



Special issue: Review

Multimodal mental imagery

Bence Nanay ^{a,b,*}^a University of Antwerp, Antwerp, Belgium^b Peterhouse, University of Cambridge, Cambridge, UK

ARTICLE INFO

Article history:

Received 1 February 2017

Reviewed 3 July 2017

Revised 5 July 2017

Accepted 7 July 2017

Published online 17 July 2017

Keywords:

Mental imagery

Multimodality

Multisensory perception

Synaesthesia

Sensory substitution

Implicit bias

ABSTRACT

When I am looking at my coffee machine that makes funny noises, this is an instance of multisensory perception – I perceive this event by means of both vision and audition. But very often we only receive sensory stimulation from a multisensory event by means of one sense modality, for example, when I hear the noisy coffee machine in the next room, that is, without seeing it. The aim of this paper is to bring together empirical findings about multimodal perception and empirical findings about (visual, auditory, tactile) mental imagery and argue that on occasions like this, we have *multimodal mental imagery*: perceptual processing in one sense modality (here: vision) that is triggered by sensory stimulation in another sense modality (here: audition). Multimodal mental imagery is not a rare and obscure phenomenon. The vast majority of what we perceive are multisensory events: events that can be perceived in more than one sense modality – like the noisy coffee machine. And most of the time we are only acquainted with these multisensory events via a subset of the sense modalities involved – all the other aspects of these multisensory events are represented by means of multisensory mental imagery. This means that multisensory mental imagery is a crucial element of almost all instances of everyday perception.

© 2017 The Author. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

When I am looking at my coffee machine that makes funny noises, this is an instance of multisensory perception – I perceive this event by means of both vision and audition. But very often we only receive sensory stimulation from a multisensory event by means of one sense modality. If I hear the noisy coffee machine in the next room, that is, without seeing it, then the question arises: how do I represent the visual aspects of this multisensory event? Do I represent them at all?

The aim of this paper is to bring together empirical findings about multimodal perception and empirical findings about (visual, auditory, tactile) mental imagery and argue that on occasions like the one described in the last paragraph, we have *multimodal mental imagery*: perceptual processing in one sense modality (here: vision) that is triggered by sensory stimulation in another sense modality (here: audition).

Multimodal mental imagery is not a rare and obscure phenomenon. The vast majority of what we perceive are multisensory events: events that can be perceived in more than one sense modality – like the noisy coffee machine. In

* University of Antwerp, D 413, Grote Kauwenberg 18, 2000 Antwerp, Belgium.,

E-mail address: bn206@cam.ac.uk.

URL: <http://uahost.uantwerpen.be/bence.nanay>

<http://dx.doi.org/10.1016/j.cortex.2017.07.006>

0010-9452/© 2017 The Author. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

fact, there are very few events that are not multisensory in this sense. And most of the time we are only acquainted with these multisensory events via a subset of the sense modalities involved – all the other aspects of these multisensory events are represented by means of multisensory mental imagery. This means that multisensory mental imagery is a crucial element of almost all instances of everyday perception: and a surprisingly neglected element.

In this paper, I will talk about three questions regarding multimodal mental imagery:

What is multimodal mental imagery? There is no firm theoretical framework at present for understanding multimodal mental imagery. The aim of this part of the paper is to provide one that is consistent with the methodology of experimental paradigms of two independent empirical fields in psychology and neuroscience: the study of multimodal perception and the study of mental imagery. What we need in order to fully understand multimodal mental imagery is a unifying framework that combines philosophical, psychological and neuroscientific perspectives.

What role does multimodal mental imagery play in everyday perception? Multimodal mental imagery is not an obscure and rare mental phenomenon. The aim of the second part of the paper is to argue that in the vast majority of cases, everyday perception depends constitutively on multimodal mental imagery. And this conclusion has wider implications to philosophy in general, for example, to epistemological questions about whether we can trust our senses.

What are the consequences of this general picture for experimental paradigms and clinical practice? Finally, focussing on multimodal mental imagery can help us to understand a number of puzzling perceptual phenomena, like sensory substitution and synaesthesia. Further, manipulating mental imagery has recently become an important clinical procedure in various branches of psychiatry as well as in counteracting implicit bias – using multimodal mental imagery rather than voluntarily and consciously conjured up mental imagery can lead to real progress in these clinical paradigms. This is the topic I address in the third part of the paper.

2. Unifying philosophical, psychological and neuroscientific perspectives on multimodal mental imagery

The first aim of the paper is to give a unified and solid theoretical framework for thinking about multimodal mental imagery, one that is consistent not only with recent empirical findings about multimodality and about mental imagery, but that also respects the experimental methodology of these disciplines. And here in the unification of philosophical, psychological and neuroscientific perspectives on multimodal mental imagery, philosophy needs to take the back seat. We should not start with some pre-existing philosophical or common sense conception of what multimodal mental imagery, or mental imagery in general, is supposed to be and cherry-pick the empirical results that match it. Instead, my aim is not only to use the concepts and implicit theoretical presuppositions of researchers working on multimodality and mental imagery, but also to respect their experimental

methodology. As a consequence, the resulting theoretical framework will be closer to standard analyses of mental imagery in psychology and neuroscience than it is to the standard philosophical concept.

2.1. Multimodal perception

Philosophers and cognitive scientists have assumed until relatively recently that we can study the senses – vision, audition, olfaction, etc. – independently from one another. The assumption was that we could study various aspects of, say, vision, without paying much attention to the other sense modalities. But there is overwhelming recent evidence that multimodal perception is the norm and not the exception – our sense modalities interact in a variety of ways (Bertelson & Gelder, 2004; O'Callaghan, 2014; Spence & Driver, 2004; Vroomen, Bertelson, & de Gelder, 2001). Information in one sense modality can influence the information processing in another sense modality at a very early stage of perceptual processing (often in the primary visual cortex in the case of vision, see Watkins, Shams, Tanaka, Haynes, & Rees, 2006).

A simple example is ventriloquism, which is commonly described as an illusory auditory experience caused by something visible (Bertelson, 1999; O'Callaghan, 2008b). It is one of the paradigmatic cases of crossmodal illusion: We experience the voices as coming from the dummy, while they in fact come from the ventriloquist. The auditory sense modality identifies the ventriloquist as the source of the voices, while the visual sense modality identifies the dummy. And, as it often (not always – see O'Callaghan, 2008a) happens in crossmodal illusions, the visual sense modality wins out: our (auditory) experience is of the voices as coming from the dummy. But there are more surprising examples: if there is a flash in your visual scene and you hear two beeps while the flash lasts, you experience it as two flashes (Shams, Kamitani, & Shimojo, 2000). These findings fly in the face of some of the most basic – and oldest – methodological assumptions of philosophical and psychological studies of perception. Philosophers, psychologists and cognitive scientists have based their analysis of perception on the methodological assumption that the senses can be studied independently from one another. But the new empirical findings show that this is a mistaken assumption.

