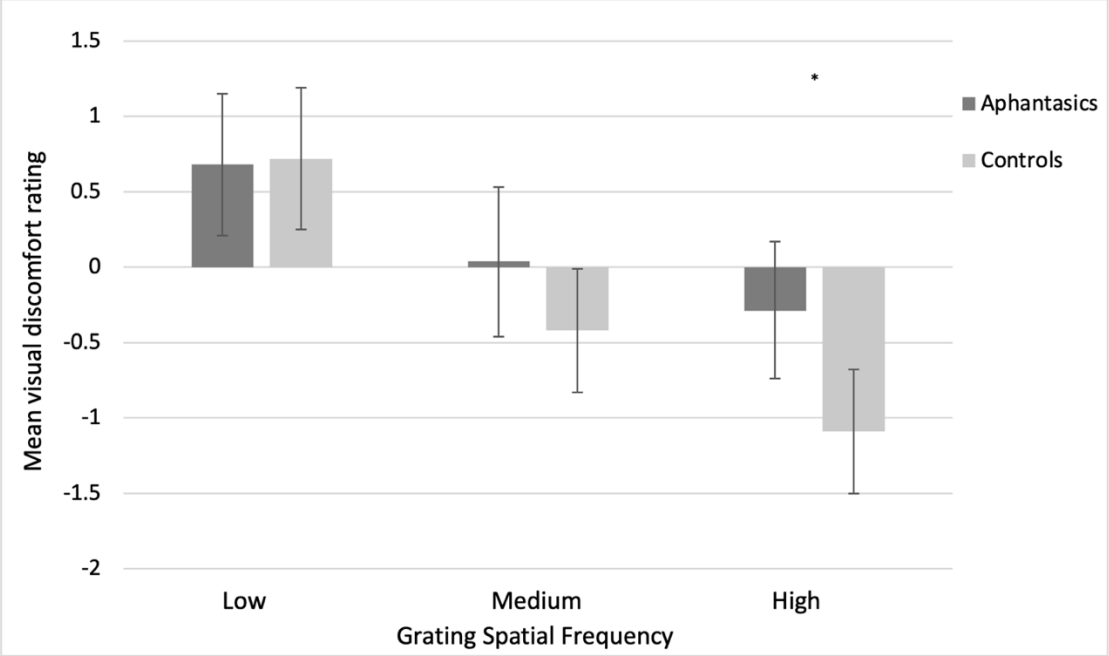

Figure 8. Mean visual discomfort ratings (with 95% confidence intervals) as a function of group (aphantasia, control) and grating spatial frequency (low, medium, high). For ease of visual display, scores were re-coded from a scale of 1-11 to a scale of -5 to 5 where scores above 0 indicate visual 'comfort', and scores below 0 indicate visual 'discomfort'. *Note.* * uncorrected p <.05, ** p <.01, *** p <.001.

We next looked at the *number* of visual effects for each participant. As before, and as expected, there was a main effect of spatial frequency, (F(1.82, 199.77) = 130.70, p <.001, η_p^2 = .543; with Greenhouse-Geisser correction) with more visual effects from high (M = 2.86, SD = 1.68) and medium gratings overall (M = 2.19, SD = 1.43), relative to the low (baseline) grating (M = .79, SD = 1.00). Our analysis also revealed a significant interaction (F(1.82, 199.77) = 6.95, p = .002, η_p^2 = .059; with Greenhouse-Geisser correction) because group differences were more pronounced for some gratings than others (e.g., for the high grating; see below and Figure 9), but no significant effect of group, (F(1, 110) = 3.18, p = .077, η_p^2 = .028).

As before, for the high spatial frequency grating, aphantasics reported significantly fewer visual effects (M = 2.50, SD = 1.44) than controls (M = 3.21, SD = 1.83), (t(104.41) = 2.30, p = .024, d = 0.43, BCa 95% CI [.161, 1.22]). However, aphantasics also reported significantly fewer visual effects (M = 1.89, SD = 1.26) than controls (M = 2.49, SD = 1.53) in response to the medium grating (t(106.03) = 2.26, p = .026, d = 0.43, BCa 95%

CI [.110, 1.06]); see Figure 9), and all differences survive correction for multiple comparisons.

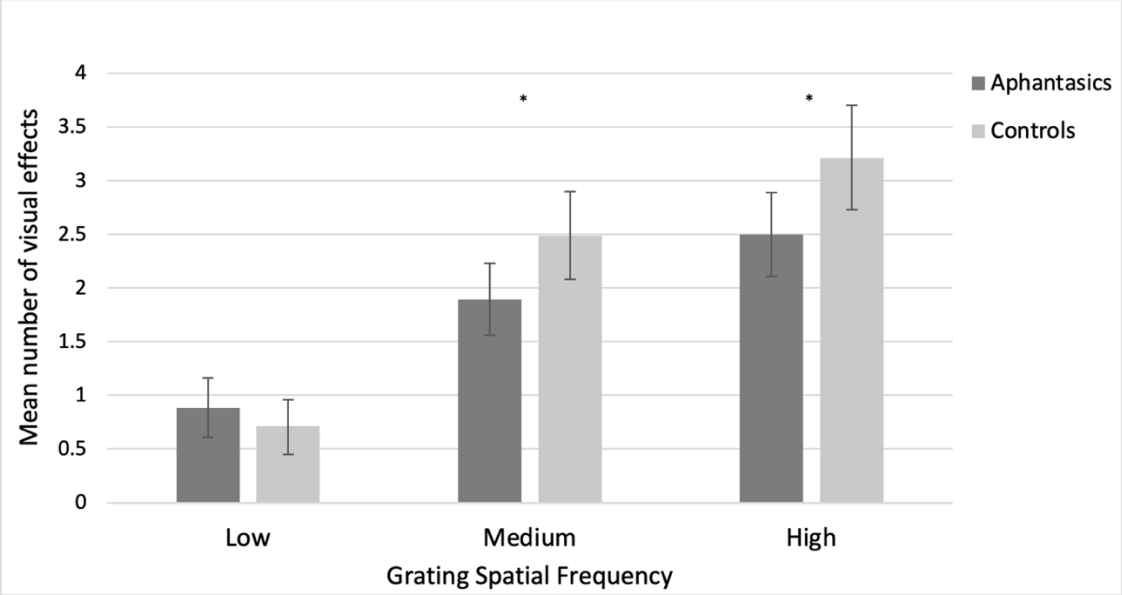

Figure 9. Mean number of visual effects (with 95% confidence intervals) as a function of group (aphantasia, control) and grating spatial frequency (low, medium, high). *Note.* * uncorrected p <.05, ** p <.01, *** p <.001.

Discussion

Here, we again present evidence that people with aphantasia experience less visual sensory sensitivity than people with visual imagery. As predicted, aphantasics reported significantly less pattern glare than controls: high/medium gratings gave them fewer visual effects, and the high gratings gave them less visual discomfort. These results provide behavioural corroboration for Experiment 1, showing that people with aphantasia experience lower levels of sensory sensitivity than imaging controls.

General Discussion

Our intention was to characterise the sensory imagery deficits and sensory sensitivities experienced by people with aphantasia, and in doing so, examine how imagery and sensory sensitivity may be linked. Until now, aphantasia has been characterised by an absence of *visual* imagery (e.g., Milton et al., 2021; Zeman et al., 2015, 2016, 2020). But in Experiment 1 we found that aphantasic individuals report significantly weaker imagery compared to controls within all of the sense domains we tested (visual, olfactory, tactile, gustatory, bodily sensation, feeling, movement). When considering imagery in other

senses in terms of the same criteria as aphantasia (i.e., imagery that is absent/dim/vague), we found that almost 100% of people with aphantasia had poor, aphantasia-like imagery in at least one other sense (i.e., 97% of aphantasics, but just 11% of controls), and as many as 62% of aphantasics had imagery-weaknesses in *all* of their senses, compared to 0% of controls. In the most extreme cases, one third of aphantasics (34%) reported *no imagery whatsoever* in *any sense whatsoever*.

Our results are in line with research showing that imagery strength is often positively correlated across sensory modalities (Andrade et al., 2014; Leclerc et al., 2019; Lima et al., 2015; White et al., 1974), and fits with recent studies using different methods. For example, Dawes et al. (2020) found that 26% of aphantasics reported a complete lack of imagery across all sense domains they tested in their own study (compared to our 34%). Their paper was an important step in understanding the broader imagery phenomenology of aphantasia, although their imagery questionnaire (*Sheehan's adapted version of Betts' Questionnaire Upon Mental Imagery*; Sheehan, 1967) has been criticised for its small number of items and outdated language (e.g., Pecher et al., 2009; Spiller et al., 2015). Nonetheless, our results converged with their own, allowing our paper to stand as a validation and replication of their findings using multiple contemporary imagery questionnaires (see Experiment 1). Our results also fit with personal reports from aphantasic writers (Watkins, 2018) and with data from 1000 aphantasics asked to self-diagnose imagery deficits, in which 54% reported imagery weaknesses across all sense modalities (Zeman et al., 2020) – compared to our own 62%. In sum, our results provide robust evidence that people with aphantasia tend to self-report imagery that is weak across multiple senses.

An important question that emerges from our findings is *why* aphantasics often experience multi-modal imagery deficits. We have suggested that understanding the neural mechanisms underlying mental imagery may shed light on this. Engaging in imagery is associated with activation in the relevant sensory cortex (e.g., visual imagery tied with visual cortex) and this has been shown for visual (Cattaneo et al., 2011; Cui et al., 2007; Daselaar, Porat, Huijbers, & Pennartz, 2010; Dijkstra et al., 2017; Ganis, Thompson, & Kosslyn, 2004; Sparing et al., 2002), auditory (Daselaar et al., 2010; Zvyagintsev et al., 2013), olfactory (Djordjevic, Zatorre, Petrides, Boyle, & Jones-Gotman, 2005; Leclerc et al., 2019; Plailly, Delon-Martin, & Royet, 2012), gustatory (Belardinelli et al., 2009;

Kobayashi, Sasabe, Shigihara, Tanaka, & Watanabe, 2011; Kobayashi et al., 2004), tactile (Schmidt, Ostwald, & Blankenburg, 2014; Yoo, Freeman, McCarthy, & Jolesz, 2003), and motor imagery (Grèzes & Decety, 2000; Hanakawa, Dimyan, & Hallett, 2008; Hétu et al., 2013). Here, some studies showed a positive correlation between sensory cortex activation and imagery vividness (Belardinelli et al., 2009; Cui et al., 2007; Herholz, Halpern, & Zatorre, 2012), while others showed that the content of visual imagery can be decoded from visual cortex activity using multivariate decoding in fMRI (Albers, Kok, Toni, Dijkerman, & De Lange, 2013; Koenig-Robert & Pearson, 2019; Naselaris, Olman, Stansbury, Ugurbil, & Gallant, 2015). Similarly, low imagers rely less on visual cortex compared to high imagers when asked to complete visual imagery tasks (mental rotation; Logie et al., 2011), and studies using TMS show that engaging in visual imagery is associated with increased excitation in visual cortex (Cattaneo et al., 2011; Sparing et al., 2002). These findings, and others, linking imagery to activity in sensory cortices, lead us to propose that aphantasia may be characterised by low visual cortex excitation, and that this may in fact be part of a wider deficit across multiple cortices within the same brain. This may give rise to the multi-modal imagery impairments seen in people with aphantasia, consistent with our model in Figure 1.

It is important to recognise that there were rare instances where aphantasics reported *intact* imagery in non-visual senses (i.e., 3% of our aphantasics did *not* have absent/vague/dim imagery in another sense). Therefore, although having low imagery in one domain (visual) is often associated with imagery impairments in other senses, this may not *always* be the case. If our model holds true, having low cortical excitability in one sensory area may increase the likelihood of low excitability in other sensory areas, but we are proposing *likelihoods* not absolutes, and therefore expect individual differences. Testing these models with neuroscientific methods will be a fruitful avenue for future research. Alternatively, however, it is possible that even these 3% did have other imagery weaknesses, but simply in sense domains we did not test (e.g., interoception, thermoception, nocioception, etc.). Therefore, testing these domains will be an important avenue for future research.

A second feature of our model was the prediction that imagery may also link to sensory sensitivity (under- and over-responsiveness to sensory stimuli entering via the sense organs). In support of our model, we found that aphantasic individuals reported fewer

sensory sensitivities, both overall (i.e., for both hypo- and hyper-sensitivity) and within every sense domain tested with the exception of audition. This effect remained even when factoring out autistic traits. In Experiment 2 we extended this finding to the general population, showing again a positive correlation between imagery and sensitivity while again factoring out AQ scores. In Experiment 3, we provided behavioural support: aphantasic individuals demonstrated less sensitivity than controls in a *Pattern Glare Task* (i.e., fewer visual effects from high/medium gratings and less visual discomfort from high gratings). Overall, our findings suggest that imagery and sensory sensitivity are linked, potentially (our model suggests) via cortical excitability. We noted that cortical excitability has been traditionally associated with levels of *hyper*-sensitivity (i.e., increased cortical excitability linked to hyper-responsiveness; Green et al., 2015, 2013; Wilkins et al., 1984). However, hyper- and hypo-sensitivities are thought to be linked via *adaptation* (Takarae & Sweeney, 2017; Ward, 2019a). This led us to predict that aphantasics – who we propose are *low* in cortical excitability – would show not only less hyper-sensitivity but also less hypo-sensitivity. This is indeed what we found. It is unclear why low sensitivity for aphantasics did not extend to auditory items (especially since sensory sensitivity is correlated across modalities). However, the auditory channel is also remarkable in being the most sensitive (Robertson & Simmons, 2013; and see Figure 5 for this clear difference). Further research is needed to better determine whether our null effect is genuine, by using alternative measures of auditory sensory sensitivity including those that employ sound stimuli and not simply questionnaire self-report.

What specific neural mechanisms might underpin low excitation across multiple sensory cortices in aphantasia? Our model fits broadly with neural noise theories from the autism literature which suggest that high levels of sensory sensitivity in people with autism are driven – at least in part – by increased levels of neural noise (excitation) within sensory cortices (Milne, 2011; Rubenstein & Merzenich, 2003; Simmons, 2019; Simmons et al., 2009). This theory links neural noise with an increase in excitatory (e.g., glutamate) synaptic activity and a reduction of inhibitory (e.g., GABA) synaptic activity (e.g., Rubenstein & Merzenich, 2003; Wood et al., 2021). Although speculative at this point, the mechanisms behind low cortical excitability in aphantasia may therefore include lower levels of excitatory glutamatergic synaptic activity, and/or higher inhibitory GABA-ergic synaptic activity in sensory cortices and areas involved in sensory modulation/regulation (e.g., thalamus; Wood et al., 2021). Depending on the sensory

cortices affected by these patterns of excitation, this may give rise to lower imagery, and sensitivity, in many, several, or a single sense domain. As a parallel, we note that differences across disparate brain regions have also been found in a number of other sensory conditions, for example in synaesthesia where hyper-connectivity of white matter can arise in multiple brain regions, giving rise to different manifestations of what is considered the same underlying condition (Rouw & Scholte, 2007; Rouw, Scholte, & Colizoli, 2011).

An important question raised by our findings is how aphantasia and autism are related. People with aphantasia report *high* levels of autism traits (Dance, Jaquiery, et al., 2021), and autism is usually linked with heightened sensory sensitivity (Ben-Sasson et al., 2009; Bogdashina, 2003; Horder et al., 2014; Robertson & Simmons, 2013, 2015; Simmons et al., 2009). But here we found the opposite: aphantasics showing *lower* levels of sensitivity. Moreover, we still found that AQ scores were positively related to sensitivity in our participants overall (when included as a covariate in our analysis of sensory sensitivity in Experiment 1). A simple explanation may lie in the fact that autism is a cluster of traits in different domains, and some populations (e.g., aphantasics) may show higher scores in only certain domains. We know already that aphantasics show weak imagination and social skills, but match controls on other autism traits such as attention-to-detail, attention switching, and communication (Dance, Jaquiery, et al., 2021). In other words, aphantasics may be like people with autism in some ways (imagination, social skills) but not in others (attention-to-detail – and indeed sensory sensitivity). Similar patterns have been found for other conditions, such as people with synaesthesia, who have higher AQ scores because they share the autistic trait of (only) attention-to-detail (Van Leeuwen et al., 2019; Ward, Brown, et al., 2018; Ward et al., 2017). It is possible that individuals with autism (or high in autistic traits) may experience somewhat less intense sensory sensitivity if they *also* have aphantasia. Teasing apart the specific instances where imagery maps directly onto sensory sensitivity will be an interesting avenue for future research. For now, we have begun to unravel the relationship between imagery and sensitivity by showing that aphantasic individuals experience lower sensory sensitivity than people who have visual imagery.

Finally, our study also provides answers about the very nature of aphantasia. Aphantasia was originally named for its links to the Greek word for imagination (φαντασία) and the

related term φάντασμα (*phántasma*, "phantasm, an appearance, image, apparition, spectre" (Online Etymology Dictionary, 2001; Thomas, 2011; Zeman et al., 2015). It also has etymological links to the word φαντάζω (*phantázō*, "I make visible"; Online Etymology Dictionary, 2001) which make it a highly useful term to describe an absence of imagery that is visual per se. However, our finding here that aphantasia is part of a broader deficit in imagery (see also Dawes et al. 2020; Zeman et al., 2020) leads us to suggest a second complementary term, *dysikonesia*. Like aphantasia, we propose that dysikonesia encompasses imagery that is absent or dim/vague, but we introduce this term to describe the *broader* phenotype, i.e., imagery deficits across multiple domains (as found in 97% of aphantasics tested here), or indeed for cases where the particular domain of the imagery deficit has not been specified. Its etymological root (*icon*[8]) has the useful meaning of a form which reflects its referent, in the same way that visual imagery has qualities that reflect visual entities in the world, and auditory imagery has qualities that reflect auditory entities in the world, and so on. People with dysikonesia, we suggest, therefore have visual and auditory (and other sensory) knowledge, but without the iconic quality of imagery. This classification makes aphantasia one sub-class of dysikonesia, and 3% of our aphantasic group experienced this subclass only (i.e., imagery deficit in only vision). Overall, our results provide an important step forward in understanding the experience of imagery in aphantasic individuals, and we open up a wider debate about the phenomenology, and indeed the causes and consequences of absent imagery.

In conclusion, our study is the first to characterise the relationship between mental imagery and sensory sensitivity, the first to model this relationship, and the first to name and model a phenotype of poor imagery cross-senses. Our findings raise questions about the best way to define aphantasia, which we now embed within a broader multi-modal imagery deficit we term *dysikonesia*. We present both self-report and behavioural methods, and propose a model linking imagery and sensory sensitivity via neural excitability within sensory cortices.

[8] In dysikonesia, we have chosen an earlier form of orthography (ikon) simply to disambiguate pronunciation (i.e., avoiding the ambiguity of /c/ = [k] or [s]).

Chapter 5

The role of visual imagery in face recognition, face matching, and the construction of facial composites. Evidence from Aphantasia

Chapter Summary

In the previous chapter I found that the imagery weakness characteristic of aphantasia often extends beyond the visual modality into other sense domains. I also found for the very first time that sensory sensitivity is linked to mental imagery. In Chapter 5, I now consider the perception and memory profile of aphantasics, in terms of face processing. I examine whether aphantasics have poorer face perception and face recognition abilities than people with visual imagery. I also assess whether aphantasics show differences in their ability to construct facial composites. This chapter tests the broader question of whether visual imagery facilitates face processing (Grueter et al., 2007; Milton et al., 2021), and has important implications for forensic settings where face recognition and face composite construction are commonly used to aid criminal investigations.

Full citation

Dance, C. J., Hole, G., & Simner, J. (2022). The role of visual imagery in face recognition, face matching, and the construction of facial composites. Evidence from Aphantasia. [Manuscript under review].

Abstract

People with *aphantasia* have a markedly impaired ability to form visual images in the mind's eye. Here, by testing people with and without aphantasia, we examine the relationship between visual imagery and face processing. We show that aphantasics have weaker face recognition than people with visual imagery, using both self-report (Prosopagnosia Index) and behavioural measures (Cambridge Face Memory Test). However, aphantasics nonetheless have a fully intact ability to construct facial composites from memory (i.e., composites produced using EFIT6 by aphantasics and imagers were rated as equally accurate in terms of their resemblance to a target face). Additionally, we show that aphantasics were less able than imagers to see the resemblance between composites and a target face, indicative of issues with face matching (perception). Finally, we show that holistic and featural methods of composite construction using EFIT6 produce equally accurate composites, as do remote (online) and in-person construction procedures. Overall, our results suggest that face recognition and face matching (but not face composite construction) are facilitated by the ability to represent visual properties as 'pictures in the mind'. Our findings have implications for the study of aphantasia, and also for forensic settings, where face composite systems are commonly used to aid criminal investigations.

Introduction

Faces provide a key source of information in everyday life, allowing us, for instance, to distinguish a friend from a stranger. In addition to its social importance, face recognition is also vital within forensic settings. Eye-witnesses are often required to identify potential suspects in police line-ups (Home Office, 2017) or construct facial composites of offenders, which can be shared with the wider police force or general public (Frowd et al., 2019; Home Office, 2017). People differ widely, however, in their face recognition abilities. Some people are incredibly good at face recognition, whereas others struggle to recognise even familiar faces (e.g., Russell, Duchaine, & Nakayama, 2009). Aside from individual differences in face-recognition, there are also individual differences in the ability to conjure up a mental image of a face in the 'mind's eye' (and other objects too). Some people have very strong visual mental imagery which is nearly as vivid as real life perception, while others – people with *aphantasia* – have little or no visual imagery at all (even though they know and remember what objects look like; e.g., Zeman et al., 2015, 2020). Here, we ask whether visual imagery and face recognition abilities are linked, by administering face recognition and face composite tasks to people with and without aphantasia. We present our experiments testing these ideas below, but begin with a brief overview of face recognition and its potential links to visual imagery, and to aphantasia.

There has been considerable scientific interest in why people differ in face memory, and relatedly, eyewitness accuracy. One explored factor, noted above, is visual mental imagery: i.e., does being able to 'see' faces in our mind's eye make us better at recognising them? Accordingly, some studies have measured whether face recognition abilities are predicted by the strength (or 'vividness') of visual imagery. In a study by McKelvie (1994), participants completed a face recognition test in which they inspected a set of unfamiliar target faces, then identified each one from pairs of subsequently presented faces. To measure their visual imagery strength, participants also completed the *Vividness of Visual Imagery Questionnaire* (*VVIQ*; Marks, 1973). The VVIQ requires participants to imagine different scenes (e.g., "A beach by the ocean on a warm summer's day"), and to rate the vividness of their visual imagery on a 5-point scale, running from 1 ("No image at all, you only "know" that you are thinking of the object"), 2 ("Vague and dim"), 3 ("Moderately clear and vivid"), 4 ("Clear and reasonably vivid"), to 5 ("Perfectly clear and as vivid as normal vision"). The VVIQ has been validated by objective

behavioural methods (e.g., using binocular rivalry; see Keogh & Pearson, 2018; Pearson et al., 2011). McKelvie (1994) showed no difference in face recognition ability for people 'low' or 'high' in visual imagery vividness. Put simply, visual imagery strength did not influence face recognition.

In contrast, other research has claimed to show a link between visual imagery and face recognition, but only when examining *imagery for faces*. For example, a study by Cabeza, Burton, Kelly, and Akamatsu (1997) claimed that the strategy of imagining[9] the face of a celebrity (compared to *not* imagining) increased accuracy in subsequent questions about the celebrity's appearance (e.g., whether the celebrity has a facial mole, or long/short hair; Cabeza et al., 1997). Consistent with this, Wu et al. (2012) found that instructing participants to visualise a face facilitated subsequent recognition of that face (compared to a mis-matching face, or a no-visualising condition). Not only did face-visualising aid subsequent recognition of the *same* face, but equally, it interfered with recognition of a *different* face. Taken together, these studies suggest that visual imagery can have a facilitatory role in face recognition.

Another area of research that demonstrates a link between visual imagery and face recognition is the study of people with *congenital prosopagnosia* (henceforth 'prosopagnosia'). People with prosopagnosia are characterised by poor facial recognition ('face blindness') that has been present from birth (although cases can also be acquired; Behrmann & Avidan, 2005). Previous research has found that people with prosopagnosia often experience visual imagery impairments, but particularly tied to the imagery of faces. For example, using a modified version of the VVIQ in which half the items were of faces, prosopagnosics reported significantly weaker visual imagery than controls (Grueter et al., 2007; Grüter et al., 2009). However, when looking at participant responses in more detail, prosopagnosics tended to report significantly weaker imagery for faces than for objects, which suggests that any impairments in visual imagery may be particularly pronounced for faces.

[9] We point out here that imagination and imagery are related, but they are not the same thing (Dance, Jaquiery et al., 2021). The instruction to 'imagine' celebrity faces in Cabeza et al. (1997) would likely have involved *visual imagery* in the majority of participants (since 96.1% of the general population are imagers; Dance, Ipser & Simner, 2022). However, it is still possible to imagine faces in the *absence* of visual imagery (which would have been the approach of the remaining 3.9% of their participants, i.e., the aphantasics).

A similar finding was presented by Tree and Wilkie (2010), whose prosopagnosic cases were impaired in what the authors described as tests of face imagery (e.g., deciding which two of three celebrities look most alike), but not object imagery (e.g., deciding which is larger, a goat or a deer), or colour imagery (e.g., comparing which is darker, an aubergine or plum; Tree & Wilkie, 2010). Naming this as an 'imagery' task may be potentially misleading because imagery is not required to mentally compare the visual features of objects (i.e., people without visual imagery can make similar judgements; Milton et al., 2021; Palermo et al., 2022). Nevertheless, prosopagnosics showed some deficits in this 'face imagery' test. Taken together with their self-reported imagery deficits (see above), this suggests prosopagnosia may be tied to some level of impairment in visual imagery. That being said, we note that there is considerable variance in the strength of imagery reported by prosopagnosics, with some reporting very strong imagery indeed. One prosopagnosic participant in the study by Grüter et al. (2009) reported "perfect" mental imagery, and suggested she could not recognise faces because "they do not always look like I imagine them" (Grüter et al., 2009, pp. 137)[10]. As such, although research into prosopagnosia may propose links between visual imagery and face recognition abilities, the relationship is not always clear cut.