Most of the multimodality research focuses on the multimodality of *perception*: on how perceptual processing in one sense modality is influenced, embellished or modified by another sense modality: how visual perceptual processing, for example, is influenced by audition. My aim is to shift this emphasis and focus on multimodal mental imagery (rather than multimodal perception): what happens when visual perceptual processing is not just modified by audition, but it is triggered by audition (for example, because there is no visual input, see Lacey & Lawson, 2013)? And in order to address this appropriately, we need to bring the multimodality literature together with another experimental research program: the one on mental imagery.

2.2. Mental imagery

Philosophical, psychological and neuroscientific approaches to mental imagery (by which I mean visual, auditory,

olfactory, tactile, etc. imagery, see [Bensafi et al., 2003](#); [Herholz, Halpern, & Zatorre, 2012](#); [Zatorre & Halpern, 2005](#)) often pull in different directions. Philosophers try to capture the intuitive concept of conjuring up an image, for example, by closing one's eyes and visualizing an apple ([Currie, 1995](#); [Kind, 2001](#); [Richardson, 1969](#)). But recent findings in neuroscience and psychology show that this is only one and not a particularly representative way of exercising mental imagery.

Recent advances in neuroimaging methodology make it possible to have a clear idea about early cortical processing in mental imagery (e.g., the primary visual cortex, see [Page, Duhamel, & Crognale, 2011](#), [Slotnick, Thompson, & Kosslyn, 2005](#)). And the retinotopy ([Grill-Spector & Malach, 2004](#)) of the early visual cortices (and their equivalent in the other sense modalities, see, e.g., [Talavage et al., 2004](#)) also makes it possible to track the content of mental imagery without having to resort to the subjects' introspective reports [a fact that highlights that mental imagery does not have to be conscious (see [Church, 2008](#); [Nanay, 2010a, 2015](#); [Phillips, 2014](#) for philosophical arguments and [Zeman et al., 2007, 2010](#); [Zeman, Dewar, Della Sala, 2015](#) for experimental evidence)]. Mental imagery, according to this paradigm, is *perceptual processing that is not triggered by corresponding sensory stimulation in a given sense modality* (see [Kosslyn, Behrmann, & Jeannerod, 1995](#); [Kosslyn, Thompson, Kim, & Alpert, 1995](#); [Nanay, 2015, 2016a, 2016b, forthcoming](#); [Pearson & Westbrook, 2015](#); [Pearson, Naselaris, Holmes, & Kosslyn, 2015](#)). Here is a representative quote from a recent review article: “We use the term ‘mental imagery’ to refer to representations [...] of sensory information without a direct external stimulus” ([Pearson et al., 2015](#)). This way of thinking about mental imagery needs some unpacking.

The last phrase ‘in a given sense modality’, is crucial in the present context: olfactory mental imagery is *olfactory* perceptual processing that is not triggered by corresponding *olfactory* sensory stimulation. Olfactory mental imagery can be (and is often) triggered by *non-olfactory* (for example, auditory) sensory stimulation. And this is exactly what I mean by multi-modal mental imagery.

By ‘sensory stimulation’ I mean the activation of the sense organ by external stimulus. In the visual sense modality, sensory stimulation amounts to the light hitting the retina. Some perceptual processing starts with sensory stimulation. But not all. Some perceptual processing – mental imagery – is not triggered by sensory stimulation (in the same sense modality).

By ‘perceptual processing’, I mean processing in the perceptual system. Some parts of the processing of the sensory stimulation are more clearly perceptual than others. To take the visual sense modality as an example ([Bullier, 2004](#); [Grill-Spector & Malach, 2004](#); [Katzner & Weigelt, 2013](#); [Van Essen, 2004](#)), in humans and nonhuman primates, the main visual pathway connects neural networks in the retina to the primary visual cortex (V1) via the lateral geniculate nucleus (LGN) in the thalamus; outputs from V1 activate other parts of the visual cortex and are also fed forward to a range of extrastriate areas (V2, V3, V4/V8, V3a, V5/MT). The earlier stages of this line of processing are more clearly perceptual than the later ones. And we can safely assume that cortical processing is perceptual processing. If we have such early

cortical processing but no corresponding sensory stimulation, we have (visual) mental imagery (see [Page et al., 2011](#); [Slotnick et al., 2005](#), but see also [Bridge, Harrold, Holmes, Stokes, & Kennard, 2012](#) for caution about how to think of ‘early cortical’ in this context).

The concept of ‘corresponding’ plays a crucial role in this way of thinking about mental imagery. We can have mental imagery even when there is sensory stimulation in the given sense modality, if it fails to correspond to the perceptual processing (we can have mental imagery of X while staring at Y). In terms of experimental methodology, correspondence is relatively easy to measure, given the retinotopy of the early visual cortices (and their equivalent in the other sense modalities, see, e.g., [Talavage et al., 2004](#)), which provides a convenient way of gaining evidence about the correspondence or lack thereof between sensory stimulation and perceptual processing. The primary visual cortex (and also many other parts of the visual cortex see [Grill-Spector & Malach, 2004](#) for a summary) is organized in a way that is very similar to the retina – it is retinotopic. So we can assess in a simple and straightforward manner whether the retinotopic perceptual processing in the primary visual cortex corresponds to the activations of the retinal cells. In the case of mental imagery, we get no such correspondence.

Mental imagery does not have anything to do with the kind of tiny images in our mind that behaviourists made fun of ([Ryle, 1949](#)). Mental imagery is not something we see: it is a certain kind of perceptual processing. So it is in no ways more mysterious than other kinds of perceptual processing (like sensory stimulation-driven perception). Nor do we need to postulate any ontologically extravagant entities (like tiny pictures in our head) to talk about mental imagery any more than we need to postulate these entities in order to talk about perception.

Defining mental imagery as perceptual processing not triggered by corresponding sensory stimulation in a given sense modality makes the example of closing one's eyes and visualizing an apple a special case of mental imagery, but it also highlights the ways in which this example is unrepresentative.

First, philosophers often take mental imagery to be necessarily conscious ([Currie, 1995](#); [Kind, 2001](#); [Richardson, 1969](#)). And visualizing an apple does indeed conjure up conscious mental imagery. But mental imagery, the way psychologists and neuroscientists use the term, is not necessarily conscious – and the experimental methodology of neither psychology nor neuroscience treats it as necessarily conscious (starting with the classic mental rotation experiment of [Shepard & Metzler, 1971](#)). Perception can be conscious or unconscious (see [Kentridge, Heywood, & Weiskrantz, 1999](#) for a summary). So it would be surprising if mental imagery had to be conscious (see also [Church, 2008](#); [Nanay, 2010a](#); [Phillips, 2014](#) for some philosophical arguments). But we also have strong empirical reasons for supposing that mental imagery can be unconscious. There are subjects (and in fact, surprisingly many of them) who have no conscious experience of mental imagery whatsoever, and at least some of these subjects are still capable of performing tasks that are assumed to require the manipulation of mental imagery ([Zeman et al., 2007, 2010, 2015](#)). Further, there is a straight

correlation between the vividness and salience of mental imagery and some straightforward (and very easily measurable) physiological features of the subject's brain (such as the size of the subject's primary visual cortex) and the relation between early cortical activities and the activities in the entire brain (see Bergmann, Genc, Kohler, Singer, & Pearson, 2016 and Cui, Jeter, Yang, Montague, & Eagleman, 2007).

I should emphasize that mental imagery can be and is very often conscious – for example, when you close your eyes and visualize an apple. And such conscious occurrences of mental imagery can be very valuable for finding out more about mental imagery in general in experimental settings (see, e.g., Pearson, Clifford, & Tong, 2008). My aim is not to replace conscious mental imagery with unconscious one, but rather to expand the category of mental imagery so that it would include both conscious and unconscious mental imagery. Further, differentiating between unconscious mental imagery and no mental imagery can be tricky – obviously we are not in a position to do so introspectively. The difference is a functional one: if I have early cortical visual processing but no visual input, I do have mental imagery, even if I have no idea that I do. And we can only figure this out for sure if we put the subject in a scanner.

Second, visualizing the apple is something you do voluntarily and intentionally. But mental imagery does not have to be voluntary or intentional. One can have flashbacks of some unpleasant scene – this is also mental imagery, but it is not a voluntary or intentional exercise of mental imagery. And some of our mental imagery is of this involuntary and unintentional kind – this is especially clear in the auditory sense modality, as demonstrated by the phenomenon of earworms: tunes that pop into our heads and that we keep on having auditory imagery of, even though we do not want to.

Third, visualizing an apple is not normally accompanied by any feeling of presence. You are not fooled by this mental imagery into thinking that there is actually an apple in front of you so that you could reach out and grab it. But, again, this is not a necessary feature of mental imagery. There is no *prima facie* reason why mental imagery could not be accompanied by the feeling of presence. In fact, dreaming, and especially lucid dreaming, which is widely considered to be a form of mental imagery (see Hobbes, 1654; Walton, 1990), is very much accompanied by the feeling of presence.

The psychological/neuroscientific concept of mental imagery (that I parse as perceptual processing not triggered by corresponding sensory stimulation) is an extension of the introspective/philosophical concept of mental imagery that focuses on examples like closing our eyes and visualizing an apple. But just how (and how far) we can extend the introspective concept of mental imagery (and where this extension should stop) is something introspection will not tell us – we need perceptual psychology and cognitive neuroscience for that.

We are now in the position to put together the multimodality part and the mental imagery part: Perceptual processing can be triggered by various things. If it is triggered by corresponding sensory stimulation in the given sense modality, we get sensory stimulation-driven perception. If it is not triggered by corresponding sensory stimulation in the given sense modality, we get mental imagery. And if it is not

triggered by corresponding sensory stimulation in the given sense modality, but it is triggered by corresponding sensory stimulation in *another* sense modality, we get multimodal mental imagery. We have very diverse evidence from neuroscience that even very early sensory cortical processing can be triggered by sensory stimulation of another sense modality (Calvert et al., 1997; Sekuler, Sekuler, & Lau, 1997; Ghazanfar & Schroeder, 2006; Hertrich, Dietrich, & Ackermann, 2011; Iurilli et al., 2012; James et al., 2002; Martuzzi et al., 2007; Pekkola et al., 2005; Zangaladze, Weisser, Stilla, Prather, & Sathian, 1999). This is the concept of multimodal mental imagery this paper is about.

3. The role of multimodal mental imagery in everyday perception

Most of the events we encounter are multisensory events: event we can perceive by means of more than one sense modality. And as our perceptual access to these multisensory events is rarely absolute (that is, encompassing all relevant sense modalities), our perceptual system needs to rely on multimodal mental imagery to represent those features of these multisensory events that we are not getting sensory stimulation from. But given that this happens in almost all real life perceptual situations, it follows that multimodal mental imagery is an integral part of perception *per se* – a claim that has echoes of Kant's 'imagination is a necessary constituent of perception itself' (see also Strawson, 1974), but that is much more specific and testable claim and one that is grounded in empirical research.

Multimodal mental imagery is very different from the kind of mental imagery that we have when we close our eyes and visualize an apple. Visualising the apple is conscious, voluntary and it is not accompanied by the feeling of presence. Multimodal mental imagery is normally involuntary and normally unconscious. But when it is conscious, it is accompanied by the feeling of presence. Nonetheless, both count as mental imagery – both involve early perceptual processing not triggered by corresponding sensory stimulation in the relevant sense modality.

Most of the time, when we form mental imagery of those parts of a multisensory event that we are not acquainted with, this mental imagery will be unattended and unconscious. But if we are really interested in them, we can attend to them and this can make them more salient. And while most of the time the properties we attribute to those aspects of the multisensory event that we are not acquainted with are very determinable, we can make them more determinate (if we are really interested in them for some reason, see Nanay, 2010b on how attention can change the determinacy of perceptual content).

Suppose that I am working in my room and I hear footsteps from downstairs (without seeing who is coming upstairs). I represent the complex multisensory event of someone coming upstairs: I perceive the auditory parts of this event and I represent the other (visual, maybe olfactory) parts of this event by means of mental imagery. But my visual and olfactory multimodal mental imagery may not be conscious – if I am not too concerned with who is coming upstairs. My

olfactory mental imagery of the olfactory aspects of the multisensory event whose auditory aspects I am acquainted with is likely to be unattended, unconscious and very determinable. But if the only two people who can come upstairs are my stinky friend X or my other friend, Y, who uses very nice perfume, and if I really want to know which one it is, I will be likely to fill in the olfactory aspects of the multisensory event in a more determinate way (which can prime me to recognize them by smell more quickly).