Our brief literature review highlights that the relationship between visual imagery and face recognition remains somewhat unclear, with no direct evidence of any link in low versus high imagers in McKelvie's (1994) study, but some evidence that imaging faces (or at the very least *imagining*/thinking about them) improves face-recognition, while being a difficult task for prosopagnosics. In the present study, we seek to build on this literature to provide clarity on the relationship between visual imagery and face recognition. We do this by taking the approach of examining people who experience a near or total absence of visual imagery, known as *aphantasia* (Zeman et al., 2015).

Around 3.9% of the population experience aphantasia (Dance, Ipser, & Simner, 2022), a term meaning that their visual imagery is either entirely absent, or fleeting/vague/dim (giving them a score of 32 or less in the VVIQ; Dance, Jaquiery, et al., 2021; Keogh & Pearson, 2018; Zeman et al., 2015). We have a growing understanding of their

[10] As we note above, imagery and imagination are related, but they are *not* the same thing (Dance, Jaquiery et al., 2021). Here, we interpret the word 'imagine' in this quote to indicate intact *visual imagery* in prosopagnosia because the case-study was asked about imagery in particular.

characteristics (e.g., Dance, Ward, et al., 2021; Dawes et al., 2020; Keogh et al., 2021; Königsmark et al., 2021; Monzel, Keidel, et al., 2021; Pounder et al., 2022; Wicken et al., 2021) and comorbidities (e.g., Dance, Jaquiery, et al., 2021; Milton et al., 2021), and a small number of studies have now investigated their face recognition abilities. Zeman et al. (2020) presented a single self-report question ("Do you have any difficulty in recognising faces?"), and found that aphantasics reported more issues with face recognition than people with average or strong visual imagery. Similarly, Milton et al. (2021) found that aphantasics reported higher prosopagnosic traits using the *Prosopagnosia Index* questionnaire (*PI20*; Shah, Gaule, Sowden, Bird, & Cook, 2015), again relative to average and strong imagers. Outside of this self-report, however, their face recognition abilities appeared to be curiously unaffected. Aphantasics scored no differently to controls in the *Famous Face Test* (where participants identify famous people such as Barack Obama from four similar faces; Milton et al., 2021), and the *Warrington Recognition Memory Test for Faces* (where participants identified a target next to a distractor, where targets had been shown earlier; Wechsler, 1997). In both cases, there were no differences between aphantasics and average/strong imagers on either task. This might suggest that self-reported face recognition deficits are unreliable, though crucially, the faces used in both tests included features external to the face (e.g., hairstyle), and paraphernalia (e.g., non-facial information such as clothing). This is of crucial importance, since face recognition generally relies on the ability to recognise faces from their *internal* features (eyes, nose, mouth, etc.) (Duchaine & Nakayama, 2006). Using external features/paraphernalia can aid recognition, meaning participants could perform well even if they had no face recognition abilities at all. Indeed, even prosopagnosics perform in the 'normal' range on the Warrington Recognition Memory Test when internal facial features are blurred out, leaving only external features and clothing visible (Duchaine & Weidenfeld, 2003). We are therefore left with the unanswered question of whether face recognition is truly impaired in aphantasia. Here, we endeavour to answer this question, and in doing so we hope to enhance our understanding of the relationship between visual imagery and face processing more generally.

In light of these measure limitations, it is important to consider what kinds of tasks might reliably measure face recognition. In the present study, we selected as our face recognition measure the widely-used *Cambridge Face Memory Test* (*CFMT*; Duchaine & Nakayama, 2006). In this computerised task, individuals are asked to memorise and

then recognise six male faces, in three blocks that increase in difficulty (see Methods). Importantly, faces exclude hair and paraphernalia to ensure performance taps recognition from internal facial features explicitly. The test has been validated in that prosopagnosics tend to perform poorly on the CFMT (Duchaine & Nakayama, 2006). Accuracy on the CFMT also correlates with scores on the Prosopagnosia Index (Gray, Bird, & Cook, 2017; Tsantani, Vestner, & Cook, 2021), meaning that self-reported face recognition maps on to this behavioural task. If visual imagery is indeed linked to face recognition, we would expect individuals with aphantasia to not only self-report more prosopagnosic traits on the Prosopagnosia Index (replicating Milton et al., 2021), but we may also expect them to perform less accurately in the CFMT, relative to people with intact visual imagery (henceforth 'imaging controls').

Although measures such as the CFMT and Prosopagnosia Index are useful in measuring face recognition abilities in general, they are perhaps less informative in testing how people perform in face recognition within forensic settings. In these settings, witnesses often use face composite systems to construct images of criminal suspects, using specialised computer software controlled by a trained composite operator (Fodarella, Kuivaniemi-Smith, Gawrylowicz, & Frowd, 2015). Composites are then shared with the wider police force and community in the hope that someone will identify the suspect (Frowd et al., 2019; Home Office, 2017). Face composite systems therefore provide a valuable method of testing the ability of an individual to reconstruct an image of a potential suspect. Important to our purposes, we are unaware of any research to date that has examined how individual differences in visual imagery may affect face composite generation, so we ask this question here for the very first time.

How might we expect aphantasics to perform in constructing face composites? Given their self-reported difficulties with face recognition (Milton et al., 2021; Zeman et al., 2020), we might expect aphantasics to produce less accurate composites of a face from memory compared to imaging controls. However, we also have reason to believe that aphantasics may *not* be impaired in their construction of composites. This is because previous research involving drawing from memory, a task qualitatively similar to constructing facial composites, shows that aphantasics are generally just as accurate as imaging controls in re-constructing the details of objects (Bainbridge et al., 2021). In this study, aphantasics and imaging controls saw a series of scenes (e.g., a kitchen, a

bedroom), then later drew them from memory. Aphantasics were just as accurate in drawing the details of objects (such as shape, material, and texture), as well as their size and location (Bainbridge et al., 2021). However, aphantasics were less accurate in the *number* of objects recalled (i.e., they were more likely to omit objects, such as an ornament in a bedroom) so their drawings contained significantly fewer objects, and they also used less colour. Nonetheless, given that faces tend to contain the *same number of objects* (i.e., they nearly always include a nose, mouth, eyes, etc.), we suggest that aphantasics constructing face composites are unlikely to be at a disadvantage from their tendency to reproduce fewer objects. Similarly, within the composite task participants will be asked whether they wish to change the colour of features (e.g., the eyes), which may prompt both groups to use a similar level of colour. Instead, their intact abilities to reconstruct the details, appearance, location, and size of objects may allow aphantasics to produce face composites that are just as accurate as those produced by imagers. Alternatively, they may even produce *more accurate* composites, since imagery has been previously linked to recalling the details of memories incorrectly, otherwise known as *false memories* (e.g., Dobson & Markham, 1993; Hyman & Pentland, 1996; Tomes & Katz, 1997). Indeed, Bainbridge et al. (2021) found that aphantasics made fewer false memories in their drawings than imaging controls, meaning they were less likely to draw objects that were not present in the original scene. Taken together, these findings suggest that aphantasic people may be impaired in producing face composites (given their self-reported face difficulties) or unimpaired (given their ability to re-construct details of objects, and their reduced tendency to produce false memories).

Up until now we have been describing face composite creation generally, but strictly speaking there are two main methods for composite construction, both of which we will use in the present study: one *featural*, and the other *holistic*. Featural systems allow the witness to select individual features (e.g., eyes, mouth, nose, etc.), whereas holistic systems require the witness to recognise *overall* facial similarity, where features are processed together in configuration (for a summary, see Fodarella et al., 2015; Zahradnikova, Duchovicova, & Schreiber, 2018). Holistic face composite systems (e.g., EVO-fit, and EFIT6; Frowd et al., 2013; Vision Metric Ltd, 2019a) are commonly used in forensic settings because they tend to produce more accurate composites than earlier featural systems such as Photofit (Frowd et al., 2015, 2019; Pike, Brace, Turner, Ness, & Vredeveldt, 2019), and the face recognition literature has consistently shown that faces

tend to be recognised holistically (Collishaw & Hole, 2000; Farah, Wilson, Drain, & Tanaka, 1998; Richler & Gauthier, 2014; Young, Hellawell, & Hay, 1987). However, there have been instances where featural systems have produced composites of equal, if not better resemblance than holistic systems (Davis, Gibson, & Solomon, 2014; Davis, Sulley, Solomon, & Gibson, 2010), suggesting that there may be variance across people in the type of construction method that produces the most accurate composite. People with aphantasia may perform better with the featural method because these systems involve the user selecting verbal categories to reveal suitable features (e.g., selecting labels to describe the shape of eyes, which then reveals appropriate options). Aphantasics have reported using verbal labels in certain tasks (e.g., to remember visual patterns or objects, Bainbridge et al., 2021; Keogh et al., 2021) and so may excel with the featural system. Alternatively, featural compositions may better suit people with visual imagery, since imagery of faces tends to better depict facial features, as opposed to holistic representations (Lobmaier & Mast, 2008). Either way, we will examine whether individual differences in visual imagery influence not only overall composite abilities, but the effectiveness of each procedure (holistic vs. featural). In doing this, we ask for the very first time how individual differences in mental imagery influence face composite construction.

In sum, we present the first in-depth study of face recognition and face composite construction in aphantasia. If visual imagery facilitates face recognition, then we expect to replicate findings that aphantasics self-report more prosopagnosic traits than imaging controls (Milton et al., 2021), and we also expect aphantasics to perform less accurately in behavioural tasks of face recognition, crucially when there are no clues from external features or paraphernalia (e.g., hairstyle and clothes). Similarly, if visual imagery aids the construction of facial composites, then we expect aphantasics to create less accurate (featural or holistic) composites, though we have also noted several reason above why they may be no different to imaging controls in this task.

Methods

In our study we were mindful to factor out confounding influences. Issues with face recognition are particularly characteristic of people with autism spectrum conditions (henceforth ‘autism’) (Griffin, Bauer, & Scherf, 2021), and higher autistic traits are also found in people with aphantasia (Dance, Jaquiery, et al., 2021; Milton et al., 2021). Given

that these links could potentially tie aphantasia to face recognition difficulties via autism alone, a subset of our participants were screened using the *Autism Spectrum Quotient* (*AQ*; Baron-Cohen et al., 2001), so that we could examine whether autistic traits influenced our results.

A second methodological consideration relates to how to assess the accuracy of facial composites. General practice is to employ 'rating participants', that is, a general population sample to rate the resemblance between each composite and its target face. Since it is at least theoretically possible that general population raters (of whom approximately 96% will have imagery; Dance et al., 2022) may somehow disadvantage aphantasics composite-makers, we engaged two groups of raters: one with, and one without aphantasia. In other words, aphantasics/imaging raters will assess aphantasic/imaging composite-makers (i.e., four groups in a fully crossed design). We may find that aphantasic raters appreciate better the composites made by aphantasics. Additionally, we may certainly find that raters with aphantasia provide lower ratings overall (for all composites) than those with imagery, simply because aphantasics' self-reported face recognition difficulties might make it more difficult for them to see the resemblance between a target face and its composite. We explore these ideas in more detail in our Results, and General Discussion below.

Participants

'Testing participants' were our main cohort in our face recognition and composite tasks below, while 'Rating participants' rated the faces created by the testing participants in the composite task. These groups are described in turn below. There was no mention of face recognition or facial composites in any recruitment materials (for either testing or rating participants) to ensure random sampling (i.e., so individuals with atypical face perception/memory were not disproportionately encouraged to take part). All participants provided informed consent prior to participating, and ethical approval came from the *University of Sussex Cross-Schools Science and Technology Ethics Board*.

Testing participants. Our participants were 52 aphantasics (38 female, 14 male; *M* age = 42.25, *SD* = 16.29) and 40 controls with intact visual imagery abilities ('imaging controls'; 30 female, 10 male; *M* age = 41.23, *SD* = 16.66). Aphantasic participants were recruited from the University of Sussex's *Imagery Lab - Aphantasia Cohort,* and had

already been screened for their visual imagery abilities using the VVIQ (described below) when joining our participant database. In line with previous research indicating that aphantasics rate their visual imagery as absent or vague/dim (Dance et al., 2022; Keogh & Pearson, 2018; Zeman et al., 2015), aphantasics were classified as having scores of 16-32 on the VVIQ (Marks, 1973; *M* = 17.29, *SD* = 3.36), where a score of 16 represents completely absent imagery across all items, and a score of 32 represents dim/vague imagery (with incremental scores in between). Control participants were recruited from the University of Sussex's *Imagery Lab - Control Cohort* and were imaging individuals who had previously signed up to our participant database, with scores of >32 on the VVIQ (*M* = 67.75, *SD* = 11.38). Age and gender were balanced groupwise, and participants were compensated £10 for our 1.5 hour study. One additional aphantasic individual began a testing session, but the session was terminated after technical issues with our face composite software (EFIT6; Vision Metric Ltd, 2019a). This participant was not included in our analyses (which included n52 aphantasics only).

Rating participants. The face composites created by our testing participants were rated for resemblance to the target face (see below) by two types of rating participants. Our imaging raters were 73 participants with intact visual imagery (40 male, 32 female, 1 another gender; *M* age = 35.15, *SD* = 9.51) recruited via the online recruitment platform testable.com. Our aphantasics raters were 50 aphantasics (20 male, 29 female, 1 another gender; *M* age = 42.94, *SD* = 14.28), recruited either from testable.com (n20) or the University of Sussex's *Imagery Lab - Aphantasia Cohort* (n30). Given time-constraints with these participants, we classified them as imaging/aphantasic using a shorter in-house imagery screener (described below). As before, aphantasics were classified as those with absent or vague/dim imagery, while imaging raters had intact visual imagery (i.e., their imagery was at least moderately clear and vivid, or better; see below for details). Raters from www.testable.com were paid $9 for our 1 hour long task, and aphantasic raters from our *Imagery Lab – Aphantasia Cohort* were entered into a prize draw worth £25.

Materials and Procedure

Participants completed the study either in-person (e.g., at Sussex University, using a HP 15.6" laptop and mouse), or remotely via a Skype Business audio call (i.e., from their

own homes, using their own computer/laptop device)[11]. Of our 52 aphantasics, 27 (20 female, 7 male; *M* age = 42.56, *SD* = 15.87) completed the study in-person, and the remaining 25 (18 female, 7 male; *M* age = 41.92, *SD* = 17.05) completed the study online (remotely via a Skype Business audio call; see below). Of our 42 control participants, 15 (11 female, 4 male; *M* age = 41.33, *SD* = 17.79) completed the experiment in-person, and the remaining 25 (19 female, 6 male; *M* age = 41.16, *SD* = 16.31) completed the study online. The same experimenter (author CD) conducted all testing sessions to ensure consistency in the administration of the tasks. For the in-person testing sessions, the experimenter brought up each task on screen for the participant to complete, and sat next to the participant when constructing the face composites (so that the experimenter could control the EFIT6 software; see below). During the remote (online) testing sessions, the experimenter sent the participant URL links to the tasks via the chat function in Skype Business. To construct the face composites during online sessions, the experimenter shared their view of their screen with the participant (using the 'share screen' function in Skype; again, this was so the experimenter could control the EFIT6 software; see below).

Participants completed the tasks below, in the order shown (with the VVIQ completed in advance of the main session, as described above). Unless otherwise stated, tasks were administered via our in-house testing platform www.syntoolkit.com. The first five tasks below were completed by our main cohort of testing participants. The final (rating) task was performed by our rating participants.

Vividness of Visual Imagery Questionnaire (VVIQ; Marks, 1973). In the VVIQ, participants were asked to imagine a series of four scenes (e.g., "A beach by the ocean on a warm summer's day"), and rate the vividness of their visual imagery for four aspects of each scene (e.g., "The appearance and colour of the water"). Visual imagery vividness was rated on a five-point scale: 1 ("No image at all, you only "know" that you are thinking of the object"), 2 ("Vague and dim"), 3 ("Moderately clear and vivid"), 4 ("Clear and reasonably vivid"), and 5 ("Perfectly clear and as vivid as normal vision"). Responses were summed giving possible scores in the range of 16-80. As noted above, participants

[11] This decision was dictated by the sudden onset of lockdown during the early months of the Covid-19 pandemic. However, we took steps in our analysis to verify that this in person/online testing made no difference whatsoever to the pattern of our findings (see Results).

were categorised as aphantasic if they scored between 16-32 (indicating absent or vague/dim imagery) with imaging controls scoring >32.

Face composite construction. In the face composite task, participants saw a single target face (a middle aged male; see Figure 2), and were required to later create two composites of that face (one holistic, one featural; see details below) using EFIT6 software (Visionmetric Ltd, 2019a). EFIT6 (formerly EFIT-V) is a particularly useful face composite system because it provides both holistic and featural methods of construction (Visionmetric Ltd, 2019a). To begin, the target face was presented in a short video, in which the target turned his face from the centre/staring straight at the camera, to the left, and then to the right (repeated twice; presented at 1280 × 720 pixels). Participants were instructed to keep their attention on the face and remember as much information as they could about it. For in-person testing sessions, the experimenter opened the video on the computer screen and pressed play. For online (remote) testing sessions, participants were sent a URL link to the video (hosted on qualtrics.com) via the chat function in Skype Business. To ensure those participating in the study remotely did not have extended access to the target face, the video URL link could only be accessed once, and the video could only be played once and could not be paused.

After seeing the video of the target face, participants were required to create two composites of the face. But in order to create a time-gap between viewing and composite creation, participants first switched to a filler task immediately after the video (the CFMT; described below). After completing this filler task, participants were reminded of the earlier video and were instructed to create a composite of the face they had seen in it. To prompt their memory of the face, participants were first asked to describe it in as much detail as possible. This recall phase is akin to the Cognitive Interview that witnesses undergo before constructing a composite in real-world forensic settings (Brace, Pike, & Turner, 2008; Fodarella et al., 2015). When participants finished recalling the face verbally, they were prompted to add any more verbal details that they could. This recall phase ended when the participant could recall no more information. Following this, the composite task began. Here, the researcher opened EFIT6 (sharing their screen if the testing session was online/remote) and asked the participant to start by identifying the target's gender, ethnicity and age range from drop-down lists available on screen. Working with the experimenter who controlled the software, the participants then

constructed composites of the face, using a holistic method and a featural method (all participants created both types of composite in a counter balanced order).

Holistic method. In the holistic method, participants were asked to consider holistic (overall) facial similarity, instead of identifying individual face features (Vision Metric Ltd, 2019b). Participants saw a set of faces, which were all variants of the same face. The participant selected the variant that most resembled the target face. A new set of variants then appeared, all based on this selection. This process of "evolution" continued until the participant decided that a good match had been found. To ensure that the first group of faces had broadly suitable attributes, participants first selected the approximate overall face-shape from 9 options presented on screen (in a 3x3 grid). Participants then did this for the overall shape of the nose, lips, eyes, and eyebrows (at this point the features lack detail, but are instead 'prototypes' for the overall shape of each feature). Next, the participant selected a hairstyle. To filter appropriate styles, the participant could select the hair length, style, and type from drop down lists available on screen (which revealed suitable sub-options at the bottom of the screen). Clicking these placed the hairstyle on the face for the participant to view. After participants chose the hairstyle, they then chose the colour using a colour pallet, adjusting lightness and contrast as necessary.

Based on the face attributes and hairstyle selected, participants were then presented with the first group of nine faces (using the 'Multi-face Easy / Evo function) displayed in a 3x3 grid. Participants were asked to select the face that most resembled the target; then, based on its attributes, the software generated another set of nine faces (one being the face selected in the previous round). This process was repeated until the participant identified a face that was as close a match to the target as possible (i.e., each round is designed to 'evolve' more suitable faces using an evolutionary algorithm; Solomon, Gibson, & Maylin, 2012; Solomon, Gibson, & Mist, 2013). Participants could also 'lock in' features seen as a good likeness (e.g., a face shape) so that they appeared in all faces in subsequent rounds. Once a face had been selected, participants could then refine its features using the 'Move' tool to adjust size/location of facial features and hairstyle (Visionmetric Ltd, 2019a, 2019b). Once the participant had completed constructing their first composite, they performed an unrelated 10 minute filler task (involving making decisions about letters, to be reported elsewhere) before constructing their second composite.

Featural method. In the featural method, participants selected face features individually instead of recognising overall facial similarity (Vision Metric Ltd, 2019b). Participants started the task by once again giving a verbal description before proceeding with the composite. Here, the phase of choosing the overall prototype shape for each facial feature was skipped, and participants began by choosing the hairstyle. Participants then selected each face feature individually (face shape, nose, mouth, eyebrows, eyes; using the 'Features – Thumbnails' function; Vision Metric Ltd, 2019a). Each feature had 20 examples to choose from, displayed at the bottom of the screen. The participant could select particular styles (e.g., type/shape), revealing suitable sub-options, and clicking on these placed the feature on the face for the participant to view. Features could again be refined further (in shape and location) using the 'Move' tool (as described above).

Finally, in both types of composite creation (holistic and featural) the ears could be altered (via the 'Ears' function) using the featural method (i.e., ears could be selected from a list of 20 examples). In both types of composite creation, the eye colour, attributes (e.g., age/skin tone/wrinkles), ethnicity, and expressions (e.g., sad/smile) could also be edited, and skin effects (e.g., stubble/eye bags/freckles) added.

Cambridge Face Memory Test (CFMT; Duchaine & Nakayama, 2006). In the CFMT, participants were asked to memorise and then identify 6 male target faces shown in black and white photographs (hosted on testable.com). The test consists of three blocks which increase in difficulty (and these blocks are called 'Same Images', 'Novel Images', and 'Novel Images with Noise') as described in turn below.

In the 'Same Images' block, participants were asked to memorise six target faces. Each face was presented individually on-screen for 9 seconds, shown first left-facing (3 seconds) then centre-facing (3 seconds), then right-facing (3 seconds). Immediately afterwards, the participant was required to identify the target in three subsequent forced-choice trials. Each forced-choice trial showed the target face alongside two other distractor faces. The correct test face was always identical (e.g., in pose and lighting etc.) to the memorised face (hence this is the 'Same Images' block). Responses were made by pressing the 1, 2 or 3 key, corresponding to the left, middle and right face, respectively. This block consisted of 18 trials in total, i.e., 3 trials for each of the 6 target faces (and the sequencing was: Face1 then trials for Face1; Face 2 then trials for Face 2, etc.). The

18 target trials were preceded by three initial practise trials with the same format, but using a cartoon face.

Next was the 'Novel Images' block, which was prefaced by a review page showing frontal shots of the 6 target faces once again, together on screen for 20 seconds. In this 'Novel Images' block, participants were immediately required to recognise the targets in 30 forced-choice trials in a random order (5 trials for each of the 6 target faces), again with three faces per trial (a target plus two distractors). This time however, the test face differed from the learning phase in lighting and/or pose. After this block, participants again saw the review page (the 6 target faces together) for 20 seconds, before completing the 'Novel Images with Noise' block. In this block, participants completed 24 forced-choice trials (4 trials for each of the 6 target faces) but this time, the test faces were both novel (as described above) and had Gaussian noise added. Noise serves to keep performance away from ceiling, and encourages strategies specific to face recognition (i.e., noise encourages holistic recognition based on the whole face, rather than features which become unclear through the noise; Duchaine & Nakayama, 2006; McKone, Martini, & Nakayama, 2001). Since each of the three phases had a different maximum score, we calculated the percentage accuracy (i.e., the percentage correct) for each block, so that we could directly compare performance block by block, as well as overall (Croydon, Pimperton, Ewing, Duchaine, & Pellicano, 2014).

Prosopagnosia Index (Shah, Gaule, et al., 2015). In this 20-item questionnaire, participants were asked to indicate how much they agreed with statements about their face recognition ability (e.g., "My face recognition ability is worse than most people", and "I often mistake people I have met before for strangers"), on a scale of 1 ("Strongly Disagree") to 5 ("Strongly agree"). Five of the items require reverse scoring, and overall scores are calculated by summing responses to all items. In line with standardised scoring, scores ≥65 were considered to be suggestive of the presence of at least mild prosopagnosia (Shah, Gaule, et al., 2015).

The Autism Spectrum Quotient (Baron-Cohen et al., 2001). This 50-item questionnaire measures five domains of autism symptomatology: attention-to-detail (e.g., "I tend to notice details that others do not"), attention switching (e.g., "I prefer to do things the same way over and over again"), communication (e.g., "I frequently find that

I don't know how to keep a conversation going"), social skills (e.g., "I find it hard to make new friends"), and imagination (e.g., "I find it difficult to imagine what it would be like to be someone else"). Participants rated the degree to which they agreed with each item on a 4-point scale ("Definitely agree", "Slightly agree", "Slightly disagree", "Definitely disagree"). Responses that indicate (either definitely or slightly) autistic symptomatology (i.e., good attention-to-detail, but poor attention switching, communication, social skills, and imagination) are scored as 1, but otherwise 0 (and approximately half of the items were reversed scored). Following previous research, we excluded one item that directly asks about aphantasia phenomenology ("I find it very easy to create a picture in my mind"), meaning overall scores could range from 0-49. This measure was only completed by our online (remote) testing participants (n25 aphantasics, and n25 controls; see above).

Face composite ratings. All face composites produced by our testing participants in the Face Composite Task were subsequently rated by a new group of participants (our Rating Participants). Raters judged the composites for their likeness to the target face, and we remind the reader that we had both imaging raters and aphantasics raters (see Participants). Raters were categorised as imagers or aphantasics using our in-house imagery screener (see Appendix B). To ensure any mention of visual imagery did not influence the resemblance ratings, the imagery screener was completed by raters after all of the composites had been rated. In the imagery screener, we asked participants to think about the building where they lived and to judge how much their 'memory of it is like a picture' in their mind. Participants responded on a scale of 0 ("I have no image or picture at all (and I don't know what it looks like)"), 1 ("I have no image or picture at all (but I still know what it looks like)")', 2 ("I have an image that is only vague or fleeting or dim"), 3 ("I have an image that is moderately clear and vivid"), 4 ("I have an image that is clear and reasonably vivid, like a relatively clear mental picture"), and 5 ("I have an image like a picture that is perfectly clear and vivid – just as if I were really standing there"). In line with previous research using this question to screen for aphantasia (Dance, Jaquiery, et al., 2021), our raters were categorised as aphantasic if they responded in the range of 0-2 (i.e., rating their visual imagery as completely absent or vague/dim/fleeting; $M = 1.14$, $SD = 0.57$), whereas imagers scored between 3-5 ($M = 3.82$, $SD = 0.69$).