Here is a nice experimental illustration of this point. The double flash illusion is one of the most striking crossmodal illusions: you are presented with one flash and two beeps simultaneously (Shams et al., 2000). So the sensory stimulation in the visual sense modality is one flash. But you experience two flashes and already in the primary visual cortex, two flashes are processed (Watkins et al., 2006). This means that the double flash illusion is really about multimodal mental imagery: we have perceptual processing in the visual sense modality (again, already in V1) that is not triggered by corresponding sensory stimulation in the visual sense modality (but by corresponding sensory stimulation in the auditory sense modality). The multimodal mental imagery that is involved in the double flash illusion is conscious, involuntary and it is accompanied by the feeling of presence. In fact, it is accompanied by the feeling of presence so much that we do take ourselves to perceive two flashes not one.

The main claim of the second half of this paper is that (almost) all perception involves multimodal mental imagery (of all the unacquainted parts of the multisensory event). This multimodal mental imagery, however, is largely unconscious.

It is important to emphasize that the reference to multimodality in the label ‘multimodal mental imagery’ does not refer to the multimodality of our phenomenology when we have multimodal mental imagery. In the case of unconscious multimodal mental imagery, there is no phenomenology whatsoever, multimodal or not. What ‘multimodal’ refers to in the name of multimodal mental imagery is the aetiology of mental imagery: mental imagery is the product of the interaction between (at least) two different sense modalities. The phenomenal feel of multimodal mental imagery, if there is one, may itself be unimodal, say, purely visual. But it is the outcome of the interaction between vision and another sense modality – it is multimodal in this sense.

Multimodal mental imagery may or may not be influenced in a top-down manner. Many of the experimental findings about multimodal mental imagery are about lateral or bottom-up effects, where, for example, the primary visual cortex is activated by the processing of sensory stimulation in the auditory cortex (and not by anything top-down) – this seems to be the case in the double flash illusion (but see Roseboom, Kawabe, & Nishida, 2013 for some further wrinkles). But in many other cases of multimodal mental imagery, top-down influences are very important.

Here is a widely researched example: seeing someone talking on television with the sound muted. The visual perception of the talking head in the visual sense modality leads to an auditory mental imagery in the auditory sense modality (e.g., Calvert et al., 1997; Hertrich et al., 2011; Pekkola et al., 2005, see also Spence & Deroy, 2013 for a summary). This auditory mental imagery will very much depend on bottom-

up factors like the lip movements of the person on the screen. But not only these.

If this person is someone you know or have heard speak, your auditory mental imagery will be influenced by this information. If it is Barack Obama (someone you have, presumably, heard before), many subjects in fact ‘hear’ him speaking with his distinctive tone of voice or intonation, for example. When I am listening to Obama’s speech with the TV muted, my auditory mental imagery is influenced by various past memories of hearing Obama speak and my expectation of how his voice would sound. I want to leave open just how far ‘top’ this information is being coded – but it is clearly further up than the visual (or auditory) cortex (see Teufel & Nanay, 2017 for a more detailed treatment of neuroscientific findings on what ‘top’ means in ‘top-down influences on perception’).

The picture we ended up with is one where perceptual processes consist of a sensory-stimulation-driven and a non-sensory-stimulation-driven component (where by sensory-stimulation-driven, I mean driven by corresponding sensory stimulation of the relevant sense modality). In other words, perception consists in mental imagery and stimulation-driven perception. And mental imagery influences the way the sensory stimulation gets processed. In some very rare examples of simple two-dimensional visual displays, the mental imagery component may be missing. But in the vast majority of perceptual scenarios, it is present and it gets combined with the sensory stimulus-driven perceptual processing. And in these cases, much of what we take ourselves to perceive we at least partly represent by means of mental imagery. And as at least some of these episodes of mental imagery are subject to top-down influences, perception per se can also be subject to top-down influences. This argument for top-down influences on perceptual processing is different from the standard philosophical ones (Macpherson, 2012; Siegel, 2011; Stokes, 2012) and it is not susceptible to some of the criticism they face (see, e.g., Firestone & Scholl, 2014, 2016).

This sensitivity to top-down influences could be thought to have far-reaching consequences for epistemology. If perception is sensitive to top-down influences because of the top-down sensitivity of multimodal mental imagery, then it is not an unbiased way of learning about the world, as our pre-existing thoughts, beliefs and expectations could influence how and what we perceive (see Siegel, 2011 on a version of this worry). So we get a form of vicious circularity: our beliefs, thoughts and expectations are supposed to be based on and justified by our perceptual states, but these perceptual states themselves are influenced by our beliefs, thoughts and expectations (because of the top-down influences on perception via multimodal mental imagery).

But I want to argue that the role multimodal mental imagery plays in everyday perception poses an even more significant worry for epistemology, regardless of whether multimodal mental imagery is influenced in a top-down manner. Knowledge is supposed to be a good thing because it tracks the truth. And perception is supposed to be a good way of acquiring knowledge because perception tracks the truth. But multimodal mental imagery does not automatically track the truth. It is, by definition, a step removed from sensory stimulation and, as a result, from the world. As a

result, it is more easily fooled – as the double flash illusion shows. And this is true of mental imagery in general, not only of mental imagery that is influenced in a top-down manner. Even those who want to deny any top-down influences on perception would need to take this challenge seriously.

The question then is whether we are justified to move from (multimodal mental imagery-tinted) perception to belief. Any such move is very far away from a ‘prima facie’ or ‘immediate’ justification some philosophers expect of perceptual justification (Huemer, 2001; Pryor, 2000). It would need to involve a close empirical examination of the reliability of the processes that give rise to multimodal mental imagery. The conclusion is that the question of perceptual justification is at least in part an empirical question – it requires the examination of the reliability of multimodal mental imagery that plays a constitutive part of perception per se. This is a sense (a fairly narrow sense, to be sure) in which epistemology needs to be naturalized (see Nanay, 2017c for more on this).

4. Experimental and clinical applications of the concept of multimodal mental imagery

So far, I gave an empirically informed theoretical account of multimodal mental imagery. In this last section, I want to discuss some consequences of this general theoretical framework for some empirical paradigms and also for some clinical practices.

One straightforward application concerns perception aided by sensory substitution. Blind subjects can be taught to navigate their environment in some sense ‘visually’ by having a camera installed on their body the images of which are fed into some other sense modality of the subject. The camera is recording images continuously and these images are transmitted to the subject in real time in the tactile sense modality, for example (it can also be done auditorily). So the images are imprinted on the subject’s skin with slight pricks as soon as they are recorded (see Bach-y-Rita & Kercel, 2003 and Sampaio, Maris, & Bach-y-Rita, 2001 for summaries).

The surprising result is that the subjects eventually start experiencing the scene in front of them ‘visually’ – they talk about visual occlusion, for example and they become very competent at navigating relatively complex terrains. What is even more important from my point of view is that there was activity in the primary visual cortex of these subjects (Murphy et al., 2016) that was clearly not triggered by sensory stimulation in the relevant sense modality – as the subjects were blind. But they were triggered by corresponding sensory stimulation in the tactile sense modality. So perception by sensory substitution would count as multimodal mental imagery.

These findings are often discussed as ammunition in the grand debate about how we should individuate the senses: about what makes, say, vision different from audition (see, e.g., Hurley & Noë, 2003). Is ‘vision’ assisted by sensory substitution really vision? Or is it tactile perception? Some of the classic ways of individuating the senses come apart in this odd case: if we individuate the senses according to the sensory stimulation, then sensory substitution assisted ‘vision’ would

count as tactile perception. If we individuate the senses according to phenomenology, then it seems to be vision.

But according to the concept of multimodal mental imagery I analysed above, sensory substitution assisted ‘vision’ is neither vision nor tactile perception. It is not perception at all. It is mental imagery – multimodal mental imagery. It is visual mental imagery triggered by tactile sensory stimulation. Sensory substitution involves perceptual processing and very clearly visual perceptual processing – as the activity in the primary visual cortex clearly shows (and this coincides with the phenomenology of the subjects). And this visual perceptual processing is induced by tactile sensory stimulation – slight pricks on the subject’s skin. As clear-cut a case of multimodal mental imagery as it gets (see Nanay 2017a for a more detailed analysis).

A more complex application of the multimodal mental imagery framework concerns synaesthesia. Some synaesthesia subjects see a colour and they experience it as having a certain pitch. Or the other way round: they hear a note and experience it as having a specific colour (Ward, Huckstep, & Tsakanikos, 2006). Synaesthesia comes in various different forms: lexical-gustatory synaesthesia (Jones et al., 2011; Ward & Simner, 2003), coloured touch synaesthesia (Ludwig & Simner, 2013), spatial time units synaesthesia (Jarick, Jensen, Dixon, & Smilek, 2011; Smilek, Callejas, Dixon, & Merikle, 2007). The list could go on.

It is widely agreed that synaesthesia is intricately connected with unusual ways of exercising one’s mental imagery. Indeed, besides introspective reports, there is neuroimaging data that makes this connection (Spiller, Jonas, Simner, & Jansari, 2015), there is also a large body of data on the involvement of the early cortical areas of the relevant sense modality in synaesthesia (see, e.g., Hubbard, Cyrus Arman, Ramachandran, & Boynton, 2005, Jones et al., 2011, Nunn et al., 2002). The question is how we can explain the ways in which subjects with synaesthesia exercise their mental imagery differently from subjects without synaesthesia.

Some instances of synaesthesia are multimodal – for example the pitch and colour synaesthesia I mentioned first. I will focus on these in the paper, bracketing (for the purpose of this paper) unimodal instances of synaesthesia (for example the quite widespread grapheme-colour synaesthesia (of having coloured mental imagery of numerals or letters)). Seeing a colour and having auditory mental imagery of a certain pitch is a clear case of multimodal mental imagery: the perceptual processing in the auditory sense modality is triggered by sensory stimulation in the visual sense modality. And hearing a pitch and having the visual mental imagery of a colour is also a clear case of multimodal mental imagery (where the perceptual processing in the visual sense modality is triggered by the sensory stimulation in the auditory sense modality).

And the only way these forms of multimodal mental imagery differ from less odd cases of multimodal mental imagery is that the ‘correspondence’ in the definition of multimodal mental imagery is different. The ‘corresponding’ sensory stimulation in the case of synaesthesia is very idiosyncratic. When we have auditory mental imagery of Barack Obama’s voice when watching the muted recording of Barack Obama speak, there is a correspondence, because we have encountered Barack Obama (well, recordings of Barack Obama

for most of us) via both the visual and the auditory sense modalities a number of times. Correspondence here is based on repeated exposure of the simultaneous presentation of the relevant auditory and visual stimuli. In the case of synaesthesia, in contrast, the ‘correspondence’ is not based on repeated exposure of any kind of multisensory event (which would be the equivalent of Obama talking). We do not encounter, let alone repeatedly encounter colours that have a certain specific pitch (and the same pitch in all contexts). So synaesthesia is multimodal mental imagery where ‘correspondence’ is unusual. And we can make progress in understanding synaesthesia if we focus on the nature and origins of the idiosyncratic correspondence relation that constitutes the difference between synaesthesia and more standard cases of multimodal mental imagery.

The self-report of many subjects indicate that they take themselves to literally hear the pitch of colours. Nonetheless, given that the auditory perceptual processing of the pitch is not triggered by corresponding auditory sensory stimulation (but rather by visual sensory stimulation), it counts as mental imagery, not perception. The fact that synaesthesia subjects can mistake one for the other indicates that the mental imagery involved in synaesthesia comes with the feeling of presence. This makes it even more similar to other ways of exercising multimodal mental imagery.¹

Finally, one of the most exciting applications of the concept of multisensory mental imagery, and the most relevant one for society at large, concerns the clinical practice of some branches of psychiatry and of reducing implicit bias.

A relatively new development in some branches of psychiatry is to manipulate the mental imagery of patients with a wide range of mental disorders by means of techniques such as ‘imaginal exposure’, ‘systematic desensitization’ and ‘imagery rescripting’, in order to improve their condition. There are reports of the success of this methodology in the case of mental disorders ranging from bipolar disorders, schizophrenia and post-traumatic stress disorder to obsessive compulsive disorder and depression (Holmes, James, Kilford, & Deepro, 2010; James et al., 2015, see Pearson et al., 2015 for a summary). While the results are very promising, there is great variability of the efficiency of this method between subjects (see, for example, Blackwell et al., 2013; Williams, Blackwell, Mackenzie, Holmes, & Andrews, 2013).

It is important that this clinical method typically relies fully on the manipulation of the conscious and voluntary mental imagery of the subjects: they ask the patients to visualize a certain event consciously and voluntarily, for

example by asking them to imagine that they are on a sunny beach.² The problem with this methodology is that voluntarily triggered mental imagery is very difficult to control and to maintain. But as we have seen, voluntary mental imagery is just one form of mental imagery and one that is dependent on a lot of factors that might prevent the patient from succeeding in visualizing what she is asked to visualize (see Clark, Holmes, Woolrich, & Mackay, 2016; Murphy, O’Donoghue, Drazich, Blackwell, & Nobre, 2015; Slofstra, Nauta, Holmes, & Bockting, 2016). Inducing multimodal mental imagery, on the other hand, could bypass these blocks and it could provide a more efficient way of interfering with the patients’ mental imagery, which is easier to control and to maintain.