In the face-rating task, our raters judged the 184 composites generated by our testing cohort. Composites were presented on-screen in a random order alongside the target face. Raters were asked to score the resemblance on a 10-point scale from 0 ("No resemblance") to 10 ("Excellent resemblance"). The face ratings were then used to generate a mean score for each of the composite-generators (i.e., our testing participants, who had generated the composites). There were four main scores for each composite generator: one average resemblance score from each rating group (imagers, aphantasics) for each composite type (featural, holistic).

Results

Since our data approximated normal distributions and homoscedasticity, we use parametric statistical tests throughout our manuscript. For completeness, we also provide robust versions of our analyses in Appendix B (and show that the overall pattern of results remains unchanged). To ensure our face processing outcomes were not affected by study type (in-person, or online), we first ran all of our analyses including study type as an additional between-subjects factor. There were no significant main effects of study type, or interactions (see Appendix B for details, and Table S1 for means and standard deviations). In our main analyses below, we therefore collapse across study type (i.e., comparing aphantasics to controls regardless of whether they were tested in-person, or online). Means and standard deviations for all of our face processing measures are provided in Table 1.

Face Composites

To examine the resemblance scores for face composites created by our aphantasic and control groups we conducted a 2x2x2 ANOVA crossing group (aphantasia, controls) with composite type (holistic, featural), and rating participant (imagers, aphantasics). Our ANOVA revealed no main effect of group, ($F(1, 90) = 1.15$, $p = .286$, $\eta_p^2 = .013$), indicating that composites produced by our aphantasics and controls resembled the target face to a similar degree. There was also no main effect of composite type, ($F(1, 90) = 0.54$, $p = .463$, $\eta_p^2 = .006$), meaning that resemblance ratings were similar for holistic and featural composites (Figure 1).

The ANOVA did, however, reveal a significant effect of rater type, ($F(1, 90) = 170.95$, $p < .001$, $\eta_p^2 = .655$), because aphantasic raters tended to rate the face composites lower in

resemblance to the target face than imaging raters. Finally, the ANOVA revealed no significant interactions (all p >.142; See Figure 1) meaning that aphantasics raters gave lower scores than imaging raters across the board. This suggests that aphantasic raters struggled more than imagers to see the resemblance between a composite and its target face, irrespective of whether composites were holistic or featural composites, and irrespective of whether composites were created by aphantasics or imagers.

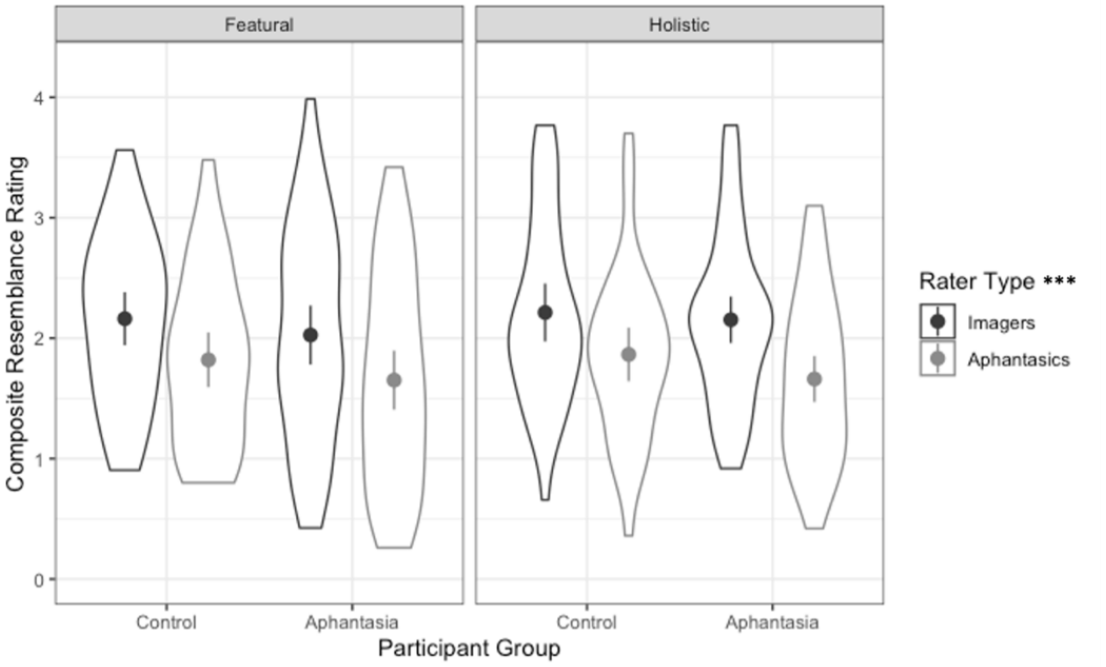

Figure 1. Face composite mean scores and distributions for aphantasics and imaging controls. Results are split by composite type (featural, holistic), and rater-type (aphantasics, imagers). Figure 1 shows that composites constructed by aphantasics and imagers resembled the target face to a similar degree; however, resemblance scores are lower when the *rater* is aphantasic (vs imager), as indicated by *** = p <.001. Figure 1 also shows that featural and holistic composites resembled the target face to a similar degree. Error bars represent 95% confidence intervals.

Table 1 shows the mean resemblance ratings for each group (aphantasics, controls) broken down by composite type (holistic, featural), and rating type (imager, aphantasic). We present these descriptive statistics to highlight that the ratings of composites were relatively low across the board (e.g., the highest group mean was 2.21 (SD = 0.75) out of 10). This is consistent with previous research using EFIT6 (e.g., Bhardwaj & Hole, 2020; Davis et al., 2014, 2010, 2016), and suggests that composites are somewhat limited in how far they accurately resemble target faces (see Figure 2 for examples of the face

composites created by participants). However, raters were in agreement about which composites showed the greatest resemblance (i.e., there was a strong positive correlation between aphantasic and imager resemblance ratings for both holistic and featural composites; respectively, $r(92) = .84$, $p < .001$, and $r(92) = .90$, $p < .001$). There was also a strong association between resemblance ratings for holistic and featural composites produced by any given participant ($r(92) = .50$, $p < .001$; collapsing across imager and aphantasic rating groups), meaning that if a participant was relatively accurate using one composite method, they also tended to be relatively accurate using the other method.

Table 1

Means and standard deviations for aphantasics and imaging controls for the Face Composite Task, the Cambridge Face Memory Test (CFMT), and the Prosopagnosia Index (collapsed across study type; in-person, and online). The Face Composite Task scores are broken down by rater type (aphantasics, and imagers).

	Group (composite-constructors)			
Measure	Aphantasia (n52)		Control (n40)	
	Mean	SD	Mean	SD
Composites				
Imager ratings				
Holistic	2.15	0.69	2.21	0.75
Featural	2.03	0.88	2.16	0.69
Aphantasia ratings				
Holistic	1.66	0.69	1.87	0.69
Featural	1.65	0.88	1.82	0.71
CFMT (%)				
Overall	69.63	15.86	77.64	12.57
Same Images	95.83	6.59	98.75	2.95
Novel Images	63.33	21.17	75.92	17.58
Noise Images	57.85	20.97	63.96	17.87
Prosopagnosia Index	58.71	16.38	38.98	9.85

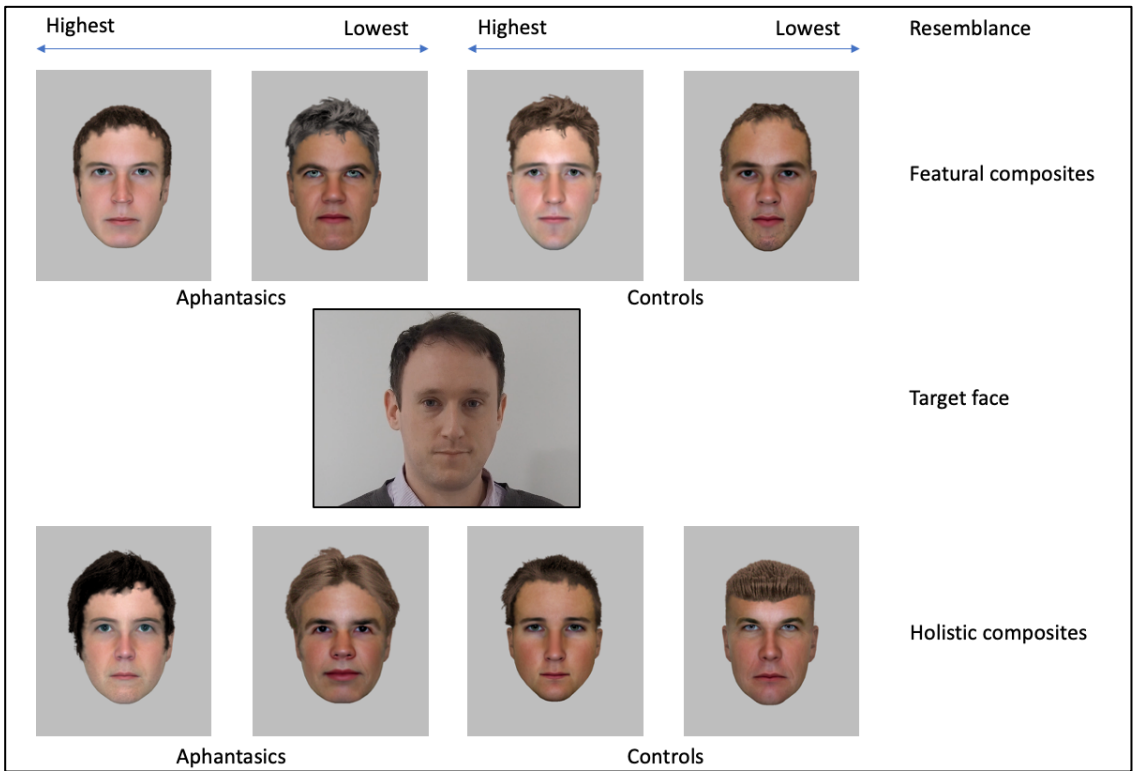

Figure 2. Example composite faces created by aphantasics and imaging controls using the holistic and featural methods in EFIT6 (Visionmetric Ltd, 2019a). We present the highest and lowest rated faces (in terms of resemblance to the target face), where ratings are collapsed across rating groups (aphantasics, and imagers).

Cambridge Face Memory Test

Using the CFMT as our behavioural test of face recognition, we conducted a 2x3 ANOVA crossing group (aphantasia, control) with CFMT block (Same Images, Novel Images, Novel Images with Noise) where each block was measured by the percentage correct. The ANOVA revealed a significant main effect of group, ($F(1, 90) = 6.75$, $p = .011$, $\eta_p^2 = .070$), driven by aphantasics performing worse on the task overall compared to controls (see Figure 3). There was also a significant main effect of CFMT block, ($F(1.71, 153.59) = 231.30$, $p < .001$, $\eta_p^2 = .720$; with Greenhouse-Geisser correction), because accuracy decreased with each block type (as expected, since blocks increased in difficulty), and a significant interaction between block type and group ($F(1.71, 153.59) = 3.89$, $p = .029$, $\eta_p^2 = .041$; with Greenhouse-Geisser correction).

To examine the significant interaction, we ran three follow-up independent samples t-tests comparing the percentage accuracy in each CFMT block within our aphantasic and control groups. Aphantasics performed significantly worse than controls in the Same

Images block, (t(74.39) = 2.84, p = .006, g = 0.54, 95% CI [.873, 4.96]; with Welch correction), and also in the Novel Images block, (t(90) = 3.04, p = .003, g = 0.63, 95% CI [4.35, 20.81]). There was, however, no significant groupwise difference in the Novel Images with Noise block, (t(90) = 1.48, p = .144, g = 0.31, 95% CI [-2.12, 14.33]; see Figure 3). Both differences here survive correction for multiple comparisons using the Benjamini-Hochberg False Discovery Rate method (FDR; we use this correction method here, and throughout our manuscript; Benjamini & Hochberg, 1995, 2000). In summary, these results show that imaging controls were more accurate than aphantasics in the CFMT overall, and in the blocks of lowest (Same Images) and medium (Novel Images) difficulty. Aphantasics did not differ from controls in the hardest block (Novel Images with Noise), although they were trending in the same direction.

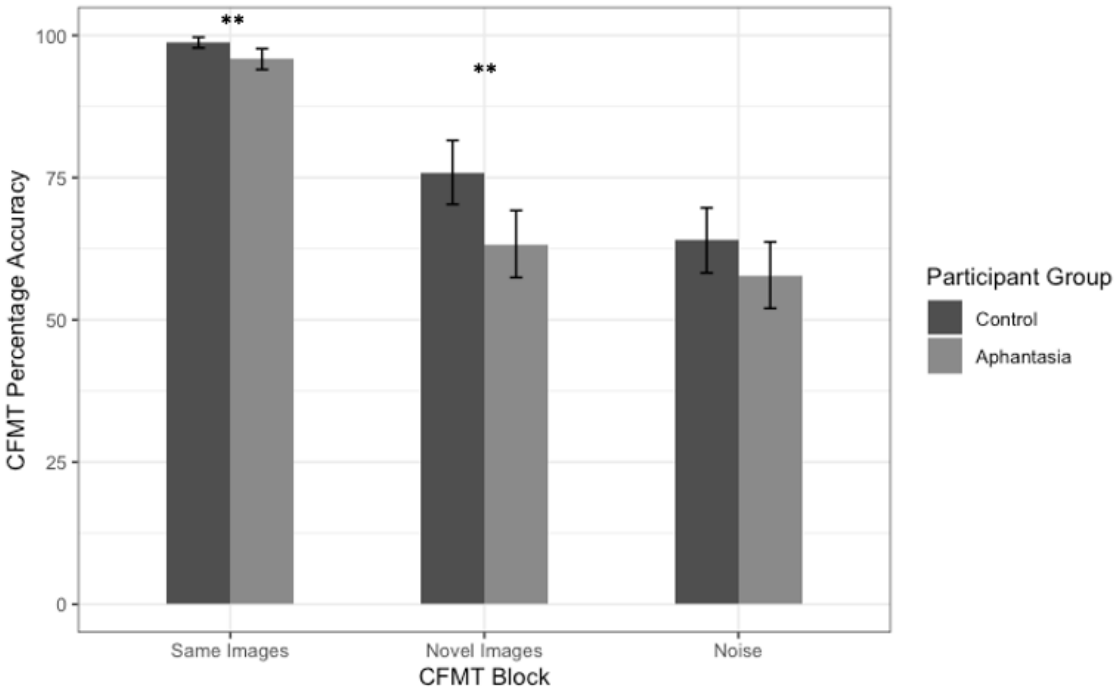

Figure 3. Cambridge Face Memory Test means (with 95% confidence intervals) in aphantasics and controls, as a function of block type (same images, novel images, novel images with noise). *Note.* * uncorrected p <.05, ** p <.01, *** p <.001.

Prosopagnosia Index

Using the Prosopagnosia Index as our measure of self-reported face recognition difficulties, a one-way ANOVA revealed a significant main effect of group, (F(1, 90) = 45.36, p <.001, η_p^2 = .335), indicating that aphantasics reported significantly weaker face

recognition than controls (i.e., aphantasics had higher Prosopagnosia Index scores; see Figure 4).

Since Prosopagnosia Index scores ≥65 indicate the possible presence of at least mild prosopagnosia (Shah, Gaule, et al., 2015), next we investigated whether there was a difference in the number of people reporting scores within this range in the aphantasic and imager groups. There were 22 people (42.3%) in our aphantasic group with an overall Prosopagnosia Index score ≥65, and 1 individual (2.5%) within our control group. This difference was highly significant, [χ^2 (1, N = 92) = 17.04, p <.001; chi-square test with Yates continuity correction].

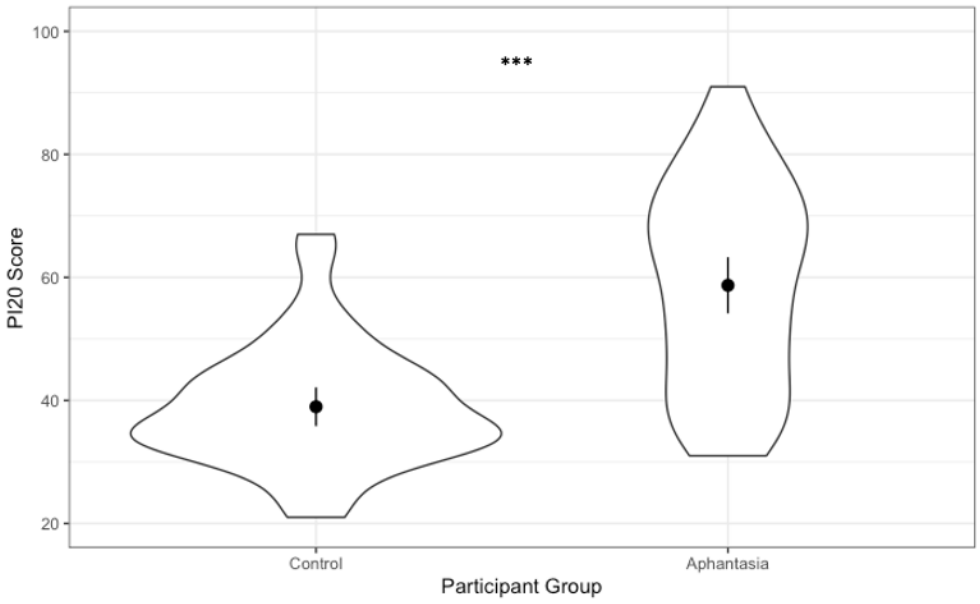

Figure 4. Prosopagnosia Index score means and distributions in aphantasics and controls (with 95% confidence intervals). *Note.* * p <.05, ** p <.01, *** p <.001.

Autism Traits

Finally, since autistic traits have been linked independently to both face recognition difficulties (Griffin et al., 2021), and aphantasia (Dance, Jaquiery, et al., 2021; Milton et al., 2021), we examined AQ scores in our participant groups to be sure that face processing differences were not driven by autism (using data from 25 aphantasics and 25 controls; see Methods). As expected, aphantasics had significantly higher AQ scores (M = 24.60, SD = 8.67) than imaging controls (M = 15.84, SD = 4.77; t(37.31) = 4.43, p <.001, g = 1.23; 95% CI [-12.77, -4.75]; with Welch correction). However, AQ scores

did not significantly correlate with any of our behavioural face processing measures, before or after correcting for multiple comparisons (CFMT, holistic composite score, featural composite score; all |*r*s| <.21, all uncorrected *p*s>.14; composite scores were collapsed across both imaging and aphantasic rating groups), or our self-report measure, the Prosopagnosia Index (The Prosopagnosia Index; *r*= .27, uncorrected *p*= .06). In sum, our AQ investigation suggests that autistic traits do not seem to be robustly linked to the outcomes of our face processing measures.

General Discussion

Our study provides the first in-depth (self-report and behavioural) investigation of face composite creation and face recognition (excluding external features and paraphernalia) in people with and without visual imagery. Our intention was to examine the relationship between visual imagery and face processing, and in doing so, further our understanding of aphantasia, as well as the broader context of forensic eyewitness testimony. We found that compared to people with visual imagery, aphantasic individuals reported more traits associated with prosopagnosia (using the *Prosopagnosia Index* questionnaire; Shah, Gaule, et al., 2015), indicative of difficulties with face recognition. This finding replicates previous research showing higher levels of self-reported face recognition impairments in aphantasia (Milton et al., 2021; Zeman et al., 2020; but see also Palermo et al., 2022). We also found significantly more aphantasics (42.3%) than controls (2.5%) with scores over the threshold of 65 on the *Prosopagnosia Index*, suggestive of the presence of at least mild prosopagnosia (Shah, Gaule, et al., 2015). Taken together, these results suggest that aphantasics report more issues with face recognition during day-to-day life than people with visual imagery.

We also found, for the very first time, that self-report findings of face difficulties in aphantasics and imaging controls mapped onto a standardised behavioural test for face recognition, namely the *Cambridge Face Memory Test* (*CFMT*; Duchaine & Nakayama, 2006). Here, aphantasics were less accurate overall at recognising a series of target faces. Interestingly, the results on this behavioural task challenge previous research by Milton et al. (2021) which showed no difference between aphantasics and imaging controls on two different face recognition tasks: the *Famous Face Test* (Milton et al., 2021), and the *Warrington Recognition Memory Test for Faces* (Wechsler, 1997). Likewise, our results contradict previous research showing no difference between people low and high in visual

imagery in the ability to remember and recognise target faces from distractor faces (McKelvie, 1994). We put this disparity down to the fact that the tests used by Milton et al. (2021) and Mckelvie (1994) may not reliably measure face recognition *per se*. This is because the faces used in their tasks included external features (e.g., hairstyle) and paraphernalia (e.g., clothing), known to facilitate recognition even in the absence of internal facial features (e.g., eyes, nose, and mouth; Duchaine & Weidenfeld, 2003). Given that the CFMT was designed to measure face recognition explicitly (i.e., the face stimuli exclude external features and paraphernalia), and that scores on this task accurately predict the presence of prosopagnosia (Duchaine & Weidenfeld, 2003; Tsantani et al., 2021), we suggest that our test provides a more sensitive index of face recognition. As such, our findings provide the very first behavioural support for aphantasia being linked to impairments in face recognition.

More broadly, our results provide evidence that visual imagery, as an individual difference, may facilitate face recognition. One interpretation of this link is that being able to 'see' a face in the 'mind's eye' may reinforce in some way the memory for the face, aiding its later recognition. This is consistent with findings from previous research which show that imagining or visualising a face can aid its subsequent recognition (Cabeza et al., 1997; Wu et al., 2012). However, we also found that the facilitatory effects of imagery somewhat diminished when the face recognition demands became very difficult. In the most difficult phase of the CFMT (block 3, with visual noise), the numerical difference between our groups was no longer significant, although the average accuracy for the aphantasic group was still lower than that of controls. The visual noise added to faces in Block 3 makes the facial features (eyes, nose, and mouth) unclear, forcing recognition from the overall holistic configuration of the face (Duchaine & Nakayama, 2006). Since face imagery has been shown to better represent the *features* of a face, as opposed to its holistic configuration (Lobmaier & Mast, 2008), the visual noise may interfere with facial recognition for our imaging controls, meaning they no longer convincingly outperform aphantasics in this final block. Nevertheless, the fact that aphantasics perform less accurately in the first two phases of the task, and overall, shows that visual imagery has – at least to some degree – a facilitatory role in face recognition.

Our findings also converge with findings from the prosopagnosia literature. Here it has been shown that people with congenital prosopagnosia (i.e., poor face recognition from

birth; Behrmann & Avidan, 2005) tend to self-report weaker visual imagery than people with intact facial recognition (Grueter et al., 2007; Grüter et al., 2009), often with particular deficits in imagery for *faces* over other objects (Tree & Wilkie, 2010). It is important to make clear, however, that although prosopagnosics as a group tend to report weaker visual imagery, there is nevertheless variation in the imagery levels reported by prosopagnosic individuals (Grueter et al., 2007; Grüter et al., 2009; Tree & Wilkie, 2010). This pattern is mirrored in our own results – although as a group aphantasics had significantly weaker face recognition abilities than imaging controls, there were some aphantasics with very good face recognition (e.g., one aphantasic got every item correct on the CFMT). It seems, therefore, that aphantasia may *often* be characterised by impaired face recognition, but these traits do *not always* co-occur. Variation such as this is to be expected, since neurodevelopmental conditions do not tend to manifest uniformly across all individuals.

We also note here that there are other noticeable parallels between prosopagnosia and aphantasia. In both congenital conditions, individuals tend to live 'normal' lives, often without realising they are in any way different until later in life (Grueter et al., 2007; Zeman et al., 2015, 2016, 2020). Effective compensatory strategies may give rise to this, such as using external facial features, gait, and voice, to accurately recognise others in cases where facial recognition is impaired (Grueter et al., 2007), or using non-imagery strategies in cases where imagery is impaired (Keogh et al., 2021). Investigating the different subtypes of aphantasia (i.e., those with, and without face recognition difficulties), and potential compensatory mechanisms, will be an interesting avenue for future research.

Our study was also the first to examine how people with and without visual imagery perform in creating facial composites, a task commonly used in forensic settings (Frowd et al., 2019; Home Office, 2017). We found that people with aphantasia were just as accurate as people with visual imagery at creating composites of a face from memory. So why might this be, given their face recognition difficulties? We noted in our Introduction that people with aphantasia are generally unimpaired when reconstructing object details of scenes from memory (e.g., shape, appearance, texture, size and location; Bainbridge et al., 2021). Although they tend to draw fewer objects from memory overall than imagers, and use less colour, these problems should not transfer to the composite task because it is

unlikely participants would omit an object (an eye, nose, mouth etc.; which are preloaded in EFIT6 in any case), and participants are explicitly asked whether they wish to alter the colour of features by the composite operator. Their performance may instead have been aided by their reduced tendency to produce false memories (Bainbridge et al., 2021), and their intact *spatial abilities* (Bainbridge et al., 2021; Dawes et al., 2020, 2022; Keogh & Pearson, 2018; Palermo et al., 2022; Pounder et al., 2022) which may have helped to ensure features were appropriately sized and placed accurately within the broader spatial configuration of the face. Overall, our results therefore show that people with, and without visual imagery, are able to produce equally accurate facial composites.

Our results also showed no difference in the accuracy of EFIT6 composites produced by featural or holistic methods, and this was true for both aphantasics and imaging controls. We also found that if a participant had a relatively good score for their holistic composite, they also tended to have a relatively good score for their featural composite (i.e., composite scores for each method were positively correlated). Interestingly, this is at odds with previous research which tends to show greater accuracy for holistic systems (Frowd et al., 2015, 2019; Pike, Brace, Turner, Ness, & Vredeveldt, 2019; but see also Davis, Gibson, & Solomon, 2014; Davis, Sulley, Solomon, & Gibson, 2010). What is novel about our study, however, is that we have been the first to compare featural and holistic composites produced *within the same composite system*. Previous research has tended to compare featural and holistic composites produced by *different* systems altogether (e.g., EVOfit vs PRO-fit, and EFIT-V vs E-FIT; Davis et al., 2014, 2010; Frowd et al., 2015, 2019; Pike et al., 2019). Importantly, during the featural procedure in EFIT6, individual facial features are chosen whilst the overall face is visible on the screen. This differs from earlier featural systems (e.g., PRO-fit), where the features are selected in isolation before the overall face is revealed. In this way, the featural method provided by EFIT6 may be a kind of 'hybrid' procedure, utilising both featural and holistic processes, since the witness reviews the *overall* face holistically whilst choosing *individual features*. This may make the featural method in EFIT6 particularly effective relative to other featural systems.

What implications can we draw from our research about composite creation in real-world forensic settings? Our findings suggest that face composites produced by people with and without visual imagery would be equally as accurate, as well as those produced via

holistic and featural methods in EFIT6. Additionally, our findings speak to whether constructing composites remotely can be effective. We showed no difference between composites produced by participants in-person, and online via a screen-share in Skype Business. This suggests that in real-world criminal investigations remote construction may be an effective option for creating composites of criminal offenders. This would perhaps be a particularly useful approach in cases where an eyewitness cannot easily travel to the relevant destination (e.g., a police station) to construct the composite in-person.

Another noteworthy finding was that our aphantasic *raters* provided lower resemblance scores than imaging raters, and this was true for both types of composite (holistic or featural) produced by both groups of testing participants (aphantasics, and imaging controls). This was true even though both rating groups were in agreement about which were accurate, and not so accurate composites (i.e., scores from each rating group were positively correlated). This finding means that aphantasic raters were less able to detect the similarities between the composite and target face (i.e., not as well as imagers). Here, we appear to be witnessing a profound novel effect which is that aphantasics are not only poorer in face recognition (as seen from our behavioural and questionnaire face recognition tasks) but they may also be poorer in face *perception*. Specifically, our findings suggest they are poorer at what is known as *face-matching*. Raters saw both faces on-screen at the same time and were not relying on long-term or episodic memory from an earlier-seen face. Instead, they could inspect (i.e., *match*) both faces on-screen simultaneously. Previous research shows that scores on the Prosopagnosia Index predict face-matching abilities (on the Glasgow face-matching test, where individuals indicate whether pairs of faces have the same, or different identities; Burton, White, & McNeill, 2010), highlighting that face- recognition and matching abilities are often linked (Shah, Sowden, Gaule, Catmur, & Bird, 2015). As such, we show for the very first time that aphantasia seems to be characterised by deficits in *both* face recognition, and face perception (matching).

In considering potential limitations of our research, it is important to consider whether another trait may explain the reduced face recognition abilities in aphantasia, other than weak visual imagery. We have already addressed – and rejected – the possibility that our findings are mediated by autism (see above; Dance, Jaquiery, et al., 2021; Griffin et al.,

2021; Milton et al., 2021). Another alternative explanation for our results could be that weak face recognition in aphantasia is driven by a broader recognition problem that is not specific to faces. We suggest that this is unlikely, however, since previous research shows an intact ability to remember and then recognise other classes of objects (e.g., words, scenes, patterns) in aphantasia (Bainbridge et al., 2021; Keogh et al., 2021; Knight et al., 2022; Pounder et al., 2022; but see also Monzel, Vetterlein, et al., (2021) who show that aphantasics may peform worse when the memory demands are very difficult). This leads us to believe that there may be something particularly difficult for aphantasics about recognising faces. Indeed, faces are unique in that they are all very similar, containing the same few core features (Richler & Gauthier, 2014). Recognising faces is therefore inherently difficult, with accurate recognition of an unfamiliar face often only occurring after we have been presented with the face multiple times, often from different angles/viewpoints (e.g., Dowsett, Sandford, & Burton, 2016; Ritchie & Burton, 2017). It is possible that visual imagery may provide an additional route by which individuals are presented with, or are able to 'view' a face. Imagers may therefore be able to "rehearse" a face by means of imagery, in a way that is unavailable to aphantasics. This may serve to enhance the memory for the face in some way (e.g., by enhancing its *memorability*; Needell & Bainbridge, 2021), improving subsequent recognition. Accordingly, due to their lack of visual imagery, aphantasics may therefore be particularly susceptible to difficulties in recognising faces (over and above other objects).

Another potential limitation relates to the methodology of the face composite task. It may be argued that the time delay between viewing the target face and constructing the composite might have interfered with any differences between aphantasics and imagers. However, it is standard procedure in face composite research to use a delay because in real-world settings eyewitnesses do not construct composites until sometime (even days) after viewing a crime (Fodarella et al., 2015; Frowd et al., 2015). Relatively speaking, our delay was therefore short, although still reflective of delays used in prior research (e.g., Davis et al., 2014; Davis, Thorniley, Gibson, & Solomon, 2016; Pike, Brace, Turner, & Vredeveldt, 2019a). Similarly, it may be questioned whether our target face was *memorable* enough to be accurately constructed as a composite, since faces have been shown to vary widely in their memorability (Bainbridge, Isola, & Oliva, 2013). We suggest our target face was memorable enough for two reasons. First, our resemblance ratings are similar to those seen in previous research (e.g., Bhardwaj & Hole, 2020; Davis

et al., 2014, 2010, 2016). Second, we recently estimated the memorability of our target face using an algorithm developed by Needell and Bainbridge (2021), which gives a score between 0 (lowest memorability) and 1 (highest memorability). Our target face had a memorability score of 0.93 (out of 1), meaning its memorability was relatively good. Taken together, our composite task was designed to measure face composite creation abilities using a forensically relevant method.

In considering potential avenues for future research, we suggest a focus on other forensically relevant face processing tasks. Eyewitnesses are often asked to identify a criminal offender out of a larger group of suspects, often presented in videos or images, or within an in-person line-up (Home Office, 2017). Future research could therefore examine how people with and without aphantasia perform in these kinds of suspect identification tasks. Given their lower levels of face recognition, we may expect aphantasics to perform less accurately than imagers. On the other hand, aphantasics may not be impaired in these kinds of tasks given that additional attributes such as hair, clothing, gait, and voice may aid successful identification. In other words, these kinds of tasks do not rely *solely* on face recognition alone. As such, examining how visual imagery may influence performance in suspect identification procedures will be an interesting, and forensically important question for future research.

In conclusion, our study provides the first detailed examination of face recognition and face composite creation in people with and without visual imagery. We have shown that aphantasics tend to have poorer face recognition abilities than people with visual imagery, but they are just as accurate in constructing facial composites of a target face from memory (using the face composite system EFIT6; Visionmetric Ltd, 2019a). We also found that aphantasics find it more difficult to see the resemblance between a face composite and its target face relative to imagers, indicative of difficulties in face-matching (perception). Of particular relevance to forensic settings, we also found that holistic and featural methods of construction produce equally accurate facial composites within EFIT6, as do online (remote) and in-person methods. Overall, our study enhances our understanding of aphantasia, prosopagnosia, and the relationship between visual imagery and face processing more broadly. Our findings also have important implications for forensic settings, shining new light on techniques used to aid criminal investigations.

Chapter 6

General Discussion

The aim of this thesis was to examine the profile of congenital aphantasia from a clinical, neuropsychological, and forensic viewpoint. In four experimental chapters, I have answered key questions about aphantasia, including its prevalence, its relationship with other neurodevelopmental conditions (synaesthesia, and autism), its cross-sensory imagery qualities, its relationship with sensory sensitivity, and its role in face processing. This body of work significantly enhances our understanding of both aphantasia, and mental imagery more widely. The purpose of this final chapter is to discuss my experimental findings as a whole. Since detailed discussions are provided within each experimental chapter above, here I will summarise my results more broadly, bringing together findings from each chapter where appropriate. I will explore the implications and limitations of my findings, and highlight important avenues for future research motivated by the present thesis.

Discussion and implications of main findings

Chapter 2 provided a prevalence rate for aphantasia in the general population. I screened for aphantasia in two separate general population samples using the aphantasia diagnostic commonly used in contemporary research (*VVIQ*; Marks, 1973). I minimised risk of recruitment biases by not mentioning aphantasia or visual imagery in any recruitment material (i.e., ensuring people with particularly high or low visual imagery were not disproportionally encouraged to take part). In the combined cohort of participants (n1004), I found that the prevalence of aphantasia where visual imagery is *absent or vague/dim* (VVIQ = 16-32) was 3.9%. I additionally found that 0.8% experienced the most extreme subtype of aphantasia, where visual imagery is *completely absent* (VVIQ = 16). The prevalence of aphantasia was also similar in males (absent or vague/dim: 3.2%; totally absent: 0.5%) and females (absent or vague/dim: 4.2%; totally absent: 1%).

The prevalence estimates provided in this thesis advance the literature significantly since up until now the true prevalence of aphantasia was unknown. In past research, the most often cited prevalence rate for aphantasia has been approximately 2% (Betts, 1909; Faw, 2009; as cited in Fulford et al., 2018; Watkins, 2018; Zeman et al., 2015, 2020, etc.). This estimate is, however, problematic, since it only represents those who experience a

complete absence of visual imagery, so does not reflect aphantasia as we recognise it today (i.e., absent *or* dim/vague visual imagery; Keogh & Pearson, 2018; Zeman et al., 2015). Similarly, more recently Zeman et al. (2020) provided a prevalence of 0.7% for *complete* aphantasia (i.e., absence of visual imagery; VVIQ=16), but this study did not provide an estimate for aphantasia where visual imagery is absent *or crucially* vague/dim. The present thesis therefore adds a novel contribution to this literature by providing a prevalence estimate for aphantasia as it is defined today (i.e., absent, or dim/vague visual imagery). This thesis has therefore furthered our understanding of how commonly aphantasia occurs, and importantly, these prevalence estimates can be used by future researchers.

Chapter 3 explored the relationship between aphantasia and two other neurodevelopmental conditions linked to sensory differences: synaesthesia, and autism. Given theories linking synaesthesia to strong mental imagery (Barnett & Newell, 2008; Price, 2009; Ward, 2019b), I asked whether synaesthesia and aphantasia can co-occur. I also asked whether aphantasia might give rise to a less vivid form of synaesthesia, characterised by 'knowing' synaesthetic colours ('associator synaesthesia'), as opposed to 'seeing' them projected into the real world ('projector synaesthesia'; Dixon et al., 2004). The presence of grapheme colour synaesthesia was measured in two separate general population samples, using the 'gold standard' consistency test, where synaesthetes must be consistent over time in selecting colours for graphemes presented on screen (i.e., individuals repeatedly pick colours for letters A-Z, and numbers 0-9; see Chapter 3; Eagleman et al., 2007). I found that synaesthesia occurred at equal rates in people with, and without aphantasia. This was true of both cohort one (Chapter 3; Experiment 1a) where aphantasia was diagnosed using a single question (in-house imagery screener), and cohort two (Chapter 3; Experiment 1b) where aphantasia was diagnosed using the *VVIQ* (Marks, 1973, 1995).

How do these results from Chapter 3 sit within the broader literature? Importantly, my findings are at odds with past research suggesting that visual imagery tends to be heightened in synaesthetes relative to non-synaesthete controls (e.g., Barnett & Newell, 2008; Chiou et al., 2018; Chun & Hupé, 2016; Mealor et al., 2016; Spiller et al., 2015). More broadly, my findings challenge previous theories linking strong imagery to the development of synaesthesia (Price, 2009; Ward, 2019b). It is possible that previous

studies might have found high levels of imagery in synaesthetes because of a referral bias (Simner, 2013). Put simply, most studies tend to test synaesthetes who self-refer, and it might be exactly these synaesthetes who have the most phenomenologically vivid synaesthesia (and vivid imagery) (Simner, 2013). In Chapter 3, I tested imagery in large samples (Experiment 1a: >1000; Experiment 1b: >16,000), which might allow for a wider range of synaesthetes with different imagery strengths, and make it more likely to find those who have low imagery. As a consequence, I found that aphantasia and synaesthesia can indeed co-occur, highlighting for the very first time that intact visual imagery is *not necessary* for experiencing synaesthesia.

In Chapter 3 I did however find that aphantasia influences the *type* of synaesthesia experienced. In Experiment 1b, participants additionally completed the *Projector-Associator Questionnaire* (Rouw & Scholte, 2007), which measures traits associated with projector synaesthesia (i.e., 'seeing' colours projected into the real world), and associator synaesthesia (i.e., 'knowing' colours in some internal way, whether that be as mental imagery or conceptual associations alone). I found that aphantasic synaesthetes reported more associator traits (relative to projector traits) than non-aphantasic synaesthetes. My results are consistent with the *Conceptual-Mediation Model* of synaesthesia (Chiou & Rich, 2014) which suggests that at their root, synaesthetic experiences are conceptual associations (e.g., a synaesthete would 'know' that the letter A is associated with the colour red, in a similar way to knowing that bananas are yellow, and the grass is green, etc.). Whether the 'known' association is also experienced as imagery or not may depend on individual differences – i.e., if someone has *very strong* imagery they may be more likely to experience projections (Amsel et al., 2017). In this sense, all synaesthetes (aphantasics and imagers) would have conceptual knowledge of their associations (characterised by 'knowing'), even if some *additionally* experience their colours 'perceptually' through mental imagery.

Interestingly, although aphantasic synaesthetes were high in associator traits as a continuous measure, I also found that when participants were divided into groups using a binary categorisation threshold, then 10% of aphantasic synaesthetes were projectors, and 90% were associators. Similar rates were found in non-aphantasic synaesthetes, and these rates are reflective of the distribution of projectors and associators more generally within synaesthesia (Dixon et al., 2004). Since aphantasics have no imagery (or almost

no imagery), this tells us something new about projector synaesthesia: we now know that synaesthetes do not need imagery to have 'projections'. So, what are these non-imagery projections like exactly? Aphantasic synaesthetes might experience their projections less 'perceptually' than synaesthetes with imagery. In other words, an individual with aphantasia might be able to experience their synaesthetic colours 'out in space', even without an image of them. This may be possible due to their intact spatial abilities (e.g., Bainbridge et al., 2021; Dawes et al., 2020; Keogh & Pearson, 2018; Palermo et al., 2022; Pounder et al., 2022). Providing a novel perspective, I therefore propose that the projector-associator distinction might be better understood not only within a framework defined by the vividness of perceptual experience (i.e., linked to visual imagery), but also the *spatial frame* (i.e., not linked to visual imagery; see Chapter 3; see also Ward, Li, et al., 2006). Examining this framework in greater detail will be an interesting, and important question for future research.

In Chapter 3 I also examined the relationship between aphantasia and autism. In Experiment 2, I tested a group of aphantasics and a group of imaging controls using the *Autism Spectrum Quotient* (Baron-Cohen et al., 2001). In line with other recent research (Milton et al., 2021), I found that aphantasics experienced more autistic traits than imaging controls. I also found that there were more people in the aphantasic group than control group over the threshold suggestive of clinical levels of autism (AQ $\geq$ 32), suggesting that autism rates might be higher within aphantasics than in the general population. In particular, I found that aphantasics showed weaknesses in the imagination and social skills domains of autistic symptomatology, but were no different from controls in attention-to-detail, attention switching, and communication. These results suggest that aphantasia and autism may be linked via weaknesses in visual imagery, and that these imagery weaknesses may contribute – at least to some degree – to reduced imagination and social skills characteristic of autism (e.g., American Psychiatric Association, 2013; Baron-Cohen et al., 2001; Holt & Yuill, 2014; Hopkins et al., 2022). For example, imagery might aid the generation of *details* in imagination (D'Argembeau & Van der Linden, 2006), or help to take on different perspectives in social settings (Amit & Greene, 2012). In sum, Chapter 3 provides insight into the neuropsychological and clinical comorbidities associated with aphantasia, by showing how aphantasia intersects with synaesthesia, and autism.

In Chapter 4, I explored whether mental imagery is linked to sensory sensitivity (i.e., under- or over-responsivity to incoming sensory stimulation from the outside world), a clinically relevant trait that is commonly characteristic of autism (Bogdashina, 2003; Robertson & Simmons, 2015), but also variable within the general population (Robertson & Simmons, 2013). In answering this question, I first examined the broader mental imagery profile in aphantasia, showing that aphantasics reported lower imagery across *all* sense domains relative to controls (i.e., poor imagery not only in the visual domain, but also in the domains of auditory imagery, olfactory imagery, tactile imagery, gustatory imagery, bodily imagery, movement imagery, and feeling imagery). I also found that the imagery weaknesses were often severe enough to be considered 'aphantasia-like' in multiple senses. Indeed, nearly all of the aphantasics (97%) tested reported absent or dim/vague imagery in *at least one other sense*, compared to just 11% of controls. Next, I asked whether the vividness of imagery in these sense domains predicts the strength of sensory sensitivity (i.e., does low imagery = low sensory sensitivity?). I found that aphantasics self-reported lower sensory sensitivity than controls in nearly all of the senses tested (with the exception of the auditory domain, especially known for its high sensitivity, Robertson & Simmons, 2013; see avenues for future research below where I explore this in further detail) and this remained when controlling for the influence of autism traits. This suggests that imagery and sensory sensitivity are linked: people low in imagery (aphantasics) also tend to have low sensory sensitivity.

In Chapter 4, I also found that the relationship between imagery and sensory sensitivity existed in a general population sample (as visual imagery scores increased, so did overall sensory sensitivity scores), meaning this link was not specific to people with aphantasia (Chapter 4; Experiment 2). I also provided behavioural support for lower sensory sensitivity in aphantasia, using a *Pattern Glare Test* (Chapter 4, Experiment 3). Compared to imaging controls, aphantasics experienced less visual discomfort and fewer visual effects in response to irritable visual gratings, indicating lower levels of visual sensory sensitivity. Together, my results show for the very first time that the way we image sensory information (mental imagery) is linked to the way we experience incoming sensory information from the outside world (sensory sensitivity). So, why might imagery and sensory sensitivity be linked? In Chapter 3 I proposed a model whereby imagery and sensory sensitivity are linked via levels of cortical excitation. This model suggests that low excitability in sensory cortex may link both low imagery and low sensory sensitivity

(it also predicts the reverse, i.e., high excitability tied to both high imagery and high sensory sensitivity). This model is based on past research which shows that imagery and sensory sensitivity are (independently) linked to sensory cortex excitation, where high imagery/sensitivity are linked to lower phosphene thresholds (indicative of higher excitability) in the relevant sensory cortex (Cattaneo et al., 2011; Mulleners et al., 2001; Sparing et al., 2002; Wilkins et al., 1984; Young et al., 2004). This model therefore proposes that cortical excitability might be a neural mechanism linking imagery to sensory sensitivity (see Chapter 4), providing a testable theory for future research.

Another implication of the findings from Chapter 4 is that I demonstrated how aphantasia might be better understood within a broader framework of imagery impairment across multiple senses (see also Dawes et al., 2020; Watkins, 2018; Zeman et al., 2020). Accordingly, I coined the term *dysikonesia* as an umbrella term for imagery weakness in any sense (or indeed multiple senses). My proposal fits well with a subsequent term introduced by Hinwar and Lambert (2021) to describe poor auditory imagery (*anauralia*). As such, we can now construe aphantasia (poor visual imagery) and anauralia (poor auditory imagery) as distinct sub-types within dysikonesia. Interestingly, my proposal for the term dysikonesia has recently been debated in a series of commentaries published in the journal *Cortex* (Lambert & Sibley, 2022; Monzel, Mitchell, Macpherson, Pearson, & Zeman, 2022a, 2022b; Simner & Dance, 2022). Monzel, Mitchell, Macpherson, Pearson, and Zeman (2022a) suggested their preference was to instead use the word ‘aphantasia’ for *all* sense modalities (e.g., visual aphantasia, auditory aphantasia, etc.). In our own reply (Simner & Dance, 2022) we pointed out that this would mean undermining a vibrant and growing literature where the term aphantasia already has a very specific meaning (poor visual imagery alone), and even overwriting the clear etymological roots of the term aphantasia (which relate to things *seen,* i.e., *phantasma*; Online Etymology Dictionary, 2001; Zeman et al., 2015, 2016). Other authors agreed: Lambert & Sibley (2022) pointed out that using the expression ‘auditory aphantasia’ (in the place of anauralia) would be akin to describing deafness as ‘auditory blindness’. Hence, I and others continue to support the notion of aphantasia as a specific visual imagery weakness within the wider framework of dysikonesia. But time itself will inevitably determine which terms prevail, as whichever become used most widely by researchers, and the aphantasia community alike.

Collectively, Chapters 3 and 4 come together more broadly to answer key questions about how aphantasia intersects with other neurodevelopmental conditions. We saw in Chapter 3 that aphantasia is associated with high levels of autistic traits, with particular weaknesses in imagination and social skills (Chapter 3). Despite this, we have also seen in Chapter 4 that aphantasics do *not* show another common autistic trait, which is high sensory sensitivity (aphantasics in fact experience *low* sensory sensitivity relative to imaging controls). Interestingly, synaesthesia has also been associated with high levels of autistic traits, with synaesthetes showing particularly heightened autistic symptomatology in the domains of (strong) attention-to-detail and (high) sensory sensitivity (Van Leeuwen et al., 2019; Ward, Brown, et al., 2018; Ward et al., 2017). This thesis furthers the literature by providing clarity on how these sensory conditions are linked. Taken together, the findings above suggest that these conditions might be tied through different *clusters* of traits: synaesthesia and autism through (heightened) attention-to-detail and sensory sensitivity, and aphantasia and autism through (poor) imagination and social skills. While bearing these differences in mind, it is also important to consider whether there might be a factor that links the conditions more broadly. One possibility is *spatial imagery*. Both synaesthesia and autism have been linked to strong spatial imagery (e.g., Bouvet et al., 2019; Ward et al., 2018), as has aphantasia (Dawes et al., 2020; Keogh & Pearson, 2018; Palermo et al., 2022). Examining in more detail the underlying commonalities and differences within these different neurodevelopmental conditions will be an important avenue for future research, motivated by some of the findings of the present thesis.

In Chapter 5, the final empirical chapter of this thesis, I investigated the forensically relevant question of whether aphantasia influences face recognition and the ability to construct facial composites. I tested people with and without aphantasia on a range of self-report and behavioural tasks. I found that aphantasics self-reported more traits associated with prosopagnosia than imaging controls on the *Prosopagnosia Index* (Shah, Gaule, et al., 2015), replicating Milton et al. (2021) (but see also Palermo et al., 2022). I also found more aphantasics than controls over the threshold suggestive of the presence of at least mild prosopagnosia on the Prosopagnosia Index (≥65). Aphantasics also performed worse overall on the *Cambridge Face Memory Test* (Duchaine & Nakayama, 2006), a behavioural test of face recognition. Interestingly, aphantasics also showed weaknesses in perceptual face-matching abilities, since our aphantasic raters of the face

composites struggled more than imagers to see the resemblance between composites and their respective target face. Taken together, these results suggest that aphantasics have poorer face recognition and face perception abilities than people with intact visual imagery.

Interestingly, the results from Chapter 5 are at odds with previous studies showing no difference between people high and low (e.g., aphantasics) in visual imagery on face recognition tasks (McKelvie, 1994; Milton et al., 2021). This disparity can perhaps be easily explained by the fact that the face recognition tasks used by Mckelvie (1994) and Milton et al. (2021) included extra non-facial features (e.g., hair, clothing, etc.). This can give rise to successful recognition even if people otherwise struggle with face recognition (Duchaine & Weidenfeld, 2003). Given that the face recognition tasks used in Chapter 5 of this thesis *excluded* non-facial features, and have been shown to reliably diagnose prosopagnosia (Duchaine & Weidenfeld, 2003; Shah, Gaule, et al., 2015; Tsantani et al., 2021), my findings provide a strong case for face recognition being impaired (on average) in aphantasia. More broadly, my results suggest that prosopagnosia and aphantasia might be linked in some way, given that low imagery and low face recognition seem to be commonly characteristic of both conditions (Grueter et al., 2007; Grüter et al., 2009; Milton et al., 2021; Tree & Wilkie, 2010).

Chapter 5 also examined whether aphantasics and people with intact visual imagery differ in their ability to construct facial composites of a target face from memory, a task commonly used in forensic settings (Frowd et al., 2019; Home Office, 2017). I found no difference in the resemblance accuracy of composites produced by aphantasics and controls, suggesting that despite weaknesses in face recognition, aphantasics have an intact ability to construct facial composites. Their generally intact ability to re-construct pictures from memory (Bainbridge et al., 2021), and their lower tendency to produce false memories (Bainbridge et al., 2021), might provide aphantasics with the skills necessary to construct composites. Overall, the findings from Chapter 5 show that intact visual imagery is not necessary for producing facial composites, but it nevertheless facilitates face recognition and face perception more generally (e.g., ‘rehearsing’ faces using imagery might in some way improve their intrinsic ‘memorability’; Bainbridge et al., 2013).

One of the recurring themes in this thesis is how aphantasia intersects with autistic traits. Chapter 5 provides further insight into this relationship, since issues with face recognition are commonly characteristic of autism (Griffin et al., 2021). In Chapter 5 I additionally asked whether face recognition problems in aphantasia might be attributed to elevated autistic traits, by administering the *Autism Spectrum Quotient* (Baron-Cohen et al., 2001) to a subset of participants. I found that autism traits did not robustly correlate with any of the face processing outcomes, suggesting that face recognition impairments in aphantasia cannot be attributed to autism. However, I point out here that even if face recognition problems *were* linked to autism in some cases, this would not necessarily negate the relationship between imagery and face recognition. On the contrary, the link between visual imagery and face recognition may even *explain* the weaknesses in face recognition seen in autism – at least to some degree. For now, though, this thesis finds no significant link between autistic traits and face recognition in aphantasia. Instead, my results provide support for face recognition being linked to visual imagery abilities more broadly.