And a similar use of multimodal mental imagery could be suggested for fighting implicit bias. Research on implicit bias shows that we have unconscious attitudes towards certain racial and gender groups (that can be very different from our conscious convictions) (Dunham, Baron, & Banaji, 2008; Greenwald & Banaji, 1995). This is demonstrated with the help of subjects’ reaction time in the Implicit Association Test, which measures how closely one associates certain images and words with racial or gender terms.

It has recently been explored how visualizing and imaginative engagement (for example, by visualizing ‘ingroup’ and ‘outgroup’ faces or by putting ourselves imaginatively into an avatar of another racial or gender group) can reduce implicit bias (Peck, Seinfeld, Aglioti, & Slater, 2013; Ratner, Dotsch, Wigboldus, van Knippenberg, & Amodio, 2014). Again, these methods, while promising, seem to be limited in their success partly because conscious and voluntary visualizing is difficult to maintain and control for. But, as in the case of the clinical applications in psychiatry, involuntarily and unconsciously generated multimodal mental imagery can be utilized in a less fragile manner.³ This should lead us to explore the methodological constraints and possibilities of doing so in future research.

Acknowledgements

This work was supported by the ERC Consolidator grant 726251, the FWO Odysseus grant G.0020.12N and the FWO research grant G0C7416N.

REFERENCES

-
- Bach-y-Rita, P., & Kercel, S. W. (2003). Sensory substitution and the human–machine interface. *Trends in Cognitive Sciences*, 7, 541–546.
- Bensafi, M., Porter, J., Pouliot, S., Mainland, J., Johnson, B., Zelano, C., et al. (2003). Olfactomotor activity during imagery

¹ Another important application of the concept of multimodal mental imagery, one I cannot explore here further, is that it can help us understand and explain the differences between the ways in which ‘concrete’ and ‘abstract’ words are processed (Paivio, 1969, 1971, 1986). The general idea would be that the emphasis on the multimodal nature of the mental imagery involved in the processing of these words can expand the scope of Paivio’s dual processing theory. I am grateful to an anonymous referee for drawing my attention to this potential application.

² Less often, subjects are asked to perform overt imagery-involving actions (like playing certain computer games, see Holmes et al., 2010) – this methodology already utilizes involuntary mental imagery, just like multimodal mental imagery.

³ Similar potential applications where multimodal mental imagery would provide a more efficient way of interfering than traditional conscious and voluntary mental imagery include treating addiction (see Papies and Barsalou, 2015), eating disorders (Cornil & Chandon, in press) and pain treatment (Berna, Tracey, & Holmes, 2012; Winterowd, Beck, & Gruener, 2003, pp. 183–207, see also Nanay, 2017b for a summary).

- mimics that during perception. *Nature Neuroscience*, 6, 1142–1144.
- Bergmann, J., Genc, E., Kohler, A., Singer, W., & Pearson, J. (2016). Smaller primary visual cortex is associated with stronger, but less precise mental imagery. *Cerebral Cortex*, 26, 3838–3850.
- Berna, C., Tracey, I., & Holmes, E. A. (2012). How a better understanding of spontaneous mental imagery linked to pain could enhance imagery-based therapy in chronic pain. *Journal of Experimental Psychopathology*, 3, 258–273.
- Bertelson, P. (1999). Ventriloquism: A case of cross-modal perceptual grouping. In G. Aschersleben, et al. (Eds.), *Cognitive contributions to the perception of spatial and temporal events* (pp. 347–362). Amsterdam: Elsevier.
- Bertelson, P., & de Gelder, B. (2004). The psychology of multimodal perception. In C. Spence, & J. Driver (Eds.), *Crossmodal space and crossmodal attention* (pp. 141–177). Oxford: Oxford University Press.
- Blackwell, S. E., Rius-Ottenheim, N., Schulte-van Maaren, Y. W., Carlier, I. V., Middelkoop, V. D., Zitman, F. G., et al. (2013). Optimism and mental imagery: A possible cognitive marker to promote wellbeing? *Psychiatry Research*, 206, 56–61.
- Bridge, H., Harrold, S., Holmes, E. A., Stokes, M., & Kennard, C. (2012). Vivid visual mental imagery in the absence of the primary visual cortex. *Journal of Neurology*, 259, 1062–1070.
- Bullier, J. (2004). Communications between cortical areas of the visual system. In L. M. Chalupa, & J. S. Werner (Eds.), *The visual neurosciences*. Cambridge, MA: MIT Press.
- Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C., McGuire, P. K., et al. (1997). Activation of auditory cortex during silent lipreading. *Science*, 276, 593–596.
- Church, J. (2008). The hidden image: A defense of unconscious imagining and its importance. *American Imago*, 65, 379–404.
- Clark, I. A., Holmes, E. A., Woolrich, M. W., & Mackay, C. E. (2016). Intrusive memories to traumatic footage: The neural basis of their encoding and involuntary recall. *Psychological Medicine*, 46, 505–518.
- Cornil, Y., & Chandon, P. (2017). Pleasure as a substitute for size: How multisensory imagery can make people happier with smaller food portions. *Journal of Marketing Research* (in press).
- Cui, X., Jeter, C. B., Yang, D., Montague, P. R., & Eagleman, D. M. (2007). Vividness of mental imagery: Individual variability can be measured objectively. *Vision Research*, 47, 4474–4478.
- Currie, G. (1995). Visual imagery as the simulation of vision. *Mind and Language*, 10, 25–44.
- Dunham, Y., Baron, A. S., & Banaji, M. R. (2008). The development of implicit intergroup cognition. *Trends in Cognitive Sciences*, 12, 248–253.
- Firestone, C., & Scholl, B. J. (2014). “Top-down” effects where none should be found: The El Greco fallacy in perception research. *Psychological Science*, 25(1), 38–46.
- Firestone, C., & Scholl, B. J. (2016). There are no top-down influences on perception. *Behavioral and Brain Sciences*, 39, e229.
- Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, 10, 278–285.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition. *Psychological Review*, 102, 4–27.
- Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience*, 27, 649–677.
- 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.
- Hertrich, I., Dietrich, S., & Ackermann, H. (2011). Cross-modal interactions during perception of audiovisual speech and nonspeech signals: An fMRI study. *Journal of Cognitive Neuroscience*, 23, 221–237.
- Hobbes, T. (1654). *De Corpore*. London.
- Holmes, E. A., James, E. L., Kilford, E. J., & DePrese, C. (2010). Key steps in developing a cognitive vaccine against traumatic flashbacks: Visuospatial tetris versus verbal pub quiz. *PLoS One*, 5, e13706.
- Hubbard, Edward M., Cyrus Arman, A., Ramachandran, Vilayanur S., & Boynton, Geoffrey M. (2005). Individual differences among grapheme-color synesthetes: Brain-behavior correlations. *Neuron*, 45(6), 975–985.
- Huemer, M. (2001). *Scepticism and the veil of perception*. Lanham: Rowman and Littlefield.
- Hurley, S., & Noë, A. (2003). Neuralplasticity and consciousness. *Biology & Philosophy*, 18, 131–168.
- Iurilli, G., Ghezzi, D., Olcese, U., Lassi, G., Nazzaro, C., Tonini, R., et al. (2012). Sound-driven synaptic inhibition in primary visual cortex. *Neuron*, 73, 814–828.
- James, T. W., Humphrey, G. K., Gati, J. S., Servos, P., Menon, R. S., & Goodale, M. A. (2002). Haptic study of three-dimensional objects activates extrastriate visual areas. *Neuropsychologia*, 40, 1706–1714.
- James, E. L., Bonsall, M. B., Hoppitt, L., Tunbridge, E. M., Geddes, J. R., Milton, A. L., et al. (2015). Computer game play reduces intrusive memories of experimental trauma via reconsolidation update mechanisms. *Psychological Science*, 26, 1201–1215.
- Jarick, Michelle, Jensen, Candice, Dixon, Michael J., & Smilek, Daniel (2011). The automaticity of vantage point shifts within a synaesthetes? *Spatial calendar. Journal of Neuropsychology*, 5(2), 333–352.
- Jones, Catherine L., Gray, Marcus A., Minati, Ludovico, Simner, Julia, Critchley, Hugo D., & Ward, Jamie (2011). The neural basis of illusory gustatory sensations: Two rare cases of lexicalgustatory synaesthesia. *Journal of Neuropsychology*, 5(2), 243–254.
- Katzner, S., & Weigelt, S. (2013). Visual cortical networks: Of mice and men. *Current Opinion in Neurobiology*, 23, 202–206.
- Kentridge, R. W., Heywood, C. A., & Weiskrantz, L. (1999). Attention without awareness in blindsight. *Proceedings of the Royal Society of London B*, 266, 1805–1811.
- Kind, A. (2001). Putting the image back to imagination. *Philosophy and Phenomenological Research*, 62, 85–109.
- Kosslyn, S. M., Behrmann, M., & Jeannerod, M. (1995). The cognitive neuroscience of mental imagery. *Neuropsychologia*, 33, 1335–1344.
- Kosslyn, Stephen M., Thompson, William L., Kim, Irene J., & Alpert, Nathaniel M. (1995). Topographical representations of mental images in primary visual cortex. *Nature*, 378, 496–498.
- Lacey, S., & Lawson, R. (2013). *Multisensory imagery*. New York: Springer.
- Ludwig, V. U., & Simner, J. (2013). What colour does that feel? Tactile–visual mapping and the development of cross-modality. *Cortex*, 49, 1089–1099.
- Macpherson, F. (2012). Cognitive penetration of colour experience. *Philosophy and Phenomenological Research*, 84, 24–62.
- Martuzzi, R., Murray, M. M., Michel, C. M., Thiran, J. P., Maeder, P. P., Clarke, S., et al. (2007). Multisensory interactions within human primary cortices revealed by BOLD dynamics. *Cerebral Cortex*, 17, 1672–1679.
- Murphy, S. E., O’Donoghue, M. C., Drazich, E. G. S., Blackwell, S. E., & Nobre, A. C. (2015). Imagining a brighter future: The effects of positive imagery training on mood, prospective mental imagery and emotional bias in older adults. *Psychiatry Research*, 230, 36–43.
- Murphy, M. C., Nau, A. C., Fisher, C., Kim, S. G., Schuman, J. S., & Chan, K. C. (2016). Top-down influence on the visual cortex of the blind during sensory substitution. *NeuroImage*, 125, 932–940.
- Nanay, B. (2010a). Perception and Imagination: Amodal perception as mental imagery. *Philosophical Studies*, 150, 239–254.