Limitations

One of the methodological barriers facing this thesis was how to recruit people with aphantasia, a condition that is relatively rare (see Chapter 2). When commencing this thesis, I set out to create the *Sussex University - Aphantasia Cohort*, a group of people with aphantasia who were interested in taking part in research by *The Imagery Lab* at Sussex University. I did this through advertising online (e.g., through social media), through word-of-mouth, and through screening the undergraduate Psychology cohorts at Sussex University. At the time of submitting this thesis, this cohort now stands at >1000 individuals world-wide, and has been a key achievement of my PhD. Creating this resource has allowed me to successfully recruit participants for my studies, both online and in-person, and they were used in all chapters apart from my prevalence study (Chapter 2; which required random sampling). However, one limitation of this cohort is that they might not represent the wider population of aphantasia per se, which strictly speaking one can only find by randomly sampling very large populations for every investigation. This suggests that future research might seek to replicate my results in a randomly sampled population. However, to have achieved my sample size across all my studies (and with an aphantasia prevalence of 3.9%), I would have had to screen >20,000 members of the population to find the necessary subset of aphantasics – a task unfortunately beyond the scope of this thesis. Alternatively, I would have had to test somewhat smaller samples,

but administer every single test, in an unfeasibly long study. With this limitation in mind, it is nevertheless important to note that self-referred samples are used widely in psychology research, and on the whole they do produce generalisable results (Casler et al., 2013; Chandler & Shapiro, 2016).

Another barrier that I faced when writing this thesis was the Covid-19 pandemic. When Covid-19 hit the UK in the early part of 2020, I was in the middle of in-person testing for two studies. One of these studies was my face processing study (Chapter 5), and the other aimed to examine the role of imagery and aphantasia in worry and heart rate variability. Since in-person testing ceased during the recurring 'lock-downs', I had to quickly adapt my thesis plan. I was able to adapt the face processing study into an online (remote) format. This meant testing a new cohort of individuals online (n25 aphantasics, n25 controls), in addition to the participants I had already tested in-person (n27 aphantasics, n15 controls). Despite the fact that the new round of online testing pushed back my thesis timeline, the 'silver lining' was that I was able to ask an additional question: is the accuracy of face composites influenced by whether they are constructed in-person or online/remotely? My findings showed that remote composite construction is just as effective as in-person construction (see Chapter 5). This novel finding is forensically relevant since it can allow future witness testimony to be confidently gathered over geographically remote areas, and would allow continuity for criminal investigations in the face of further lockdowns. Overall, in terms of the face processing study, I was able to adapt effectively to barriers caused by the Covid-19 pandemic.

Unfortunately, my second in-person study – examining the role of imagery and aphantasia in worry and heart rate variability – had to be stopped indefinitely. This was because key measures (e.g., heart rate) were more difficult to gather remotely, and more importantly, the exact process I was measuring in aphantasia (anxiety/worry) was exacerbated in the general population during the pandemic (Santabárbara et al., 2021). This meant I would have been unable to compare data from pre- and post-pandemic participants. Moreover, the prolonged uncertainty about when in-person testing would re-start meant I could not make plans to repeat, or run another in-person study within my thesis timeline. Again, however, a 'silver lining' has been that I was able to use my preliminary data from this study to seed a new postdoctoral research grant. As such, I was again able to adapt effectively to setbacks caused by the pandemic.

Thinking about limitations more broadly, some have suggested that one direction for future research is the development of easy to administer 'objective' measures of aphantasia. This is because the wider contemporary literature generally relies on self-report measures such as the VVIQ (Marks, 1973, 1995), which have been criticised for their 'subjectivity'. Indeed, one recent study criticised the VVIQ since it did not correlate with a new physiological measure of aphantasia. Using pupillometry, Kay et al. (2021) found the pupils' of people with aphantasia do not constrict when they imagine bright stimuli relative to dark stimuli, a physiological index of imagery seen in imagers. They also found that overall VVIQ scores *did not* correlate with changes in pupil diameter. However, I point out here that one of the measures used to *categorise* participants (as aphantasics or controls) in this very study was the VVIQ, so arguably VVIQ scores do map onto *group-based* differences in pupillary responses. Moreover, trial-by-trial imagery vividness ratings also predicted changes in pupil diameter, meaning that self-reported imagery mapped onto the physiological measure of imagery at least to some degree. Other research too shows that people tend to have good metacognition about their imagery abilities, since self-reported visual imagery maps onto objective lab-based imagery binocular rivalry tasks (Keogh & Pearson, 2018; Pearson et al., 2011). In these tasks, imaging a colour (e.g., green or red) means it is more likely to become dominant in a subsequent binocular rivalry display of the two colours, and the strength of this priming effect is indicative of visual imagery strength (Pearson et al., 2008). Although it could be questioned whether the imagery measured within this task (i.e., imaging a *colour*) is truly reflective of visual imagery in everyday life (which usually involves more complex representations of objects, scenes etc.), this dominance effect maps onto self-reported visual imagery trial-by-trial (i.e., how vivid the colour is reported to be) and overall visual imagery abilities as measured by the VVIQ (Keogh & Pearson, 2018; Pearson et al., 2008, 2011). Moreover, this effect has been replicated multiple times in imagers (e.g., Pearson et al., 2008, 2011), and is not present in people with aphantasia (Keogh & Pearson, 2018). Taken together, this suggests that this binocular rivalry task does indeed tap into visual imagery abilities. Adapting these kinds of 'objective' measures so that they can be administered easily online will therefore be a valuable step for future research. However, these findings also highlight that self-report measures (such as the VVIQ) still have great utility in the study of aphantasia. Indeed, self-report measures may even pick up on more nuanced differences that objective measures fail to detect (e.g., an individual might struggle with *one type of imagery* in particular such as

imagery for faces, as often seen in prosopagnosia; Grueter et al., 2007; Grüter et al., 2009). As such, both self-report and 'objective' measures are likely to retain their value in future research.

Questions for future research

One avenue for future research motivated by the findings of the present thesis is exploring in further detail how aphantasia intersects with other clinical or neurodevelopmental conditions. One question is whether aphantasic populations show a higher rate of diagnosed autism, given their high levels of autistic traits, and the higher number of aphantasics passing the AQ threshold suggestive of clinical autism (see Chapter 3). I suggest this avenue for future research since the AQ cannot be used as a diagnostic measure per se, so it would be useful to confirm against rates of clinical diagnosis. A second key question is how aphantasia might link to other sensory traits such as *misophonia*, a sensory sensitivity in the auditory domain (Rinaldi, Simner, Koursarou, & Ward, 2022; Schröder, Vulink, & Denys, 2013; Simner, Koursarou, Rinaldi, & Ward, 2021). People with misophonia have extreme aversions to certain sounds (e.g., human bodily sounds), and often have comorbidities with other auditory sensitivities (e.g., *hyperacusis*: a sensitivity to loud noises; Baguley & McFerran, 2011; Jastreboff & Jastreboff, 2015). This thesis showed that aphantasics reported lower sensory sensitivity than imaging controls in all senses *except* the auditory domain (Chapter 4). This raises the question of whether misophonia/hyperacusis are the only sensory sensitivities against which aphantasics are not protected. Examining links between aphantasia/imagery and misophonia and how this affects auditory sensory sensitivity will therefore be an interesting avenue for future research.

Another important avenue for future research is examining how aphantasia influences clinical psychopathology. A small number of studies have started to examine this. For example, compared to people with intact visual imagery, aphantasics are less likely to experience intrusive imagery of past traumatic events (Dawes et al., 2020), and physiological fear responses when reading frightening passages (Wicken et al., 2021). They are also less susceptible to experiencing pseudo-hallucinations (Königsmark et al., 2021). Given the role of imagery in depression (e.g., imagery of suicidal acts increasing the likelihood of an attempt; Holmes, Crane, Fennell, & Williams, 2007), and in anxiety disorders such as post-traumatic stress disorder (i.e., flashbacks; Clark & Mackay, 2015)

and social anxiety disorder (e.g., imagery of social situations going wrong; Hirsch, Clark, & Mathews, 2006), it will be important to examine comorbidities between aphantasia and mental health conditions, assessing whether aphantasia offers a level of *protection* against traits linked to imagery.

Another important area for future research is exploring the neural underpinnings of aphantasia. Previous research has shown that compared to average imagers and hyperphantasics, people with aphantasia show less BOLD activity in the left anterior parietal cortex when asked to imagine (as compared to perceive) famous faces/places (Milton et al., 2021). Relative to hyperphantasics, aphantasics also have reduced (resting state fMRI) connectivity between top-down control (prefrontal cortex) and visual-occipital areas (Milton et al., 2021). Building on this initial neuroimaging research, future work might wish to examine activation in the visual cortex in more detail, since differences in activation here between aphantasics and imaging controls might further clarify the *function* of visual imagery, within the context of 'The Imagery Debate' (i.e., depictive vs. propositional representation; see Chapter 1). Previous research has shown that it is possible to decode the content of a visual image from activation patterns in the primary visual cortex (Albers et al., 2013; Harrison & Tong, 2009; Serences, Ester, Vogel, & Awh, 2009). Since the primary visual cortex is retinotopic, that is, regions of neurons located next to each other process stimuli that are next to each other in space (Mountcastle, 1997), this suggests that *depictive* (as opposed to propositional) representations might give rise to imagery (Kosslyn et al., 2006; Pearson & Kosslyn, 2015). If we were to find the same depictive activation in aphantasics when they are asked to imagine a scene (albeit, without the conscious experience of visual imagery), this might suggest that aphantasics, too, have intact depictive representations. Crucially, however, this would also show that the conscious experience of imagery is *not necessary* in experiencing depictive representations. Instead, visual imagery might be an internal 'display' of information that is stored in a depictive format in the brain, seen only in imagers. This way, both aphantasics and imagers would have intact depictive representations, but only imagers would additionally 'see' this information as visual imagery. Alternatively, future research might also find individual differences in the way information is stored, including propositional formats (Pylyshyn, 1973, 2002), or a mixture of the two (Pearson & Kosslyn, 2015). Examining the neural mechanisms of

aphantasia therefore has important implications for our understanding the causes of imagery absence, but also the function of mental imagery within cognition more broadly.

Another avenue for future research is examining the presentation of aphantasia and mental imagery cross-culturally. One important question is whether current measures of aphantasia (e.g., the VVIQ) are applicable across cultures. Although I have recruited aphantasic individuals living in world-wide destinations within the *Sussex University - Aphantasia Cohort*, and see no reason within the literature that imagery would not exist across cultures, we may find that the *language* for mental imagery, or indeed the *concept* of this sensory experience, does not exist universally. In support of this, adapting questionnaire measures so they can be used effectively across cultures can be challenging given differences in language and conceptual understandings (Epstein, Santo, & Guillemin, 2015). This highlights an important avenue for future imagery research, and provides us with an important reminder that as researchers we need to be mindful of the generalisability and applicability of mental imagery and aphantasia measures across cultures.

Finally, when thinking about how to define aphantasia, and other imagery differences (e.g., hyperphantasia, dysikonesia, anauralia, etc.), researchers need to be sensitive to the appropriateness of language used more broadly. As we saw in the Introduction of this thesis (Chapter 1), aphantasia is currently defined within the existing literature as a 'congenital condition'. Some may question whether we can accurately categorise aphantasia as congenital, or even 'neurodevelopmental', since we are relying on retrospective self-report for diagnosis. That is, we assume that if an individual has no recollection of ever experiencing imagery, then this is due to an absence of visual imagery from birth. An alternative explanation is that imagery might have been experienced within infant years, but not accurately recollected at a later date. Despite these inherent limitations, retrospective self-report is, however, used widely to other categorise congenital conditions (such as congenital prosopagnosia; Behrmann & Avidan, 2005). Indeed, self-report is often thought to have great utility in measuring the phenomenology of conditions characterised by 'inner world' or 'subjective' experiences, that are often less easily tapped by objective measures. Currently, retrospective self-report therefore remains a valid way of identifying life-long versus acquired cases of aphantasia.

Relatedly, some may question whether it is accurate to define aphantasia as a 'condition', or alternatively, whether aphantasia might be better understood as the lower end of a continuous visual imagery spectrum (Pearson, 2019). As we have seen throughout this thesis and in past research, aphantasia tends to be characterised by a distinct set of characteristics, and these characteristics do not always have a linear relationship with visual imagery abilities. As such, current understandings define aphantasia as a categorical condition, akin to the way other sensory differences (e.g., synaesthesia) characterised by distinct profiles of traits are also defined as conditions. It is nevertheless important to stress that the use of the word condition is not in any way meant to imply that aphantasia is linked to any kind of inherent 'deficit'. Related to this, imagery researchers have an ongoing responsibility to ensure we explain our use of language, ensuring that it is seen as appropriate and useful in the eyes of individuals who experience imagery differences. Indeed, aphantasia researchers tend work closely with people who experience imagery differences, otherwise known as 'expert-by-experience'. This collaborative framework is highlighted by the *Extreme Imagination Conference*, which brings together researchers and experts-by-experience to share developments in research, and discuss personal experiences of sensory differences (e.g., see Zeman et al., 2019). This approach is also illustrated by the *Aphantasia Network*, an online community of experts-by-experience, researchers, and individuals interested in learning more about mental imagery differences. This network promotes collaboration between science and the community, including dissemination of research findings (e.g., see https://aphantasia.com/sensory-overwhelm/ which shares results relating to Chapter 4 of this thesis). Taken together, this collaborative approach helps to ensure that our research questions, and the language we use to define mental imagery differences, are informed by the very community who experience these sensory traits (see Monzel, Mitchell, et al., 2022a; Monzel, Mitchell, Macpherson, Pearson, & Zeman, 2022b; Monzel, Vetterlein, & Reuter, 2022; Simner & Dance, 2022). As research develops, it is essential that researchers continue to collaborate with experts-by-experience in this way; because first and foremost, the overarching aim of mine and others' research is to *help and support* individuals who experience sensory differences to better understand themselves and others, so actively listening and valuing their perspectives is of the utmost importance.

Conclusions

This thesis set out to examine the profile of aphantasia from a clinical, neuropsychological, and forensic perspective. I have answered key questions about aphantasia including its prevalence in the general population, and how it links with other clinical and neuropsychological conditions linked to sensory differences (synaesthesia, and autism). I have also shown that aphantasia is commonly tied to low imagery in other (non-visual) sense domains, and have introduced the term *dysikonesia* to account for this (characterising poor imagery across one or more senses, where aphantasia is the *visual* subtype). I have also shown for the very first time that mental imagery predicts sensory sensitivity both within aphantasic individuals, and in the general population, and have proposed a novel model based on sensory cortex excitation to explain this relationship. Finally, I asked the forensically important question of whether aphantasia influences face processing. Aphantasia is tied to weaknesses in face perception and face recognition, but intact face composite construction abilities. This body of work has provided a significant step forwards in our understanding of congenital aphantasia, and the role of mental imagery in cognition and sensation more broadly, and importantly, my findings provide motivation for numerous future research questions.

References

Addis, D. R., Wong, A. T., & Schacter, D. L. (2007). Remembering the past and imagining the future: Common and distinct neural substrates during event construction and elaboration. *Neuropsychologia*, *45*(7), 1363–1377. https://doi.org/10.1016/j.neuropsychologia.2006.10.016

Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & De Lange, F. P. (2013). Shared representations for working memory and mental imagery in early visual cortex. *Current Biology*, *23*(15), 1427–1431. https://doi.org/10.1016/j.cub.2013.05.065

Allbutt, J., Ling, J., Rowley, M., & Shafiullah, M. (2011). Vividness of visual imagery and social desirable responding: Correlations of the vividness of visual imagery questionnaire with the balanced inventory of desirable responding and the Marlowe-Crowne scale. *Behavior Research Methods*, *43*, 791. https://doi.org/10.3758/s13428-011-0086-8

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Amit, E., & Greene, J. D. (2012). You See, the Ends Don't Justify the Means: Visual Imagery and Moral Judgment. *Psychological Science*, *23*(8), 861–868. https://doi.org/10.1177/0956797611434965

Amsel, B. D., Kutas, M., & Coulson, S. (2017). Projectors, associators, visual imagery, and the time course of visual processing in grapheme-color synesthesia. *Cognitive Neuroscience*, *8*(4), 206–223. https://doi.org/10.1080/17588928.2017.1353492

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

Arcangeli, M. (2020). The Two Faces of Mental Imagery. *Philosophy and Phenomenological Research*, *101*(2), 304–322. https://doi.org/10.1111/phpr.12589

Aurora, S. K., Cao, Y., Bowyer, S. M., & Welch, K. M. A. (1999). The occipital cortex is hyperexcitable in migraine: Experimental evidence. *Headache*, *39*(7), 469–476. https://doi.org/10.1046/j.1526-4610.1999.3907469.x

Aurora, S. K., Welch, K. M. A., & Al-Sayed, F. (2003). The threshold for phosphenes is lower in migraine. *Cephalalgia*, *23*(4), 258–263. https://doi.org/10.1046/j.1468-

2982.2003.00471.x

Aurora, S. K., & Wilkinson, F. (2007). The brain is hyperexcitable in migraine. *Cephalalgia*, *27*(12), 1442–1453. https://doi.org/10.1111/j.1468-2982.2007.01502.x

Baddeley, A. D., & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology: General*, *129*(1), 126–145. https://doi.org/10.1037/0096-3445.129.1.126

Baguley, D. M., & McFerran, D. J. (2011). Hyperacusis and Disorders of Loudness Perception. In A. R. Møller, B. Langguth, D. De Ridder, & T. Kleinjung (Eds.), *Textbook of Tinnitus* (pp. 13–23). https://doi.org/10.1007/978-1-60761-145-5_3

Bainbridge, W. A., Isola, P., & Oliva, A. (2013). The intrinsic memorability of face photographs. *Journal of Experimental Psychology: General*, *142*(4), 1323–1334. https://doi.org/10.1037/a0033872

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

Barnett, K. J., & Newell, F. N. (2008). Synaesthesia is associated with enhanced, self-rated visual imagery. *Consciousness and Cognition*, *17*(3), 1032–1039. https://doi.org/10.1016/j.concog.2007.05.011

Baron-Cohen, S. (1987). Autism and symbolic play. *British Journal of Developmental Psychology*, *5*(2), 139–148. https://doi.org/10.1111/j.2044-835x.1987.tb01049.x

Baron-Cohen, S. (2000). Theory of mind and autism: A review. *International Review of Research in Mental Retardation*, *23*, 169–184. https://doi.org/https://doi.org/10.1016/S0074-7750(00)80010-5

Baron-Cohen, S., Ashwin, E., Ashwin, C., Tavassoli, T., & Chakrabarti, B. (2009). Talent in autism: Hyper-systemizing, hyper-attention to detail and sensory hypersensitivity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1522), 1377–1383. https://doi.org/10.1098/rstb.2008.0337

Baron-Cohen, S., Johnson, D., Asher, J., Wheelwright, S., Fisher, S. E., Gregersen, P. K., & Allison, C. (2013). Is synaesthesia more common in autism? *Molecular Autism*, *4*, 40. https://doi.org/10.1186/2040-2392-4-40

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, *21*(1), 37–46. https://doi.org/10.1016/0010-

0277(85)90022-8

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Males and Females, Scientists and Mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5–17. https://doi.org/10.1023/A:1005653411471

Beeli, G., Esslen, M., & Jäncke, L. (2005). When coloured words taste sweet. *Nature*, *434*, 38. https://doi.org/10.1038/434038a

Behrmann, M., & Avidan, G. (2005). Congenital prosopagnosia: Face-blind from birth. *Trends in Cognitive Sciences*, *9*(4), 180–187. https://doi.org/10.1016/j.tics.2005.02.011

Belardinelli, M. O., Palmiero, M., Sestieri, C., Nardo, D., Di Matteo, R., Londei, A., … Romani, G. L. (2009). An fMRI investigation on image generation in different sensory modalities: The influence of vividness. *Acta Psychologica*, *132*(2), 190–200. https://doi.org/10.1016/j.actpsy.2009.06.009

Ben-Sasson, A., Hen, L., Fluss, R., Cermak, S. A., Engel-Yeger, B., & Gal, E. (2009). A meta-analysis of sensory modulation symptoms in individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *39*, 1–11. https://doi.org/10.1007/s10803-008-0593-3

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Benjamini, Y., & Hochberg, Y. (2000). On the adaptive control of the false discovery rate in multiple testing with independent statistics. *Journal of Educational and Behavioral Statistics*, *25*(1), 60–83. https://doi.org/10.3102/10769986025001060

Berger, C. C., & Ehrsson, H. H. (2017). The content of imagined sounds changes visual motion perception in the cross-bounce illusion. *Scientific Reports*, *7*, 40123. https://doi.org/10.1038/srep40123

Betts, G. H. (1909). *The distribution and functions of mental imagery*. Columbia University, New York.

Bhardwaj, K., & Hole, G. (2020). Effect of racial bias on composite construction. *Applied Cognitive Psychology*, *34*, 616–627. https://doi.org/10.1002/acp.3655

Blackwell, S. E., Rius-Ottenheim, N., Schulte-van Maaren, Y. W. M., Carlier, I. V. E.,

Middelkoop, V. D., Zitman, F. G., … Giltay, E. J. (2013). Optimism and mental imagery: A possible cognitive marker to promote well-being? *Psychiatry Research*, *206*(1), 56–61. https://doi.org/10.1016/j.psychres.2012.09.047

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

Bogdashina, O. (2003). *Sensory perceptual issues in autism and Asperger syndrome: Different sensory experiences – different perceptual worlds.* London, UK: Jessica Kingsley.

Borduin, B. J., Borduin, C. M., & Manley, C. M. (1994). The use of Imagery Training to Improve Reading Comprehension of Second Graders. *Journal of Genetic Psychology*, *155*(1), 115–118. https://doi.org/10.1080/00221325.1994.9914764

Bouvet, L., Amsellem, F., Maruani, A., Tonus-Vic Dupont, A., Mathieu, A., Bourgeron, T., … Mottron, L. (2019). Synesthesia & autistic features in a large family: Evidence for spatial imagery as a common factor. *Behavioural Brain Research*, *362*, 266–272. https://doi.org/10.1016/j.bbr.2019.01.014

Brace, N. A., Pike, G. E., & Turner, J. A. (2008). Holistic facial composite systems: are they compatible with witness recall? *Cognitive Technology Journal*, *13*(2), 42–49.

Braithwaite, J. J., Broglia, E., Bagshaw, A. P., & Wilkins, A. J. (2013). Evidence for elevated cortical hyperexcitability and its association with out-of-body experiences in the non-clinical population: New findings from a pattern-glare task. *Cortex*, *49*(3), 793–805. https://doi.org/10.1016/j.cortex.2011.11.013

Braithwaite, J. J., Broglia, E., Brincat, O., Stapley, L., Wilkins, A. J., & Takahashi, C. (2013). Signs of increased cortical hyperexcitability selectively associated with spontaneous anomalous bodily experiences in a nonclinical population. *Cognitive Neuropsychiatry*, *18*(6), 549–573. https://doi.org/10.1080/13546805.2013.768176

Brang, D., & Ahn, E. (2019). Double-blind study of visual imagery in grapheme-color synesthesia. *Cortex*, *117*, 89–95. https://doi.org/10.1016/j.cortex.2019.02.025

Brigo, F., Storti, M., Nardone, R., Fiaschi, A., Bongiovanni, L. G., Tezzon, F., & Manganotti, P. (2012). Transcranial magnetic stimulation of visual cortex in migraine patients: A systematic review with meta-analysis. *Journal of Headache and Pain*, *13*, 339–349. https://doi.org/10.1007/s10194-012-0445-6

Burton, A. M., White, D., & McNeill, A. (2010). The glasgow face matching test. *Behavior Research Methods*, *42*(1), 286–291.

https://doi.org/10.3758/BRM.42.1.286

Cabeza, R., Burton, A. M., Kelly, S. W., & Akamatsu, S. (1997). Investigating the Relation between Imagery and Perception: Evidence from Face Priming. *The Quaterly Journal of Experimental Psychology*, *50A*(2), 274–289.