- Nanay, B. (2010b). Attention and perceptual content. *Analysis*, 70, 263–270.
- Nanay, B. (2015). Perceptual content and the content of mental imagery. *Philosophical Studies*, 172, 1723–1736.
- Nanay, B. (2016a). Imagination and perception. In A. Kind (Ed.), *Routledge handbook of philosophy of imagination*. London: Routledge.
- Nanay, B. (2016b). Hallucination as mental imagery. *Journal of Consciousness Studies*, 23(7–8), 65–81.
- Nanay, B. (2017a). Sensory substitution and multimodal mental imagery. *Perception*. <http://dx.doi.org/10.1177/0301006617699225>. in print.
- Nanay, B. (2017b). Pain as mental imagery. *The Monist*. in print.
- Nanay, B. (2017c). Multimodal completion and naturalized epistemology. In D. Gatzia, & B. Brogaard (Eds.), *The rational roles of perceptual experience: Beyond vision*. New York: Oxford University Press.
- Nanay, B. (2018). *Seeing things you don't see*. Oxford: Oxford University Press (forthcoming).
- Nunn, J. A., Gregory, L. J., Brammer, M., Williams, S. C. R., Parslow, D. M., Morgan, M. J., et al. (2002). Functional magnetic resonance imaging of synesthesia: Activation of V4/V8 by spoken words. *Nature Neuroscience*, 5(4), 371–375.
- O'Callaghan, C. (2008a). Object perception: Vision and audition. *Philosophy Compass*, 3, 803–829.
- O'Callaghan, C. (2008b). Seeing what you hear: Cross-modal illusions and perception. *Philosophical Issues*, 18(1), 316–338.
- O'Callaghan, C. (2014). Intermodal binding awareness. In D. Bennett, & C. Hill (Eds.), *Sensory integration and the unity of consciousness*. Cambridge, MA: MIT Press.
- Page, J. W., Duhamel, P., & Crognale, M. A. (2011). ERP evidence of visualization at early stages of visual processing. *Brain and Cognition*, 75(2), 141–146.
- Paivio, A. (1969). Mental Imagery in associative learning and memory. *Psychological Review*, 76(3), 241–263.
- Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart, and Winston.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford, England: Oxford University Press.
- Papies, E. K., & Barsalou, L. W. (2015). Grounding desire and motivated behavior: A theoretical framework and empirical evidence. In W. Hofmann, & L. F. Nordgren (Eds.), *The psychology of desire*. NY: Guilford.
- Pearson, J., Clifford, C. W. G., & Tong, F. (2008). The functional impact of mental imagery on conscious perception. *Current Biology*, 18(13), 982–986.
- Pearson, J., & Westbrook, F. (2015). Phantom perception: Voluntary and involuntary nonretinal vision. *Trends in Cognitive Sciences*, 19, 278–284.
- Pearson, Joel, Naselaris, Thomas, Holmes, Emily A., & Kosslyn, Stephen M. (2015). Mental imagery: Functional mechanisms and clinical applications. *Trends in Cognitive Sciences*, 19, 590–602.
- Peck, T. C., Seinfeld, S., Aglioti, S. M., & Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and Cognition*, 22, 779–787.
- Pekkola, J., Ojanen, V., Autti, T., Jääskeläinen, I. P., Möttönen, R., Tarkiainen, A., et al. (2005). Primary auditory cortex activation by visual speech: An fMRI study at 3 T. *NeuroReport*, 16, 125–128.
- Phillips, I. (2014). Lack of imagination: Individual differences in mental imagery and the significance of consciousness. In J. Kallestrup, & M. Sprevak (Eds.), *New waves in philosophy of mind*. Palgrave Macmillan.
- Pryor, J. (2000). The skeptic and the dogmatist. *Noûs*, 34(4), 517–549.
- Ratner, K. G., Dotsch, R., Wigboldus, D., van Knippenberg, A., & Amodio, D. M. (2014). Visualizing minimal ingroup and outgroup faces: Implications for impressions, attitudes, and behavior. *Journal of Personality and Social Psychology*, 106, 897–911.
- Richardson, A. (1969). *Mental imagery*. New York: Springer.
- Roseboom, W., Kawabe, T., & Nishida, S. (2013). The cross-modal double flash illusion depends on featural similarity between cross-modal inducers. *Scientific Reports*, 3, 3437–4359.
- Ryle, G. (1949). *The concept of mind*. London: Hutchinson.
- Sampaio, E., Maris, S., & Bach-y-Rita, P. (2001). Brain plasticity: 'Visual' acuity of blind persons via the tongue. *Brain Research*, 908(2), 204–207.
- Sekuler, R., Sekuler, A. B., & Lau, R. (1997). Sound alters visual motion perception. *Nature*, 285, 308.
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature*, 408, 788.
- Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171, 701–703.
- Siegel, S. (2011). Cognitive penetrability and perceptual justification. *Nous*, 46, 201–222.
- Slofstra, C., Nauta, M. H., Holmes, E. A., & Bockting, C. L. H. (2016). Imagery rescripting: The impact of conceptual and perceptual changes to aversive autobiographical memories. *PLoS One*, 11(8), e0160235.
- Slotnick, S. D., Thompson, W. L., & Kosslyn, S. M. (2005). Visual mental imagery induces retinotopically organized activation of early visual areas. *Cerebral Cortex*, 15, 1570–1583.
- Smilek, Daniel, Callejas, Alicia, Dixon, Mike J., & Merikle, Philip M. (2007). Ovals of time: Timespace associations in synaesthesia. *Consciousness and Cognition*, 16(2), 507–519.
- Spence, C., & Deroy, O. (2013). Crossmodal imagery. In S. Lacey, & R. Lawson (Eds.), *Multisensory imagery* (pp. 157–183). New York: Springer.
- Spence, C., & Driver, J. (Eds.). (2004). *Crossmodal space and crossmodal attention*. Oxford: Oxford University Press.
- 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.
- Stokes, D. (2012). Perceiving and desiring: A new look at the cognitive penetrability of experience. *Philosophical Studies*, 158, 479–492.
- Strawson, P. F. (1974). Imagination and perception. In P. F. Strawson (Ed.), *Freedom and resentment* (pp. 45–65). London: Methuen.
- Talavage, T. M., Sereno, M. I., Melcher, J. R., Ledden, P. J., Rosen, B. R., & Dale, A. M. (2004). Tonotopic organization in human auditory cortex revealed by progressions of frequency sensitivity. *Journal of Neurophysiology*, 91, 1282–1296.
- Teufel, C., & Nanay, B. (2017). How to (and how not to) think about top-down influences on perception. *Consciousness and Cognition*, 47, 17–25.
- Van Essen, D. C. (2004). Organization of visual areas in macaque and human cerebral cortex. In L. Chalupa, & J. S. Werner (Eds.), *The visual neurosciences*. MIT Press.
- Vroomen, J., Bertelson, P., & de Gelder, B. (2001). Auditory-visual spatial interactions. In B. de Gelder, et al. (Eds.), *Out of mind* (pp. 140–150). Oxford: Oxford University Press.
- Walton, K. (1990). *Mimesis as make-believe*. Cambridge, MA: Harvard University Press.
- Ward, Jamie, Huckstep, Brett, & Tsakanikos, Elias (2006). Sound-colour synaesthesia: To what extent does it use cross-modal mechanisms common to us all? *Cortex*, 42(2), 264–280.
- Ward, J., & Simner, J. (2003). Lexical-gustatory synaesthesia: Linguistic and conceptual factors. *Cognition*, 89(3), 237–261.
- Watkins, S., Shams, L., Tanaka, S., Haynes, J. D., & Rees, G. (2006). Sound alters activity in human V1 in association with illusory visual perception. *NeuroImage*, 31, 1247–1256.
- Williams, A. D., Blackwell, S. E., Mackenzie, A., Holmes, E. A., & Andrews, G. (2013). Combining imagination and reason in the