Campos, A. (2014). Gender differences in imagery. *Personality and Individual Differences*, *59*, 107–111. https://doi.org/10.1016/j.paid.2013.12.010

Carmichael, D. A., Down, M. P., Shillcock, R. C., Eagleman, D. M., & Simner, J. (2015). Validating a standardised test battery for synesthesia: Does the Synesthesia Battery reliably detect synesthesia? *Consciousness and Cognition*, *33*, 375–385. https://doi.org/10.1016/j.concog.2015.02.001

Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, *29*(6), 2156–2160. https://doi.org/10.1016/j.chb.2013.05.009

Cattaneo, Z., Pisoni, A., Papagno, C., & Silvanto, J. (2011). Modulation of visual cortical excitability by working memory: Effect of luminance contrast of mental imagery. *Frontiers in Psychology*, *2*, 29. https://doi.org/10.3389/fpsyg.2011.00029

Center, Y., Freeman, L., Robertson, G., & Outhred, L. (1999). The effect of visual imagery training on the reading and listening comprehension of low listening comprehenders in Year 2. *Journal of Research in Reading*, *22*(3), 241–256. https://doi.org/10.1111/1467-9817.00088

Chadaide, Z., Arlt, S., Antal, A., Nitsche, M. A., Lang, N., & Paulus, W. (2007). Transcranial direct current stimulation reveals inhibitory deficiency in migraine. *Cephalalgia*, *27*(7), 833–839. https://doi.org/10.1111/j.1468-2982.2007.01337.x

Chandler, J., & Shapiro, D. (2016). Conducting Clinical Research Using Crowdsourced Convenience Samples. *Annual Review of Clinical Psychology*, *12*, 53–81. https://doi.org/10.1146/annurev-clinpsy-021815-093623

Chiou, R., & Rich, A. N. (2014). The role of conceptual knowledge in understanding synesthesia: Evaluating contemporary findings from a "hub-and-spokes" perspective. *Frontiers in Psychology*, *5*, 105. https://doi.org/10.3389/fpsyg.2014.00105

Chiou, R., Rich, A. N., Rogers, S., & Pearson, J. (2018). Exploring the functional nature of synaesthetic colour: Dissociations from colour perception and imagery. *Cognition*, *177*, 107–121. https://doi.org/10.1016/j.cognition.2018.03.022

Chun, C. A., & Hupé, J. M. (2016). Are synesthetes exceptional beyond their synesthetic associations? A systematic comparison of creativity, personality, cognition, and mental imagery in synesthetes and controls. *British Journal of Psychology*, *107*(3), 397–418. https://doi.org/10.1111/bjop.12146

Clark, I. A., & Mackay, C. E. (2015). Mental imagery and post-traumatic stress disorder: A neuroimaging and experimental psychopathology approach to intrusive memories of trauma. *Frontiers in Psychiatry*, *6*, 104. https://doi.org/10.3389/fpsyt.2015.00104

Collishaw, S. M., & Hole, G. J. (2000). Featural and configurational processes in the recognition of faces of different familiarity. *Perception*, *29*(8), 893–909. https://doi.org/10.1068/p2949

Commodari, E., Guarnera, M., Di Stefano, A., & Di Nuovo, S. (2020). Children Learn to Read: How Visual Analysis and Mental Imagery Contribute to the Reading Performances at Different Stages of Reading Acquisition. *Journal of Psycholinguistic Research*, *49*, 59–72. https://doi.org/10.1007/s10936-019-09671-w

Conlon, E., Lovegrove, W., Barker, S., & Chekaluk, E. (2001). Visual discomfort: The influence of spatial frequency. *Perception*, *30*(5), 571–581. https://doi.org/10.1068/p2954

Conson, M., Mazzarella, E., Esposito, D., Grossi, D., Marino, N., Massagli, A., & Frolli, A. (2015). "Put Myself Into Your Place": Embodied Simulation and Perspective Taking in Autism Spectrum Disorders. *Autism Research*, *8*(4), 454–466. https://doi.org/10.1002/aur.1460

Coutts, L. V., Cooper, C. E., Elwell, C. E., & Wilkins, A. J. (2012). Time course of the haemodynamic response to visual stimulation in migraine, measured using near-infrared spectroscopy. *Cephalalgia*, *32*(8), 621–629. https://doi.org/10.1177/0333102412444474

Croydon, A., Pimperton, H., Ewing, L., Duchaine, B., & Pellicano, E. (2014). The Cambridge Face Memory Test for Children (CFMT-C): A new tool for measuring face recognition skills in childhood. *Neuropsychologia*, *62*, 60–67. https://doi.org/10.1016/j.neuropsychologia.2014.07.008

Cui, X., Jeter, C. B., Yang, D., Montague, P. R., & David, M. (2007). Vividness of mental imagery: individual variability can be measured objectively. *Vision Research*, *47*(4), 474–478. https://doi.org/10.1016/j.visres.2006.11.013

D'Argembeau, A., & Van der Linden, M. (2006). Individual differences in the phenomenology of mental time travel: The effect of vivid visual imagery and emotion regulation strategies. *Consciousness and Cognition*, *15*(2), 342–350. https://doi.org/10.1016/j.concog.2005.09.001

Dalton, P. (2000). Psychophysical and behavioral characteristics of olfactory adaptation. *Chemical Senses*, *25*(4), 487–492. https://doi.org/10.1093/chemse/25.4.487

Dance, C. J., Ipser, A., & Simner, J. (2022). The prevalence of aphantasia (imagery weakness) in the general population. *Consciousness and Cognition*, *97*, 103243. https://doi.org/10.1016/j.concog.2021.103243

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

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

Daselaar, S. M., Porat, Y., Huijbers, W., & Pennartz, C. M. A. (2010). Modality-specific and modality-independent components of the human imagery system. *NeuroImage*, *52*(2), 677–685. https://doi.org/10.1016/j.neuroimage.2010.04.239

Davis, J. P., Gibson, S., & Solomon, C. J. (2014). The positive influence of creating a holistic facial composite on video line-up identification. *Applied Cognitive Psychology*, *28*(5), 634–639. https://doi.org/10.1002/acp.3045

Davis, J. P., Sulley, L., Solomon, C. J., & Gibson, S. (2010). A comparison of individual and morphed facial composites created using different systems. *2010 International Conference on Emerging Security Technologies*, 56–60. https://doi.org/10.1109/EST.2010.29

Davis, J. P., Thorniley, S., Gibson, S., & Solomon, C. J. (2016). Holistic facial composite construction and subsequent lineup identification accuracy: Comparing adults and children. *The Journal of Psychology: Interdisciplinary and Applied*, *150*(1), 102–118. https://doi.org/10.1080/00223980.2015.1009867

Davis, P. E., Simon, H., Meins, E., & Robins, D. L. (2018). Imaginary Companions in Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *48*, 2790–2799. https://doi.org/10.1007/s10803-018-3540-y

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*, 10022. 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*, 105192. https://doi.org/10.1016/j.cognition.2022.105192

Dawson, G., & Fernald, M. (1987). Perspective-taking ability and its relationship to the social behavior of autistic children. *Journal of Autism and Developmental Disorders*, *17*, 487–498. https://doi.org/10.1007/BF01486965

de Vito, S., & Bartolomeo, P. (2016). Refusing to imagine? On the possibility of psychogenic aphantasia. A commentary on Zeman et al. (2015). *Cortex*, *74*, 334–335. https://doi.org/10.1016/j.cortex.2015.06.013

Decety, J., & Grèzes, J. (2006). The power of simulation: Imagining one's own and other's behavior. *Brain Research*, *1079*(1), 4–14. https://doi.org/10.1016/j.brainres.2005.12.115

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*, 781. https://doi.org/10.3389/fpsyg.2014.00781

Dijkstra, N., Ambrogioni, L., Vidaurre, D., & van Gerven, M. (2020). Neural dynamics of perceptual inference and its reversal during imagery. *ELife*, *9*, 1–19. https://doi.org/10.7554/eLife.53588

Dijkstra, N., Bosch, S. E., & van Gerven, M. A. J. (2017). Vividness of visual imagery depends on the neural overlap with perception in visual areas. *Journal of Neuroscience*, *37*(5), 1367–1373. https://doi.org/10.1523/JNEUROSCI.3022-16.2016

Dijkstra, N., Bosch, S. E., & van Gerven, M. A. J. (2019). Shared Neural Mechanisms of Visual Perception and Imagery. *Trends in Cognitive Sciences*, *23*(5), 423–434. https://doi.org/10.1016/j.tics.2019.02.004

Dixon, M. J., Smilek, D., & Merikle, P. M. (2004). Not all synaesthetes are created equal: Projector versus associator synaesthetes. *Cognitive, Affective and Behavioral Neuroscience*, *4*(3), 335–343. https://doi.org/10.3758/CABN.4.3.335

Djordjevic, J., Zatorre, R. J., Petrides, M., Boyle, J. A., & Jones-Gotman, M. (2005). Functional neuroimaging of odor imagery. *NeuroImage*, *24*(3), 791–801. https://doi.org/10.1016/j.neuroimage.2004.09.035

Dobson, M., & Markham, R. (1993). Imagery ability and source monitoring: Implications for eyewitness memory. *British Journal of Psychology*, *84*(1), 111–

118. https://doi.org/10.1111/j.2044-8295.1993.tb02466.x

Downey, J. E. (1911). A Case of Colored Gustation. *The American Journal of Psychology*, *22*(4), 528–539. https://doi.org/10.2307/1412797

Dowsett, A. J., Sandford, A., & Burton, A. M. (2016). Face learning with multiple images leads to fast acquisition of familiarity for specific individuals. *The Quarterly Journal of Experimental Psychology*, *69*(1), 1–10. https://doi.org/10.1080/17470218.2015.1017513

Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia*, *44*(4), 576–585. https://doi.org/10.1016/j.neuropsychologia.2005.07.001

Duchaine, B., & Weidenfeld, A. (2003). An evaluation of two commonly used tests of unfamiliar face recognition. *Neuropsychologia*, *41*(6), 713–720. https://doi.org/10.1016/S0028-3932(02)00222-1

Durndell, A. J., & Wetherick, N. E. (1975). Reported Imagery and Social Desirability. *Perceptual and Motor Skills*, *41*(3), 987–992. https://doi.org/10.2466/pms.1975.41.3.987

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

Epstein, J., Santo, R. M., & Guillemin, F. (2015). A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *Journal of Clinical Epidemiology*, *68*(4), 435–441. https://doi.org/10.1016/j.jclinepi.2014.11.021

Eriksson, K. (2013). Autism-spectrum traits predict humor styles in the general population. *Humor*, *26*(3), 461–475. https://doi.org/10.1515/humor-2013-0030

Esposito, G., Dellantonio, S., Mulatti, C., & Job, R. (2016). Axiom, anguish, and Amazement: How autistic traits modulate emotional mental imagery. *Frontiers in Psychology*, *7*, 757. https://doi.org/10.3389/fpsyg.2016.00757

Evans, B. J. W., & Stevenson, S. J. (2008). The Pattern Glare Test: A review and determination of normative values. *Ophthalmic and Physiological Optics*, *28*(4), 295–309. https://doi.org/10.1111/j.1475-1313.2008.00578.x

Fadiga, L., Buccino, G., Craighero, L., Fogassi, L., Gallese, V., & Pavesi, G. (1999). Corticospinal excitability is specifically modulated by motor imagery: a magnetic

stimulation study. *Neuropsychologia*, *37*(2), 147–158. https://doi.org/10.1016/S0028-3932(98)00089-X

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

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

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review*, *105*(3), 482–498. https://doi.org/10.1037//0033-295x.105.3.482

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

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Fletcher, P. C., Frith, C. D., Baker, S. C., Shallice, T., Frackowiak, R. S. J., & Dolan, R. J. (1995). The Mind's Eye - Precuneus Activation in Memory-Related Imagery. *NeuroImage*, *2*(3), 195–200. https://doi.org/10.1006/nimg.1995.1025

Fodarella, C., Kuivaniemi-Smith, H., Gawrylowicz, J., & Frowd, C. D. (2015). Forensic procedures for facial-composite construction. *Journal of Forensic Practice*, *17*(4), 259–270. https://doi.org/10.1108/JFP-10-2014-0033

Frith, U. (2001). Mind blindness and the brain in autism. *Neuron*, *32*(6), 969–979. https://doi.org/10.1016/S0896-6273(01)00552-9

Frowd, C. D., Erickson, W. B., Lampinen, J. M., Skelton, F. C., McIntyre, A. H., & Hancock, P. J. B. (2015). A decade of evolving composites: regression- and meta-analysis. *Journal of Forensic Practice*, *17*(4), 319–334. https://doi.org/10.1108/JFP-08-2014-0025

Frowd, C. D., Portch, E., Killeen, A., Mullen, L., Martin, A. J., & Hancock, P. J. B. (2019). EvoFIT Facial composite images: A detailed assessment of impact on forensic practitioners, police investigators, victims, witnesses, offenders and the media. *2019 Eighth International Conference on Emerging Security Technologies (EST)*, 1–7. https://doi.org/10.1109/EST.2019.8806211

Frowd, C. D., Skelton, F., Hepton, G., Holden, L., Minahil, S., Pitchford, M., …

Hancock, P. J. B. (2013). Whole-face procedures for recovering facial images from memory. *Science and Justice*, *53*(2), 89–97. https://doi.org/10.1016/j.scijus.2012.12.004

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*, *105*, 26–40. https://doi.org/10.1016/j.cortex.2017.09.014

Galton, F. (1880). Mind a quarterly review of psychology and philosophy. *Mind*, *os-V*(19), 301–318. https://doi.org/10.1093/mind/os-V.19.301

Ganis, G., & Schendan, H. E. (2008). Visual mental imagery and perception produce opposite adaptation effects on early brain potentials. *NeuroImage*, *42*(4), 1714–1727. https://doi.org/10.1016/j.neuroimage.2008.07.004

Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2004). Brain areas underlying visual mental imagery and visual perception: An fMRI study. *Cognitive Brain Research*, *20*(2), 226–241. https://doi.org/10.1016/j.cogbrainres.2004.02.012

Gauthier, S., Anzalone, S. M., Cohen, D., Zaoui, M., Chetouani, M., Villa, F., … Xavier, J. (2018). Behavioral own-body-transformations in children and adolescents with typical development, autism spectrum disorder, and developmental coordination disorder. *Frontiers in Psychology*, *9*, 676. https://doi.org/10.3389/fpsyg.2018.00676

Gilbert, A. N., Voss, M. M., & Kroll, J. J. (1997). Vividness of olfactory mental imagery: Correlations with sensory response and consumer behavior. *Chemical Senses*, *22*, 686.

González, M. A., Campos, A., & Pérez, M. J. (1997). Mental imagery and creative thinking. *Journal of Psychology*, *131*(4), 357–364. https://doi.org/10.1080/00223989709603521

Grandin, T. (2009). How does visual thinking work in the mind of a person with autism? A personal account. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1522), 1437–1442. https://doi.org/10.1098/rstb.2008.0297

Gray, K. L. H., Bird, G., & Cook, R. (2017). Robust associations between the 20-item prosopagnosia index and the Cambridge face memory test in the general population. *Royal Society Open Science*, *4*(3), 160923. https://doi.org/10.1098/rsos.160923

Green, S. A., Hernandez, L., Tottenham, N., Krasileva, K., Bookheimer, S. Y., &

Dapretto, M. (2015). Neurobiology of sensory overresponsivity in youth with autism spectrum disorders. *JAMA Psychiatry*, *72*(8), 778–786. https://doi.org/10.1001/jamapsychiatry.2015.0737

Green, S. A., Rudie, J. D., Colich, N. L., Wood, J. J., Shirinyan, D., Hernandez, L., … Bookheimer, S. Y. (2013). Overreactive brain responses to sensory stimuli in youth with autism spectrum disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, *52*(11), 1158–1172. https://doi.org/10.1016/j.jaac.2013.08.004

Grèzes, J., & Decety, J. (2000). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: A meta-analysis. *Human Brain Mapping*, *12*(1), 1–19. https://doi.org/10.1002/1097-0193(200101)12:1<1::AID-HBM10>3.0.CO;2-V

Griffin, J. W., Bauer, R., & Scherf, K. S. (2021). A Quantitative Meta-Analysis of Face Recognition Deficits in Autism: 40 Years of Research. *Psychological Bulletin*, *147*(3), 268–292. https://doi.org/10.1037/bul0000310

Grueter, M., Grueter, T., Bell, V., Horst, J., Laskowski, W., Sperling, K., … Kennerknecht, I. (2007). Hereditary prosopagnosia: The first case series. *Cortex*, *43*(6), 734–749. https://doi.org/10.1016/S0010-9452(08)70502-1

Grüter, T., Grüter, M., Bell, V., & Carbon, C. C. (2009). Visual mental imagery in congenital prosopagnosia. *Neuroscience Letters*, *453*(3), 135–140. https://doi.org/10.1016/j.neulet.2009.02.021

Hanakawa, T., Dimyan, M. A., & Hallett, M. (2008). Motor planning, imagery, and execution in the distributed motor network: A time-course study with functional MRI. *Cerebral Cortex*, *18*(12), 2775–2788. https://doi.org/10.1093/cercor/bhn036

Harrell, F. E. (2021). *Hmisc*. Retrieved from https://cran.r-project.org/package=Hmisc

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, *458*(7238), 632–635. https://doi.org/10.1038/nature07832.

Hashimoto, R., & Rothwell, J. C. (1999). Dynamic changes in corticospinal excitability during motor imagery. *Experimental Brain Research*, *125*(1), 75–81. https://doi.org/10.1007/s002210050660

Hassabis, D., Kumaran, D., & Maguire, E. A. (2007). Using imagination to understand the neural basis of episodic memory. *Journal of Neuroscience*, *27*(52), 14365–14374. https://doi.org/10.1523/JNEUROSCI.4549-07.2007

Havlik, A. M., Carmichael, D. A., & Simner, J. (2015). Do sequence-space synaesthetes have better spatial imagery skills? Yes, but there are individual differences. *Cognitive Processing*, *16*, 245–253. https://doi.org/10.1007/s10339-015-0657-1

Herholz, S. C., Halpern, A. R., & Zatorre, R. J. (2012). Neuronal correlates of perception, imagery, and memory for familiar tunes. *Journal of Cognitive Neuroscience*, *24*(6), 1382–1397. https://doi.org/10.1162/jocn_a_00216

Hétu, S., Grégoire, M., Saimpont, A., Coll, M. P., Eugène, F., Michon, P. E., & Jackson, P. L. (2013). The neural network of motor imagery: An ALE meta-analysis. *Neuroscience and Biobehavioral Reviews*, *37*(5), 930–949. https://doi.org/10.1016/j.neubiorev.2013.03.017

Hinwar, R. P., & Lambert, A. J. (2021). Anauralia: The Silent Mind and Its Association With Aphantasia. *Frontiers in Psychology*, *12*, 744213. https://doi.org/10.3389/fpsyg.2021.744213

Hirsch, C. R., Clark, D. M., & Mathews, A. (2006). Imagery and Interpretations in Social Phobia: Support for the Combined Cognitive Biases Hypothesis. *Behavior Therapy*, *37*, 223–236. https://doi.org/10.1016/j.beth.2006.02.001

Hirsch, C. R., & Holmes, E. A. (2007). Mental imagery in anxiety disorders. *Psychiatry*, *6*(4), 161–165. https://doi.org/10.1016/j.mppsy.2007.01.005

Hirschfeld, G., Feldker, K., & Zwitserlood, P. (2012). Listening to “flying ducks”: Individual differences in sentence-picture verification investigated with ERPs. *Psychophysiology*, *49*(3), 312–321. https://doi.org/10.1111/j.1469-8986.2011.01315.x

Hiscock, M. (1978). Imagery assessment through self-report: What do imagery questionnaires measure? *Journal of Consulting and Clinical Psychology*, *46*(2), 223–230. https://doi.org/10.1037/0022-006X.46.2.223

Holmes, E. A., Crane, C., Fennell, M. J. V., & Williams, J. M. G. (2007). Imagery about suicide in depression-"Flash-forwards"? *Journal of Behavior Therapy and Experimental Psychiatry*, *38*(4), 423–434. https://doi.org/10.1016/j.jbtep.2007.10.004

Holmes, E. A., & Mathews, A. (2005). Mental imagery and emotion: A special relationship? *Emotion*, *5*(4), 489–497. https://doi.org/10.1037/1528-3542.5.4.489

Holmes, E. A., & Mathews, A. (2010). Mental imagery in emotion and emotional disorders. *Clinical Psychology Review*, *30*(3), 349–362. https://doi.org/10.1016/j.cpr.2010.01.001

Holt, S., & Yuill, N. (2014). Facilitating other-awareness in low-functioning children with autism and typically-developing preschoolers using dual-control technology. *Journal of Autism and Developmental Disorders*, *44*(1), 236–248. https://doi.org/10.1007/s10803-013-1868-x

Home Office. (2017). *Police and Criminal Evidence Act 1984 (PACE) Code D Revised Code of Practice for the identification of persons by Police Officers* (pp. 1–53). pp. 1–53. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/903812/pace-code-d-2017.pdf

Hopkins, Z. L., Yuill, N., & Branigan, H. P. (2022). Autistic children's language imitation shows reduced sensitivity to ostracism. *Journal of Autism and Developmental Disorders*, *52*(5), 1929–1941. https://doi.org/10.1007/s10803-021-05041-5

Horder, J., Wilson, C. E., Mendez, M. A., & Murphy, D. G. (2014). Autistic traits and abnormal sensory experiences in adults. *Journal of Autism and Developmental Disorders*, *44*, 1461–1469. https://doi.org/10.1007/s10803-013-2012-7

Huang, J., Cooper, T. G., Satana, B., Kaufman, D. I., & Cao, Y. (2003). Visual distortion provoked by a stimulus in migraine associated with hyperneuronal activity. *Headache*, *43*(6), 664–671. https://doi.org/10.1046/j.1526-4610.2003.03110.x

Huang, J., Zong, X., Wilkins, A., Jenkins, B., Bozoki, A., & Cao, Y. (2011). FMRI evidence that precision ophthalmic tints reduce cortical hyperactivation in migraine. *Cephalalgia*, *31*(8), 925–936. https://doi.org/10.1177/0333102411409076

Hughes, J. E. A., Ward, J., Gruffydd, E., Baron-Cohen, S., Smith, P., Allison, C., & Simner, J. (2018). Savant syndrome has a distinct psychological profile in autism. *Molecular Autism*, *9*, 53. https://doi.org/10.1186/s13229-018-0237-1

Hyman, I. E., & Pentland, J. (1996). The role of mental imagery in the creation of false childhood memories. *Journal of Memory and Language*, *35*(2), 101–117. https://doi.org/10.1006/jmla.1996.0006

Ibáñez-Marcelo, E., Campioni, L., Phinyomark, A., Petri, G., & Santarcangelo, E. L. (2019). Topology highlights mesoscopic functional equivalence between imagery and perception: The case of hypnotizability. *NeuroImage*, *200*(May), 437–449. https://doi.org/10.1016/j.neuroimage.2019.06.044

Inquisit 5. (2016). Retrieved from www.millisecond.com

Ipser, A., Ward, J., & Simner, J. (2020). The MULTISENSE Test of Lexical–Gustatory Synaesthesia: An automated online diagnostic. *Behavior Research Methods*, *52*, 544–560. https://doi.org/10.3758/s13428-019-01250-0

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. https://doi.org/10.1111/j.2044-8295.1994.tb02536.x

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

Jarrold, C., Boucher, J., & Smith, P. K. (1996). Generativity deficits in pretend play in autism. *British Journal of Developmental Psychology*, *14*(3), 275–300. https://doi.org/10.1111/j.2044-835X.1996.tb00706.x

Jastreboff, P. J., & Jastreboff, M. M. (2015). Chapter 21 - Decreased sound tolerance: hyperacusis, misophonia, diplacousis, and polyacousis. In M. J. Aminoff, F. Boller, & D. F. Swaab (Eds.), *The Human Auditory System* (pp. 375–387). https://doi.org/https://doi.org/10.1016/B978-0-444-62630-1.00021-4

Joffe, V. L., Cain, K., & Marić, N. (2007). Comprehension problems in children with specific language impairment: Does mental imagery training help? *International Journal of Language and Communication Disorders*, *42*(6), 648–664. https://doi.org/10.1080/13682820601084402

Kana, R. K., Keller, T. A., Cherkassky, V. L., Minshew, N. J., & Just, M. A. (2006). Sentence comprehension in autism: thinking in pictures with decreased functional connectivity. *Brain*, *129 (Pt 9)*, 2484–2493. https://doi.org/10.1093/brain/awl164

Kay, L., Keogh, R., Andrillon, T., & Pearson, J. (2021). 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

Keogh, R., Bergmann, J., & Pearson, J. (2020). Cortical excitability controls the strength of mental imagery. *ELife*, *9*, e50232. https://doi.org/10.7554/eLife.50232

Keogh, R., & Pearson, J. (2011). Mental imagery and visual working memory. *PLoS ONE*, *6*(12), e29221. https://doi.org/10.1371/journal.pone.0029221

Keogh, R., & Pearson, J. (2014). The sensory strength of voluntary visual imagery predicts visual working memory capacity. *Journal of Vision*, *14*(12), 7. https://doi.org/10.1167/14.12.7

Keogh, R., & Pearson, J. (2018). The blind mind: No sensory visual imagery in aphantasia. *Cortex*, *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. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1817), 20190688. https://doi.org/10.1098/rstb.2019.0688

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

Kessler, K., & Wang, H. (2012). Spatial perspective taking is an embodied process, but not for everyone in the same way: Differences predicted by sex and social skills score. *Spatial Cognition and Computation*, *12*(2–3), 133–158. https://doi.org/10.1080/13875868.2011.634533

Kirby, J. R., Moore, P. J., & Schofield, N. J. (1988). Verbal and visual learning styles. *Contemporary Educational Psychology*, *13*(2), 169–184. https://doi.org/10.1016/0361-476X(88)90017-3

Knäuper, B., McCollam, A., Rosen-Brown, A., Lacaille, J., Kelso, E., & Roseman, M. (2011). Fruitful plans: Adding targeted mental imagery to implementation intentions increases fruit consumption. *Psychology and Health*, *26*(5), 601–617. https://doi.org/10.1080/08870441003703218

Knight, K., Milton, F., & Zeman, A. (2022). Memory without Imagery: No Evidence of Visual Working Memory Impairment in People with Aphantasia. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *44*. Retrieved from https://escholarship.org/uc/item/0b16s06v

Kobayashi, M., Sasabe, T., Shigihara, Y., Tanaka, M., & Watanabe, Y. (2011). Gustatory imagery reveals functional connectivity from the prefrontal to insular cortices traced with magnetoencephalography. *PLoS ONE*, *6*(7), e21736. https://doi.org/10.1371/journal.pone.0021736

Kobayashi, M., Takeda, M., Hattori, N., Fukunaga, M., Sasabe, T., Inoue, N., … Watanabe, Y. (2004). Functional imaging of gustatory perception and imagery: "Top-down" processing of gustatory signals. *NeuroImage*, *23*(4), 1271–1282. https://doi.org/10.1016/j.neuroimage.2004.08.002

Koenig-Robert, R., & Pearson, J. (2019). Decoding the contents and strength of imagery before volitional engagement. *Scientific Reports*, *9*, 3504.

https://doi.org/10.1038/s41598-019-39813-y

Königsmark, V. T., Bergmann, J., & Reeder, R. R. (2021). The Ganzflicker experience: High probability of seeing vivid and complex pseudo-hallucinations with imagery but not aphantasia. *Cortex*, *141*, 522–534. https://doi.org/10.1016/j.cortex.2021.05.007

Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The Case for Mental Imagery*. New York: Oxford University Press.

Kunda, M., & Goel, A. K. (2008). How thinking in pictures can explain many characteristic behaviors of autism. *2008 IEEE 7th International Conference on Development and Learning, Monterey, CA*, 304–309. https://doi.org/10.1109/DEVLRN.2008.4640847

Kunda, M., & Goel, A. K. (2011). Thinking in pictures as a cognitive account of autism. *Journal of Autism and Developmental Disorders*, *41*(9), 1157–1177. https://doi.org/10.1007/s10803-010-1137-1

Lambert, A. J., & Sibley, C. G. (2022). On the importance of consistent terminology for describing sensory imagery and its absence: A response to Monzel et al. (2022). *Cortex*, *152*, 153–156. https://doi.org/10.1016/j.cortex.2022.03.012

Leclerc, M. P., Kellermann, T., Freiherr, J., Clemens, B., Habel, U., & Regenbogen, C. (2019). Externalization errors of olfactory source monitoring in healthy controls — An fMRI study. *Chemical Senses*, *44*(8), 593–605. https://doi.org/10.1093/chemse/bjz055

Lima, C. F., Lavan, N., Evans, S., Agnew, Z., Halpern, A. R., Shanmugalingam, P., … Scott, S. K. (2015). Feel the noise: Relating individual differences in auditory imagery to the structure and function of sensorimotor systems. *Cerebral Cortex*, *25*(11), 4638–4650. https://doi.org/10.1093/cercor/bhv134

Lobmaier, J. S., & Mast, F. W. (2008). Face imagery is based on featural representations. *Experimental Psychology*, *55*(1), 47–53. https://doi.org/10.1027/1618-3169.55.1.47

Logie, R. H., Pernet, C. R., Buonocore, A., & Della Sala, S. (2011). Low and high imagers activate networks differentially in mental rotation. *Neuropsychologia*, *49*(11), 3071–3077. https://doi.org/10.1016/j.neuropsychologia.2011.07.011

Low, J., Goddard, E., & Melser, J. (2009). Generativity and imagination in autism spectrum disorder: Evidence from individual differences in children's impossible entity drawings. *British Journal of Developmental Psychology*, *27*(2), 425–444.

https://doi.org/10.1348/026151008X334728

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

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

Mazza, M., Mariano, M., Peretti, S., Masedu, F., Pino, M. C., & Valenti, M. (2017). The Role of Theory of Mind on Social Information Processing in Children With Autism Spectrum Disorders: A Mediation Analysis. *Journal of Autism and Developmental Disorders*, *47*, 1369–1379. https://doi.org/10.1007/s10803-017-3069-5

McCrae, R. R., & Costa, P. T. (1987). Validation of the Five-Factor Model of Personality Across Instruments and Observers. *Journal of Personality and Social Psychology*, *52*(1), 81–90. https://doi.org/10.1037/0022-3514.52.1.81

McGeown, W. J., Venneri, A., Kirsch, I., Nocetti, L., Roberts, K., Foan, L., & Mazzoni, G. (2012). Suggested visual hallucination without hypnosis enhances activity in visual areas of the brain. *Consciousness and Cognition*, *21*(1), 100–116. https://doi.org/10.1016/j.concog.2011.10.015

McKelvie, S. J. (1994). The Vividness of Visual Imagery Questionnaire as a predictor of facial recognition memory performance. *British Journal of Psychology*, *85*(1), 93–104. https://doi.org/10.1111/j.2044-8295.1994.tb02510.x

McKone, E., Martini, P., & Nakayama, K. (2001). Categorical perception of face identity in noise isolates configural processing. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(3), 573–599. https://doi.org/10.1037/0096-1523.27.3.573

Mealor, A. D., Simner, J., Rothen, N., Carmichael, D. A., & Ward, J. (2016). Different dimensions of cognitive style in typical and atypical cognition: New evidence and a new measurement tool. *PLoS ONE*, *11*(5), e0155483. https://doi.org/10.1371/journal.pone.0155483

Meier, B., & Rothen, N. (2013). Grapheme-color synaesthesia is associated with a distinct cognitive style. *Frontiers in Psychology*, *4*, 632. https://doi.org/10.3389/fpsyg.2013.00632

Milne, E. (2011). Increased intra-participant variability in children with autistic spectrum disorders: Evidence from single-trial analysis of evoked EEG. *Frontiers*

in Psychology, *2*, 51. https://doi.org/10.3389/fpsyg.2011.00051

Milton, F., Fulford, J., Dance, C., Gaddum, J., Heuerman-Williamson, B., Jones, K., … Zeman, A. (2021). Behavioral and Neural Signatures of Visual Imagery Vividness Extremes: Aphantasia vs. Hyperphantasia. *Cerebral Cortex Communications*, *2*(2), tgab035. https://doi.org/10.1093/texcom/tgab035

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

Monzel, M., Mitchell, D., Macpherson, F., Pearson, J., & Zeman, A. (2022a). Aphantasia, dysikonesia, anauralia: call for a single term for the lack of mental imagery – Commentary on Dance et al. (2021) and Hinwar and Lambert (2021). *Cortex*. https://doi.org/10.1016/j.cortex.2022.02.002

Monzel, M., Mitchell, D., Macpherson, F., Pearson, J., & Zeman, A. (2022b). Proposal for a consistent definition of aphantasia and hyperphantasia: A response to Lambert and Sibley (2022) and Simner and Dance (2022). *Cortex*, *152*, 74–76. https://doi.org/10.1016/j.cortex.2022.04.003

Monzel, M., Vetterlein, A., & Reuter, M. (2021). Memory deficits in aphantasics are not restricted to autobiographical memory – Perspectives from the Dual Coding Approach. *Journal of Neuropsychology*, *16*(2), 444–461. https://doi.org/10.1111/jnp.12265

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*, *Epub ahead*, 36463494. https://doi.org/10.1111/sjop.12887

Morey, R. D., & Rouder, J. N. (2018). *BayesFactor: Computation of Bayes Factors for Common Designs*. Retrieved from https://cran.r-project.org/package=BayesFactor

Mountcastle, V. B. (1997). The columnar organization of the neocortex. *Brain*, *120*(4), 701–722. https://doi.org/10.1093/brain/120.4.701

Mulleners, W. M., Chronicle, E. P., Palmer, J. E., Koehler, P. J., & Vredeveld, J. W. (2001). Visual cortex excitability in migraine with and without aura. *Headache*, *41*(6), 565–572. https://doi.org/10.1046/j.1526-4610.2001.041006565.x

Naselaris, T., Olman, C. A., Stansbury, D. E., Ugurbil, K., & Gallant, J. L. (2015). A voxel-wise encoding model for early visual areas decodes mental images of remembered scenes. *NeuroImage*, *105*, 215–228. https://doi.org/10.1016/j.neuroimage.2014.10.018

Needell, C., & Bainbridge, W. (2022). Embracing New Techniques in Deep Learning for Estimating Image Memorability. *Computational Brain & Behavior*, *5*, 168–184. https://doi.org/10.1007/s42113-022-00126-5

Neufeld, J., Roy, M., Zapf, A., Sinke, C., Emrich, H. M., Prox-Vagedes, V., … Zedler, M. (2013). Is synesthesia more common in patients with Asperger syndrome? *Frontiers in Human Neuroscience*, *7*, 847. https://doi.org/10.3389/fnhum.2013.00847

OECD. (2021). Adult education level (indicator). https://doi.org/10.1787/36bce3fe-en

Office for National Statistics. (2011). 2011 Census: Key Statistics and Quick Statistics for Local Authorities in the United Kingdom. Retrieved January 20, 2021, from https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/bulletins/keystatisticsandquickstatisticsforlocalauthoritiesintheunitedkingdom/2013-12-04#qualifications

Okuda, J., Fujii, T., Ohtake, H., Tsukiura, T., Tanji, K., Suzuki, K., … Yamadori, A. (2003). Thinking of the future and past: The roles of the frontal pole and the medial temporal lobes. *NeuroImage*, *19*(4), 1369–1380. https://doi.org/10.1016/S1053-8119(03)00179-4

Online Etymology Dictionary. (2001). *Online Etymology Dictionary*. Retrieved from https://www.etymonline.com/search?q=phantasm

Osterreith, P., & Rey, A. (1944). Le test de copie d'une figure complexe. *Arch Psychol*, *30*, 205–220.

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

Palmer, C. J., Paton, B., Enticott, P. G., & Hohwy, J. (2015). 'Subtypes' in the Presentation of Autistic Traits in the General Adult Population. *Journal of Autism and Developmental Disorders*, *45*, 1291–1301. https://doi.org/10.1007/s10803-014-2289-1

Pearson, A., Marsh, L., Hamilton, A., & Ropar, D. (2014). Spatial transformations of bodies and objects in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *44*(9), 2277–2289. https://doi.org/10.1007/s10803-014-2098-6

Pearson, J. (2014). New Directions in Mental-Imagery Research: The Binocular-Rivalry

Technique and Decoding fMRI Patterns. *Current Directions in Psychological Science*, *23*(3), 178–183. https://doi.org/10.1177/0963721414532287

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., Clifford, C. W. G., & Tong, F. (2008). The Functional Impact of Mental Imagery on Conscious Perception. *Current Biology*, *18*(13), 982–986. https://doi.org/10.1016/j.cub.2008.05.048

Pearson, J., & Kosslyn, S. M. (2015). The heterogeneity of mental representation: Ending the imagery debate. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(33), 10089–10092. https://doi.org/10.1073/pnas.1504933112

Pearson, J., Naselaris, T., Holmes, E. A., & Kosslyn, S. M. (2015). Mental Imagery: Functional Mechanisms and Clinical Applications. *Trends in Cognitive Sciences*, *19*(10), 590–602. https://doi.org/10.1016/j.tics.2015.08.003

Pearson, J., Rademaker, R. L., & Tong, F. (2011). Evaluating the mind's eye: The metacognition of visual imagery. *Psychological Science*, *22*(12), 1535–1542. https://doi.org/10.1177/0956797611417134

Pecher, D., van Dantzig, S., & Schifferstien, H. N. J. (2009). Concepts are not represented by conscious imagery. *Psychonomic Bulletin and Review*, *16*(5), 914–919. https://doi.org/10.3758/PBR.16.5.914

Perner, J., Frith, U., Leslie, A. M., & Leekam, S. R. (1989). Exploration of the Autistic Child 's Theory of Mind: Knowledge, Belief, and Communication. *Child Development*, *60*(3), 689–700. https://doi.org/10.2307/1130734

Pike, G. E., Brace, N. A., Turner, J., Ness, H., & Vredeveldt, A. (2019). Advances in facial composite technology, utilizing holistic construction, do not lead to an increase in eyewitness misidentifications compared to older feature-based systems. *Frontiers in Psychology*, *10*, 1962. https://doi.org/10.3389/fpsyg.2019.01962

Pike, G. E., Brace, N. A., Turner, J., & Vredeveldt, A. (2019). The Effect of Facial Composite Construction on Eyewitness Identification Accuracy in an Ecologically Valid Paradigm. *Criminal Justice and Behavior*, *46*(2), 319–336. https://doi.org/10.1177/0093854818811376

Plailly, J., Delon-Martin, C., & Royet, J. P. (2012). Experience induces functional reorganization in brain regions involved in odor imagery in perfumers. *Human*

Brain Mapping, *33*(1), 224–234. https://doi.org/10.1002/hbm.21207

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

Price, M. C. (2009). Spatial forms and mental imagery. *Cortex*, *45*(10), 1229–1245. https://doi.org/10.1016/j.cortex.2009.06.013

Pylyshyn, Z. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, *80*(1), 1–24. https://doi.org/10.1037/h0034650

Pylyshyn, Z. (2002). Mental Imagery: In search of a theory. *Behavioral and Brain Sciences*, *25*(2), 157–182. https://doi.org/10.1017/s0140525x02000043

R Core Team. (2018). *R: A Language and Environment for Statistical Computing*. Retrieved from https://www.r-project.org/

Rademaker, R. L., & Pearson, J. (2012). Training visual imagery: Improvements of metacognition, but not imagery strength. *Frontiers in Psychology*, *3*, 224. https://doi.org/10.3389/fpsyg.2012.00224

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

Richardson, A. (1969). *Mental imagery*. New York: Springer.

Richardson, A. (1978). Subject, task, and tester variables associated with initial eye movement responses. *Journal of Mental Imagery*, *2*(1), 85–100.

Richardson, J. T. E. (1995). Gender differences in the Vividness of Visual Imagery Questionnaire: A meta-analysis. *Journal of Mental Imagery*, *19*(3–4), 177–187.

Richler, J. J., & Gauthier, I. (2014). A meta-analysis and review of holistic face processing. *Psychological Bulletin*, *140*(5), 1281–1302. https://doi.org/10.1037/a0037004

Rinaldi, L. J., Simner, J., Koursarou, S., & Ward, J. (2022). Autistic traits, emotion regulation, and sensory sensitivities in children and adults with Misophonia. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-022-05623-x

Ritchie, K. L., & Burton, A. M. (2017). Learning faces from variability. *The Quarterly Journal of Experimental Psychology*, *70*(5), 897–905.

https://doi.org/10.1080/17470218.2015.1136656

Roberts, R., Callow, N., Hardy, L., Markland, D., & Bringer, J. (2008). Movement Imagery Ability: Development and Assessment of a Revised Version of the Vividness of Movement Imagery Questionnaire. *Journal of Sport and Exercise Psychology*, *30*(2), 200–221. https://doi.org/10.1123/jsep.30.2.200

Robertson, A. E., & Simmons, D. R. (2013). The relationship between sensory sensitivity and autistic traits in the general population. *Journal of Autism and Developmental Disorders*, *43*(4), 775–784. https://doi.org/10.1007/s10803-012-1608-7

Robertson, A. E., & Simmons, D. R. (2015). The sensory experiences of adults with autism spectrum disorder: A qualitative analysis. *Perception*, *44*(5), 569–586. https://doi.org/10.1068/p7833

Rothen, N., Seth, A. K., Witzel, C., & Ward, J. (2013). Diagnosing synaesthesia with online colour pickers: Maximising sensitivity and specificity. *Journal of Neuroscience Methods*, *215*(1), 156–160. https://doi.org/10.1016/j.jneumeth.2013.02.009

Rouw, R., & Scholte, H. S. (2007). Increased structural connectivity in grapheme-color synesthesia. *Nature Neuroscience*, *10*, 792–797. https://doi.org/10.1038/nn1906

Rouw, R., & Scholte, H. S. (2010). Neural basis of individual differences in synesthetic experiences. *Journal of Neuroscience*, *30*(18), 6205–6213. https://doi.org/10.1523/JNEUROSCI.3444-09.2010

Rouw, R., Scholte, H. S., & Colizoli, O. (2011). Brain areas involved in synaesthesia: A review. *Journal of Neuropsychology*, *5*(2), 214–242. https://doi.org/10.1111/j.1748-6653.2011.02006.x

Rubenstein, J. L. R., & Merzenich, M. M. (2003). Model of autism: Increased ratio of excitation/inhibition in key neural systems. *Genes, Brain and Behavior*, Vol. 2, pp. 255–267. https://doi.org/10.1034/j.1601-183X.2003.00037.x

Russell, R., Duchaine, B., & Nakayama, K. (2009). Super-recognizers: People with extraordinary face recognition ability. *Psychonomic Bulletin & Review, 16*(2), 252–257. https://doi.org/10.3758/PBR.16.2.252

Santabárbara, J., Lasheras, I., Lipnicki, D. M., Bueno-Notivol, J., Pérez-Moreno, M., López-Antón, R., … Gracia-García, P. (2021). Prevalence of anxiety in the COVID-19 pandemic: An updated meta-analysis of community-based studies. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *109*, 110207.

https://doi.org/10.1016/j.pnpbp.2020.110207

Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: Remembering the past and imagining the future. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481), 773–786. https://doi.org/10.1098/rstb.2007.2087

Schmidt, T. T., Ostwald, D., & Blankenburg, F. (2014). Imaging tactile imagery: Changes in brain connectivity support perceptual grounding of mental images in primary sensory cortices. *NeuroImage*, *98*, 216–224. https://doi.org/10.1016/j.neuroimage.2014.05.014

Schröder, A., Vulink, N., & Denys, D. (2013). Misophonia: Diagnostic Criteria for a New Psychiatric Disorder. *PloS One*, *8*(1), e54706. https://doi.org/10.1371/journal.pone.0054706

Scott, F. J., & Baron-Cohen, S. (1996). Imagining real and unreal things: Evidence of a dissociation in autism. *Journal of Cognitive Neuroscience*, *8*(4), 371–382. https://doi.org/10.1162/jocn.1996.8.4.371

Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science*, *20*(2), 207–214. https://doi.org/10.1111/j.1467-9280.2009.02276.x

Seron, X., Pesenti, M., Nöel, M., Deloche, G., & Cornet, J. (1992). Images of numbers, or "When 98 is upper left and 6 sky blue." *Cognition*, *44*(1–2), 159–196. https://doi.org/10.1016/0010-0277(92)90053-K

Shah, P., Gaule, A., Sowden, S., Bird, G., & Cook, R. (2015). The 20-item prosopagnosia index (PI20): A self-report instrument for identifying developmental prosopagnosia. *Royal Society Open Science*, *2*(6), 14034. https://doi.org/10.1098/rsos.140343

Shah, P., Sowden, S., Gaule, A., Catmur, C., & Bird, G. (2015). The 20 item prosopagnosia index (PI20): Relationship with the Glasgow face-matching test. *Royal Society Open Science*, *2*(11), 150305. https://doi.org/10.1098/rsos.150305

Sheehan, P. W. (1967). A shortened form of Betts' questionnaire upon mental imagery. *Journal of Clinical Psychology*, *23*(3), 386–389. https://doi.org/10.1002/1097-4679(196707)23:3<386::aid-jclp2270230328>3.0.co;2-s

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*(3972), 701–703.

Shine, J. M., Keogh, R., O'Callaghan, C., Muller, A. J., Lewis, S. J. G., & Pearson, J.

(2015). Imagine that: Elevated sensory strength of mental imagery in individuals with Parkinson's disease and visual hallucinations. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1798), 20142047. https://doi.org/10.1098/rspb.2014.2047

Simmons, D. R. (2019). Some clarifications on neural noise and sensory sensitivities in Autism. *Cognitive Neuroscience*, *10*(3), 169–171. https://doi.org/10.1080/17588928.2019.1598349

Simmons, D. R., Robertson, A. E., McKay, L. S., Toal, E., McAleer, P., & Pollick, F. E. (2009). Vision in autism spectrum disorders. *Vision Research*, *49*(22), 2705–2739. https://doi.org/10.1016/j.visres.2009.08.005

Simner, J. (2013). Why are there different types of synesthete? *Frontiers in Psychology*, *4*, 558. https://doi.org/10.3389/fpsyg.2013.00558

Simner, J. (2019). *Synaesthesia: A Very Short Introduction*. Oxford University Press.

Simner, J., & Carmichael, D. A. (2015). Is synaesthesia a dominantly female trait? *Cognitive Neuroscience*, *6*(2–3), 68–76. https://doi.org/10.1080/17588928.2015.1019441

Simner, J., & Dance, C. J. (2022). Dysikonesia or aphantasia? Understanding the impact and history of names. A reply to Monzel et al. (2022). *Cortex*, *153*, 220–223. https://doi.org/10.1016/j.cortex.2022.04.013

Simner, J., Glover, L., & Mowat, A. (2006). Linguistic determinants of word colouring in grapheme-colour synaesthesia. *Cortex*, *42*(2), 281–289. https://doi.org/10.1016/S0010-9452(08)70353-8

Simner, J., & Hubbard, E. M. (2013). Oxford Handbook of Synesthesia. In *Oxford Handbook of Synesthesia*. https://doi.org/10.1093/oxfordhb/9780199603329.001.0001

Simner, J., Koursarou, S., Rinaldi, L. J., & Ward, J. (2021). Attention, flexibility, and imagery in misophonia: Does attention exacerbate everyday disliking of sound? *Journal of Clinical and Experimental Neuropsychology*, *43*(10), 1006–1017. https://doi.org/10.1080/13803395.2022.2056581

Simner, J., Mulvenna, C., Sagiv, N., Tsakanikos, E., Witherby, S. A., Fraser, C., … Ward, J. (2006). Synaesthesia: The prevalence of atypical cross-modal experiences. *Perception*, *35*(8), 1024–1033. https://doi.org/10.1068/p5469

Skelton, R., Ludwig, C., & Mohr, C. (2009). A novel, illustrated questionnaire to distinguish projector and associator synaesthetes. *Cortex*, *45*(6), 721–729.

https://doi.org/10.1016/j.cortex.2008.02.006

Smees, R., Rinaldi, L. J., Simmons, D. R., & Simner, J. (2020). *Measuring Sensory Sensitivities in Children: The Parent-completed Glasgow Sensory Questionnaire (GSQ-P)*. Manuscript submitted for publication.

Smith, B. H., Campbell, A., Linksted, P., Fitzpatrick, B., Jackson, C., Kerr, S. M., … Morris, A. D. (2013). Cohort profile: Generation scotland: Scottish family health study (GS: SFHS). The study, its participants and their potential for genetic research on health and illness. *International Journal of Epidemiology*, *42*(3), 689–700. https://doi.org/10.1093/ije/dys084

Smith, B. H., Campbell, H., Blackwood, D., Connell, J., Connor, M., Deary, I. J., … Morris, A. D. (2006). Generation Scotland: The Scottish Family Health Study; a new resource for researching genes and heritability. *BMC Medical Genetics*, *7*, 74. https://doi.org/10.1186/1471-2350-7-74

Smith, W. G. (2008). *Does Gender Influence Online Survey Participation? A Record-Linkage Analysis of University Faculty Online Survey Response Behavior.* Unpublished manuscript.

Solomon, C. J., Gibson, S. J., & Maylin, M. (2012). EFIT-V: evolutionary algorithms and computer composites. In C. Wilkinson & C. Rynn (Eds.), *Craniofacial identification* (pp. 24–41). Cambridge: Cambridge University Press.

Solomon, C. J., Gibson, S. J., & Mist, J. J. (2013). Interactive evolutionary generation of facial composites for locating suspects in criminal investigations. *Applied Soft Computing Journal*, *13*(7), 3298–3306. https://doi.org/10.1016/j.asoc.2013.02.010

Soulières, I., Zeffiro, T. A., Girard, M. L., & Mottron, L. (2011). Enhanced mental image mapping in autism. *Neuropsychologia*, *49*(5), 848–857. https://doi.org/10.1016/j.neuropsychologia.2011.01.027

Sparing, R., Mottaghy, F. M., Ganis, G., Thompson, W. L., Töpper, R., Kosslyn, S. M., & Pascual-Leone, A. (2002). Visual cortex excitability increases during visual mental imagery - A TMS study in healthy human subjects. *Brain Research*, *938*(1–2), 92–97. https://doi.org/10.1016/S0006-8993(02)02478-2

Spiller, M. J., Harkry, L., McCullagh, F., Thoma, V., & Jonas, C. (2019). Exploring the relationship between grapheme colour-picking consistency and mental imagery. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*, 20190023. https://doi.org/10.1098/rstb.2019.0023

Spiller, M. J., & Jansari, A. S. (2008). Mental imagery and synaesthesia: Is synaesthesia

from internally-generated stimuli possible? *Cognition*, *109*(1), 143–151. https://doi.org/10.1016/j.cognition.2008.08.007

Spiller, M. J., Jonas, C. N., Simner, J., & Jansari, A. (2015). Beyond visual imagery: How modality-specific is enhanced mental imagery in synesthesia? *Consciousness and Cognition*, *31*, 73–85. https://doi.org/10.1016/j.concog.2014.10.010

Stewart, L. M., Walsh, V., & Rothwell, J. C. (2001). Motor and phosphene thresholds: A transcranial magnetic stimulation correlation study. *Neuropsychologia*, *39*(4), 415–419. https://doi.org/10.1016/S0028-3932(00)00130-5

Stewart, M. E., & Austin, E. J. (2009). The structure of the Autism-Spectrum Quotient (AQ): Evidence from a student sample in Scotland. *Personality and Individual Differences*, *47*(3), 224–228. https://doi.org/10.1016/j.paid.2009.03.004

Stinear, C. M., Byblow, W. D., Steyvers, M., Levin, O., & Swinnen, S. P. (2006). Kinesthetic, but not visual, motor imagery modulates corticomotor excitability. *Experimental Brain Research*, *168*(1–2), 157–164. https://doi.org/10.1007/s00221-005-0078-y

Szpunar, K. K., Watson, J. M., & McDermott, K. B. (2007). Neural substrates of envisioning the future. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(2), 642–647. https://doi.org/10.1073/pnas.0610082104

Takarae, Y., & Sweeney, J. (2017). Neural hyperexcitability in autism spectrum disorders. *Brain Sciences*, *7*(10), 129. https://doi.org/10.3390/brainsci7100129

Ten Eycke, K. D., & Müller, U. (2014). Brief Report: New Evidence for a Social-Specific Imagination Deficit in Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *45*, 213–220. https://doi.org/10.1007/s10803-014-2206-7

Terhune, D. B., Murray, E., Near, J., Stagg, C. J., Cowey, A., & Kadosh, R. C. (2015). Phosphene perception relates to visual cortex glutamate levels and covaries with atypical visuospatial awareness. *Cerebral Cortex*, *25*(11), 4341–4350. https://doi.org/10.1093/cercor/bhv015

Terhune, D. B., Song, S. M., & Cohen Kadosh, R. (2015). Transcranial alternating current stimulation reveals atypical 40 Hz phosphene thresholds in synaesthesia. *Cortex*, *63*, 267–270. https://doi.org/10.1016/j.cortex.2014.09.006

Terhune, D. B., Tai, S., Cowey, A., Popescu, T., & Cohen Kadosh, R. (2011). Enhanced cortical excitability in grapheme-color synesthesia and its modulation. *Current*

Biology, *21*(23), 2006–2009. https://doi.org/10.1016/j.cub.2011.10.032

Thomas, N. J. T. (1999). Are Theories of Imagery Theories of Imagination?: An Active Perception Approach to Conscious Mental Content. *Cognitive Science*, *23*(2), 207–245. https://doi.org/10.1016/S0364-0213(99)00004-X

Thomas, N. J. T. (2014). "Mental Imagery." In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy* (Fall 2014). Retrieved from https://plato.stanford.edu/archives/fall2014/entries/mental-imagery/

Thompson, R. F., & Spencer, W. A. (1966). Habituation: A model phenomenon for the study of neuronal substrates of behavior. *Psychological Review*, *73*(1), 16–43. https://doi.org/10.1037/h0022681

Thurstone, L. L., & Thurstone, T. G. (1989). *Aptitudes menfales primarias [Primary mental abilities]*. Madrid: TEA Ediciones.

Tiggemann, M., & Kemps, E. (2005). The phenomenology of food cravings: The role of mental imagery. *Appetite*, *45*(3), 305–313. https://doi.org/10.1016/j.appet.2005.06.004

Tomes, J. L., & Katz, A. N. (1999). Habitual Susceptibility to Misinformation and Individual Differences in Eyewitness Memory. *Applied Cognitive Psychology*, *11*(3), 233–251. https://doi.org/10.1002/(SICI)1099-0720(199706)11:3<233::AID-ACP447>3.0.CO;2-V

Torrance, E. P., & Ball, O. E. (1992). *Torrance Tests of Creative Thinking: Streamlined (revised) manual*. Bensenville, IL: Scholastic Testing Services.