- treatment of depression: A randomized controlled trial of internet-based cognitive-bias modification and internet-CBT for depression. *Journal of Consulting and Clinical Psychology*, 81, 793–799.
- Winterowd, C., Beck, A. T., & Gruener, D. (2003). *Cognitive therapy with chronic pain patients, Eliciting and modifying imagery*. New York: Springer Publishing Co.
- Zangaladze, A., Weisser, V. D., Stilla, E., Prather, S. C., & Sathian, K. (1999). Involvement of visual cortex in tactile discrimination of orientation. *Nature*, 401, 587–590.
- Zatorre, R. J., & Halpern, A. R. (2005). Mental concert: Musical imagery and auditory cortex. *Neuron*, 47, 9–12.
- Zeman, A., McGonigle, D., Gountouna, E., Torrens, L., Della Sala, S., & Logie, R. (2007). “Blind imagination”: Brain activation after loss of the mind's eye. *Journal of Neurology Neurosurgery & Psychiatry*, 78(2), 209.
- Zeman, A. Z. J., Della Sala, S., Torrens, L. A., Gountouna, V.-E., McGonigle, D. J., & Logie, R. H. (2010). Loss of imagery phenomenology with intact visuo-spatial task performance: A case of 'blind imagination'. *Neuropsychologia*, 48, 145–155.
- Zeman, A., Dewar, M., & Della Sala, S. (2015). Lives without imagery: Congenital aphasia. *Cortex*, 73, 378–380.