Tree, J. J., & Wilkie, J. (2010). Face and object imagery in congenital prosopagnosia: A case series. *Cortex*, *46*(9), 1189–1198. https://doi.org/10.1016/j.cortex.2010.03.005

Tsantani, M., Vestner, T., & Cook, R. (2021). The Twenty Item Prosopagnosia Index (PI20) provides meaningful evidence of face recognition impairment. *Royal Society Open Science*, *8*(11), 202062. https://doi.org/10.1098/rsos.202062

UK Comission for Employment and Skills. (2016). UK labour market projections: 2014 to 2024. Retrieved July 26, 2021, from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/513801/Working_Futures_final_evidence_report.pdf

Van Leeuwen, T. M., Van Petersen, E., Burghoorn, F., Dingemanse, M., & Van Lier, R. (2019). Autistic traits in synaesthesia: Atypical sensory sensitivity and enhanced perception of details. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1787), 20190024. https://doi.org/10.1098/rstb.2019.0024

Vasquez, N. A., & Buehler, R. (2007). Seeing future success: Does imagery perspective influence achievement motivation? *Personality and Social Psychology Bulletin*, *33*(10), 1392–1405. https://doi.org/10.1177/0146167207304541

Visionmetric Ltd. (2019a). *EFIT6 [Computer Software]*. Retrieved from https://visionmetric.com/efit6/

Visionmetric Ltd. (2019b). *EFIT6 User Guide*. Retrieved from https://visionmetric.com/wp-content/uploads/2019/10/EFIT6_user_guide.1.4.4.12.12Sep2019-converted_compressed.pdf

Vogeley, K., & Fink, G. R. (2003). Neural correlates of the first-person-perspective. *Trends in Cognitive Sciences*, *7*(1), 38–42. https://doi.org/10.1016/S1364-6613(02)00003-7

Ward, J. (2019a). Individual differences in sensory sensitivity: A synthesizing framework and evidence from normal variation and developmental conditions. *Cognitive Neuroscience*, *10*(3), 139–157. https://doi.org/10.1080/17588928.2018.1557131

Ward, J. (2019b). Synaesthesia: A distinct entity that is an emergent feature of adaptive neurocognitive differences. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*, 20180351. https://doi.org/10.1098/rstb.2018.0351

Ward, J., Brown, P., Sherwood, J., & Simner, J. (2018). An autistic-like profile of attention and perception in synaesthesia. *Cortex*, *107*, 121–130. https://doi.org/10.1016/j.cortex.2017.10.008

Ward, J., Hoadley, C., Hughes, J. E. A., Smith, P., Allison, C., Baron-Cohen, S., & Simner, J. (2017). Atypical sensory sensitivity as a shared feature between synaesthesia and autism. *Scientific Reports*, *7*, 41155. https://doi.org/10.1038/srep41155

Ward, J., Huckstep, B., & Tsakanikos, E. (2006). Sound-Colour Synaesthesia: to What Extent Does it Use Cross-Modal Mechanisms Common to us All? *Cortex*, *42*(2), 264–280. https://doi.org/10.1016/S0010-9452(08)70352-6

Ward, J., Ipser, A., Phanvanova, E., Brown, P., Bunte, I., & Simner, J. (2018). The prevalence and cognitive profile of sequence-space synaesthesia. *Consciousness and Cognition*, *61*, 79–93. https://doi.org/10.1016/j.concog.2018.03.012

Ward, J., Li, R., Salih, S., & Sagiv, N. (2006). Varieties of grapheme-colour synaesthesia: A new theory of phenomenological and behavioural differences.

Consciousness and Cognition, *16*(4), 913–931. https://doi.org/10.1016/j.concog.2006.09.012

Ward, J., & Simner, J. (2003). Lexical-gustatory synaesthesia: Linguistic and conceptual factors. *Cognition*, *89*(3), 237–261. https://doi.org/10.1016/S0010-0277(03)00122-7

Ward, J., Simner, J., & Auyeung, V. (2005). A comparison of lexical-gustatory and grapheme-colour synaesthesia. *Cognitive Neuropsychology*, *22*(1), 28–41. https://doi.org/10.1080/02643290442000022

Watkins, N. W. (2018). (A)phantasia and severely deficient autobiographical memory: Scientific and personal perspectives. *Cortex*, *105*, 41–52. https://doi.org/10.1016/j.cortex.2017.10.010

Wechsler, D. (1997). *Wechsler memory scale III*. San Antonio (TX): The Psychological Corporation.

White, K., Ashton, R., & Law, H. (1974). Factor analyses of the shortened form of Betts' questionnaire upon mental imagery. *Australian Journal of Psychology*, *26*(3), 183–190. https://doi.org/10.1080/00049537408255229

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

Wilkins, A., Binnie, C. D., & Darby, C. E. (1980). Visually-induced seizures. *Progress in Neurobiology*, *15*(2), 85–117. https://doi.org/10.1016/0301-0082(80)90004-0

Wilkins, A., Nimmo-smith, I., Tait, A., Mcmanus, C., Sala, S. Della, Tilley, A., … Scott, S. (1984). A neurological basis for visual discomfort. *Brain*, *107*(4), 989–1017. https://doi.org/10.1093/brain/107.4.989

Willander, J., & Baraldi, S. (2010). Development of a new clarity of auditory imagery scale. *Behavior Research Methods*, *42*(3), 785–790. https://doi.org/10.3758/BRM.42.3.785

Winlove, C. I. P., Milton, F., Ranson, J., Fulford, J., MacKisack, M., Macpherson, F., & Zeman, A. (2018). The neural correlates of visual imagery: A co-ordinate-based meta-analysis. *Cortex*, *105*, 4–25. https://doi.org/10.1016/j.cortex.2017.12.014

Wood, E. T., Cummings, K. K., Jung, J., Patterson, G., Okada, N., Guo, J., … Green, S. A. (2021). Sensory over-responsivity is related to GABAergic inhibition in thalamocortical circuits. *Translational Psychiatry*, *11*, 39.

https://doi.org/10.1038/s41398-020-01154-0

Wu, J., Duan, H., Tian, X., Wang, P., & Zhang, K. (2012). The effects of visual imagery on face identification: An ERP study. *Frontiers in Human Neuroscience*, *6*, 305. https://doi.org/10.3389/fnhum.2012.00305

Yoo, S. S., Freeman, D. K., McCarthy, J. J., & Jolesz, F. A. (2003). Neural substrates of tactile imagery: a functional MRI study. *NeuroReport*, *14*(4), 581–585. https://doi.org/10.1097/00001756-200303240-00011

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, *16*, 747–759. https://doi.org/10.1068/p160747

Young, W. B., Oshinsky, M. L., Shechter, A. L., Gebeline-Myers, C., Bradley, K. C., & Wassermann, E. M. (2004). Consecutive Transcranial Magnetic Stimulation: Phosphene Thresholds in Migraineurs and Controls. *Headache*, *44*(2), 131–135. https://doi.org/10.1111/j.1526-4610.2004.04028.x

Zahradnikova, B., Duchovicova, S., & Schreiber, P. (2018). Facial composite systems: review. *Artificial Intelligence Review*, *49*, 131–152. https://doi.org/10.1007/s10462-016-9519-1

Zandt, F., Prior, M., & Kyrios, M. (2007). Repetitive behaviour in children with high functioning autism and obsessive compulsive disorder. *Journal of Autism and Developmental Disorders*, *37*(2), 251–259. https://doi.org/10.1007/s10803-006-0158-2

Zeman, A., Della Sala, S., Torrens, L. A., Gountouna, V., 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., … Winlove, C. (2020). Phantasia - The Psychological Significance of Lifelong Imagery Vividness Extremes. *Cortex*, *130*, 426–440. https://doi.org/10.1016/j.cortex.2020.04.003

Zeman, A., Onians, J., Macpherson, F., Aldworth, S., Winlove, C., & MacKisack, M. (2019). *Extreme Imagination - inside the mind's eye*. [Exhibition Catalogue].

Exhibitied at Royal Albert Memorial Museum, Exeter 30/03/19-02/06/19.

Zvyagintsev, M., Clemens, B., Chechko, N., Mathiak, K. A., Sack, A. T., & Mathiak, K. (2013). Brain networks underlying mental imagery of auditory and visual information. *European Journal of Neuroscience*, *37*(9), 1421–1434. https://doi.org/10.1111/ejn.12140

Appendices

Appendix A

Table S1

Groups means in mental imagery (with standard deviations, SDs) for Aphantasics and Controls, in each of the sense domains. The table shows independent samples t-tests with Welch correction, with bootstrapped 95% confidence intervals (BCa bootstrapping performed 1000 times). Uncorrected p values are shown, with significant differences after correcting for multiple comparisons (using the Benjamini-Hochberg False Discovery Rate method; Benjamini & Hochberg, 1995, 2000) highlighted in bold (i.e., all survive).

Sense domain	Measure	Aphantasics		Controls		*t*	*df*	*p*	*g*	BCa 95% CIs	
		M	*SD*	*M*	*SD*					Lower	Upper
Visual	Psi-Q	.45	1.18	7.70	2.02	37.25	212.11	**<.001**	4.48	6.87	7.66
Auditory	CAIS	24.66	16.31	61.73	12.60	22.26	297.89	**<.001**	2.51	33.74	40.34
	Psi-Q	1.76	2.94	7.87	2.00	21.39	287.71	**<.001**	2.39	5.51	6.68
Tactile	Betts-ad	16.85	9.73	45.85	9.54	26.09	293.09	**<.001**	3.00	26.56	31.11
	Psi-Q	1.18	2.43	7.81	2.00	26.01	299.87	**<.001**	2.95	6.14	7.13
Olfactory	VOIQ	20.82	11.31	55.96	13.60	24.13	266.85	**<.001**	2.83	32.30	37.69
	Psi-Q	.83	1.97	7.42	2.11	27.88	283.15	**<.001**	3.23	6.14	7.05
Gustatory	Betts-ad	15.61	8.41	45.52	9.34	29.00	278.73	**<.001**	3.37	27.75	31.98
	Psi-Q	.92	2.03	7.47	2.15	27.03	284.72	**<.001**	3.13	6.09	7.01
Bodily	Betts-ad	19.98	11.41	44.01	8.79	20.65	297.74	**<.001**	2.33	21.83	26.20
	Psi-Q	1.34	2.36	7.51	1.98	24.67	300.00	**<.001**	2.80	5.69	6.66
Movement	VMIQ	17.27	9.93	44.12	10.97	22.13	279.42	**<.001**	2.57	24.64	29.13
Feeling	Psi-Q	2.57	3.43	7.56	1.82	16.14	256.10	**<.001**	1.77	4.29	5.65

Table S2

Groups means in sensory sensitivity (with standard deviations, SDs) for Aphantasics and Controls, in each of the sense domains (GSQ). The table shows independent samples t-tests with Welch correction, with bootstrapped 95% confidence intervals (BCa bootstrapping performed 1000 times). Uncorrected p values are shown, with significant differences after correcting for multiple comparisons (using the Benjamini-Hochberg False Discovery Rate method; Benjamini & Hochberg, 1995, 2000) highlighted in bold (i.e., all survive).

Sense domain	Aphantasics		Controls		*t*	*df*	*p*	*g*	BCa 95% CIs	
	M	*SD*	*M*	*SD*					Lower	Upper
Visual	7.32	4.17	9.37	5.20	3.72	261.33	**<.001**	0.44	1.01	3.11
Auditory	11.95	4.63	12.09	4.65	.265	290.88	.791	0.03	-.924	1.19
Tactile	6.96	4.12	8.84	5.13	3.46	261.48	**.001**	0.41	.791	2.94
Olfactory	6.35	3.50	8.74	5.35	4.49	228.46	**<.001**	0.54	1.37	3.36
Gustatory	7.00	4.02	8.82	4.98	3.45	262.25	**.001**	0.41	.698	2.86
Proprioception	5.22	3.54	7.57	5.75	4.18	219.52	**<.001**	0.50	1.25	3.45
Vestibular	5.72	3.81	8.20	5.60	4.41	234.64	**<.001**	0.52	1.46	3.57

Results of Experiment 1 when including age as a covariate

In the two sections below (re: imagery and sensory sensitivity), we repeat our analyses of Experiment 1, but include now a covariate of age. In our group sampling, we found an unintended groupwise difference in age (aphantasics: $M = 42.35$, $SD = 15.95$;

controls $M = 37.39$, $SD = 13.83$; $t(299.72) = -2.89$, $p = .004$. To ensure this ~5 year age gap did not influence our results, we repeat our analyses here, with age as a covariate. To anticipate our findings, our overall pattern of results remained unchanged, for both imagery and sensitivity.

Do Aphantasics have Poor Imagery in Multiple Domains?

To examine if age influenced our results for imagery, we replicated our main analysis for the Psi-Q, but this time including age as a covariate. We conducted a 2x7 ANCOVA crossing group (aphantasics, controls) with sense domain (Psi-Q imagery subscales; visual, auditory, olfactory, gustatory, tactile, body sensation, and feeling), adding age as a covariate. As before, there was a significant main effect of group ($F(1, 299) = 834.51$, $p < .001$, $\eta_p^2 = .736$), a significant main effect of sense domain ($F(3.93, 1174.08) = 7.85$, $p < .001$, $\eta_p^2 = .026$; with Greenhouse-Geisser correction), and a significant interaction between the two ($F(3.93, 1174.08) = 21.35$, $p < .001$, $\eta_p^2 = .067$; with Greenhouse-Geisser correction. There was no significant effect of age ($F(1, 299) = .035$, $p = .852$, $\eta_p^2 = .000$), and no interaction between age and sense domain ($F(3.93, 1174.08) = 1.63$, $p = .165$, $\eta_p^2 = .005$; with Greenhouse-Geisser correction). Overall, our results show that when controlling for the influence of age our pattern of results for imagery remain the same.

Does Imagery Predict Sensory Sensitivity?

Next, we repeated our sensory sensitivity (GSQ) analysis, by conducting a 2x2x7 ANCOVA crossing group (aphantasic, control) with sensitivity type (hyper-, hypo-sensitivity) and sense domain (GSQ subscales; visual, auditory, olfactory, tactile, proprioception, vestibular, gustatory). As before, we included AQ scores as a covariate to control for the influence of autism traits, and this time we included age as an additional covariate to control for the influence of participant age. The ANCOVA revealed a significant main effect of age ($F(1, 298) = 18.08$, $p < .001$, $\eta_p^2 = .057$), indicating that overall sensory sensitivity reduced with increasing age. Despite this effect of age, our main pattern of results emerged as before. Replicating our main analysis, the ANCOVA revealed a significant main effect of group, ($F(1, 298) = 20.78$, $p < .001$, $\eta_p^2 = .065$), a significant interaction between group and sensitivity-type (hyper/hypo), ($F(1, 298) = 6.32$, $p = .012$, $\eta_p^2 = .021$; with Greenhouse-Geisser correction), and a significant interaction between group and sense domain, ($F(5.16, 1538.58) = 5.96$, $p < .001$, $\eta_p^2 = .020$; Greenhouse-Geisser correction). This time, there was also a significant main effect

of sensitivity type, ($F(1, 298) = 8.55$, $p = .004$, $\eta_p^2 = .028$; with Greenhouse-Geisser correction), reflecting that overall participants tended to report more hyper-sensitivities than hypo-sensitivities. As before, the ANCOVA also revealed a number of other effects which were unrelated to our hypotheses (which again we did not explore further to reduce proliferation of multiple comparisons). For example, there was a significant main effect of AQ score, ($F(1, 298) = 89.88$, $p < .001$, $\eta_p^2 = .232$), and a significant effect of sense domain, ($F(5.16, 1538.58) = 3.34$, $p = .005$, $\eta_p^2 = .011$; Greenhouse-Geisser correction).

In sum, our results show that when controlling for the influence of age, the overall pattern of results for both our analyses (imagery and sensitivity) remain unchanged.

Appendix B

Below, we replicate our analyses with study type (in-person, or online via Skype Business) included as an additional between-subjects factor. We show that study type does not affect any of our face processing outcomes (see Table S1 for means and standard deviations). For completeness, we also replicate our main analyses using statistical methods (using R Core Team, 2021; RStudio Team, 2021) that are robust to violations of the assumptions of parametric tests (e.g., normality of distribution, homoscedasticity; Mair & Wilcox, 2020), showing again that the overall pattern of results remains unchanged.

Analyses Including Study Type

Prosopagnosia Index

We conducted a 2x2 ANOVA crossing group (aphantasia, control) with study type (in-person, online). Replicating our main results, the ANOVA revealed a significant main effect of group, ($F(1, 88) = 42.18$, $p < .001$, $\eta_p^2 = .324$). There was, however, no main effect of study type, ($F(1, 88) = .31$, $p = .577$, $\eta_p^2 = .004$), showing that scores were similar in participants completing the study in-person and online (remotely). There was also no interaction, ($F(1, 88) = .052$, $p = .820$, $\eta_p^2 = .001$).

Cambridge Face Memory Test

We conducted a 2x2x3 ANOVA crossing group (aphantasia, control) with study type (in-person, control), and CFMT block (Same Images, Novel Images, Novel Images with Noise; where each block was measured by the percentage correct, see Methods). The ANOVA revealed no main effect of study type, ($F(1, 88) = 3.39$, $p = .069$, $\eta_p^2 = .037$), showing that performance on the CFMT was similar in participants completing the task in-person, and online. Replicating our main results, there was a significant main effect of group, ($F(1, 88) = 5.57$, $p = .020$, $\eta_p^2 = .060$), a significant main effect of CFMT block, ($F(1.70, 149.27) = 224.68$, $p < .001$, $\eta_p^2 = .719$; with Greenhouse-Geisser correction), and a significant interaction between block type and aphantasia, ($F(1.70, 149.27) = 3.90$, $p = .029$, $\eta_p^2 = .042$; with Greenhouse-Geisser correction). There were no further interactions (all $p > .395$).

Face Composites

We conducted a 2x2x2x2 ANOVA crossing group (aphantasia, controls) with composite type (holistic, featural), rating type (imagers, aphantasics), and study type (in-person, online). Our ANOVA revealed no main effect of study type, ($F(1, 88) = .77$, $p = .383$, $\eta_p^2 = .009$), showing that resemblance ratings were similar for composites created in-person, and online (remotely). Replicating our main results, there was also no main effect of group, ($F(1, 88) = .63$, $p = .431$, $\eta_p^2 = .007$), and no main effect of composite type, ($F(1, 88) = .21$, $p = .651$, $\eta_p^2 = .002$), but there was a significant effect of rater type, ($F(1, 88) = 161.11$, $p < .001$, $\eta_p^2 = .647$). There were no significant interactions (all $p > .092$).

Robust Analyses

Prosopagnosia Index

Using the t2way function from the WRS2 package (Mair & Wilcox, 2020), we conducted a 2x2 robust ANOVA for Prosopagnosia Index scores based on 20% trimmed means, crossing group (aphantasia, control) with study type (in-person, online). Consistent with our main results, the ANOVA revealed a significant main effect of group, ($F = 38.66$, $p = .001$), but no significant main effect of study type, ($F = .54$, $p = .467$), or interaction, ($F = .14$, $p = .713$).

Cambridge Face Memory Test

Since the WRS2 package (Mair & Wilcox, 2020) only provides a robust version for a two-way mixed ANOVA, for analyses where our factors exceeded n2 we instead replicated follow-up pairwise comparisons using robust t-tests (based on 20% trimmed means). For the CFMT, we therefore conducted robust independent samples t-tests (yuen function; WRS2 package; Mair & Wilcox, 2020) comparing the percentage accuracy in each of the CFMT blocks within our aphantasic and control groups. Replicating our main results, the aphantasic group performed significantly worse than controls in both the Same Images, (*20% Trimmed M difference* = 1.74; $t(31) = 2.81$, $p = .009$, 95% CI [.474, 3.00]), and Novel Images blocks, (*20% Trimmed M difference* = 14.27; $t(53.61) = 3.10$, $p = .003$, 95% CI [5.05, 23.49]), but there was no difference between groups in the Novel Images with Noise block (*20% Trimmed M difference* = 8.68; $t(53.85) = 1.75$, $p = .085$, 95% CI [-1.25, 18.61]). These differences survive correction for multiple comparisons (FDR; Benjamini & Hochberg, 1995, 2000).

Face Composites

For our robust analyses of face composite scores, we conduct pairwise comparisons comparing aphantasic raters to imaging raters for each composite type (holistic vs featural), and each constructor type (aphantasics vs imaging controls), using dependent samples t-tests based on 20% trimmed means (yuend function; WRS2 package; Mair & Wilcox, 2020). Consistent with our main results, aphantasic raters gave significantly lower resemblance scores for both holistic (*20% Trimmed M difference* = 0.40; $t(55) = 7.30$, $p < .001$, 95% CI [.228, .506]) and featural composites (*20% Trimmed M difference* = 0.41; $t(55) = 8.59$, $p < .001$, 95% CI [.315, .507]), relative to imaging raters. Aphantasic raters also gave significantly lower resemblance scores (collapsed across featural and holistic scores) than imaging raters for composites constructed by aphantasics (*20% Trimmed M difference* = .43; $t(31) = 8.46$, $p < .001$, 95% CI [.328, .536]), and imaging controls (*20% Trimmed M difference* = .33; $t(23) = 5.76$, $p < .001$, 95% CI [.214, .453]). Again, the differences here survive correction for multiple comparisons.

Robust Correlations

We replicated our correlational analyses using the percentage bend method (Wilcox, 1994; using the dplyr package; Wickham, François, Henry, & Müller, 2020). Replicating our main results, we found a significant positive correlation between holistic and featural composite scores for any given participant, ($r(92) = .48$, $p < .001$; where ratings were collapsed across aphantasic and imager raters). Both rating groups (aphantasics and imagers) were also again in agreement about the relative resemblance of featural ($r(92) = .89$, $p < .001$) and holistic ($r(92) = .82$, $p < .001$) composites. Replicating our main results, AQ score did not correlate with the CFMT, or featural composite scores (all $|rs| < .17$, all uncorrected $p > .25$). Correlations between the AQ and the Prosopagnosia Index ($r = .29$, uncorrected $p = .04$), and AQ and holistic composite scores ($r = -.29$, uncorrected $p = .041$) did not survive correction for multiple comparisons (FDR; Benjamini & Hochberg, 1995, 2000). As such, mirroring our main analyses, autistic traits do not seem to be robustly linked to the outcomes of our face processing measures.

Imagery screener (used to categorise rating participants as aphantasic or imagers)

Some people can form clear pictures in their mind. Others can't form any pictures in their mind at all (even though they can still remember what things looked like). And some people are inbetween. Let's find out where you are on this scale.

Close your eyes and imagine the building where you live. Try to form a picture of it in your mind, as if you were standing in front of it. Click one option below to describe how much your memory of it is like a picture.

- I have an image like a picture that is perfectly clear and vivid - just as if I were really standing there
- I have an image that is clear and reasonably vivid, like a relatively clear mental picture
- I have an image that is moderately clear and vivid
- I have an image that is only vague or fleeting or dim
- I have no image or picture at all (but I still know what it looks like)
- I have no image or picture at all (and I DON'T know what it looks like)

Table S1

Means and standard deviations for aphantasics and imaging controls for the Face Composite Task, the Cambridge Face Memory Test (CFMT), and the Prosopagnosia Index, as a function of study type (in-person, or online) and overall. The Face Composite Task scores are broken down by rater type (aphantasics, and imagers).

	In-person						Online						Overall			
Measure	Aphantasia (n27)		Control (n15)		Overall (n42)		Aphantasia (n25)		Control (n25)		Overall (n50)		Aphantasia (n52)		Control (n40)	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	Mean	SD	Mean	SD
Prosopagnosia Index	59.85	15.20	39.60	11.25	52.62	16.91	57.48	17.81	38.60	9.14	48.04	16.95	58.71	16.38	38.98	9.85
CFMT (%)																
Overall	66.31	14.31	75.28	12.58	69.51	14.24	73.22	16.94	79.06	12.60	76.14	15.07	69.63	15.86	77.64	12.57
Same Images	94.24	7.31	97.41	4.13	95.37	6.48	97.56	5.34	99.56	1.54	98.56	4.02	95.83	6.59	98.75	2.95
Novel Images	59.01	19.14	75.11	14.63	64.76	19.14	68.00	22.63	76.40	19.41	72.20	21.29	63.33	21.17	75.92	17.58
Noise Images	54.48	19.47	58.89	19.72	56.05	19.44	61.50	22.29	67.00	16.31	64.25	19.53	57.85	20.97	63.96	17.87
Composites																
Imager ratings																
Holistic	2.21	0.73	1.96	0.75	2.12	0.74	2.09	0.66	2.36	0.73	2.23	0.71	2.15	0.69	2.21	0.75
Featural	2.00	0.82	2.03	0.66	2.01	0.75	2.06	0.96	2.24	0.71	2.15	0.84	2.03	0.88	2.16	0.69
Aphant ratings																
Holistic	1.77	0.74	1.55	0.66	1.69	0.71	1.55	0.62	2.05	0.66	1.80	0.68	1.66	0.69	1.87	0.69
Featural	1.62	0.80	1.78	0.65	1.68	0.75	1.68	0.97	1.84	0.75	1.76	0.86	1.65	0.88	1.82	0.71

References

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Benjamini, Y., & Hochberg, Y. (2000). On the adaptive control of the false discovery rate in multiple testing with independent statistics. *Journal of Educational and Behavioral Statistics*, *25*(1), 60–83. https://doi.org/10.3102/10769986025001060

Mair, P., & Wilcox, R. (2020). Robust statistical methods in R using the WRS2 package. *Behavior Research Methods*, *52*(2), 464–488. https://doi.org/10.3758/s13428-019-01246-w

R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

RStudio. (2021). *RStudio: Integrated Development for R*. Boston, MA: RStudio, PBC.

Wickham, H., François, R., Henry, L., & Müller, K. (2020). *dplyr: A Grammar of Data Manipulation.* Retrieved from http://cran.r-project.org/package=dplyr

Wilcox, R. R. (1994). The percentage bend correlation coefficient. *Psychometrika*, *59*(4), 601–616.