



## Review

**Cite this article:** Deroy O, Chen Y-C, Spence C. 2014 Multisensory constraints on awareness. *Phil. Trans. R. Soc. B* **369**: 20130207. <http://dx.doi.org/10.1098/rstb.2013.0207>

One contribution of 13 to a Theme Issue 'Understanding perceptual awareness and its neural basis'.

### Subject Areas:

neuroscience

### Keywords:

perceptual awareness, multisensory, crossmodal, binding, integration, unity of consciousness

### Author for correspondence:

Ophelia Deroy  
e-mail: [ophelia.deroy@sas.ac.uk](mailto:ophelia.deroy@sas.ac.uk)

# Multisensory constraints on awareness

Ophelia Deroy<sup>1</sup>, Yi-Chuan Chen<sup>2</sup> and Charles Spence<sup>3</sup>

<sup>1</sup>Centre for the Study of the Senses and Institute of Philosophy, School of Advanced Study, University of London, London, UK

<sup>2</sup>Department of Psychology, Neuroscience and Behaviour, McMaster University, Hamilton, Ontario, Canada

<sup>3</sup>Department of Experimental Psychology, Oxford University, Oxford, UK

Given that multiple senses are often stimulated at the same time, perceptual awareness is most likely to take place in multisensory situations. However, theories of awareness are based on studies and models established for a single sense (mostly vision). Here, we consider the methodological and theoretical challenges raised by taking a multisensory perspective on perceptual awareness. First, we consider how well tasks designed to study unisensory awareness perform when used in multisensory settings, stressing that studies using binocular rivalry, bistable figure perception, continuous flash suppression, the attentional blink, repetition blindness and backward masking can demonstrate multisensory influences on unisensory awareness, but fall short of tackling multisensory awareness directly. Studies interested in the latter phenomenon rely on a method of subjective contrast and can, at best, delineate conditions under which individuals report experiencing a multisensory object or two unisensory objects. As there is not a perfect match between these conditions and those in which multisensory integration and binding occur, the link between awareness and binding advocated for visual information processing needs to be revised for multisensory cases. These challenges point at the need to question the very idea of multisensory awareness.

If we think that phenomenal experience is of any relevance to psychology or cognitive science or plays any role in our mental life or behaviour, then we will have to have some grasp on it. Marcel [1, p. 171].

## 1. Introduction

Given the frequency of cases in which different senses are stimulated (especially if we include the vestibular and proprioceptive senses which are almost constantly sending signals to the brain), there are reasons to consider that the paradigmatic case of perceptual awareness that we have to account for is likely to be multisensory. As stressed by several authors

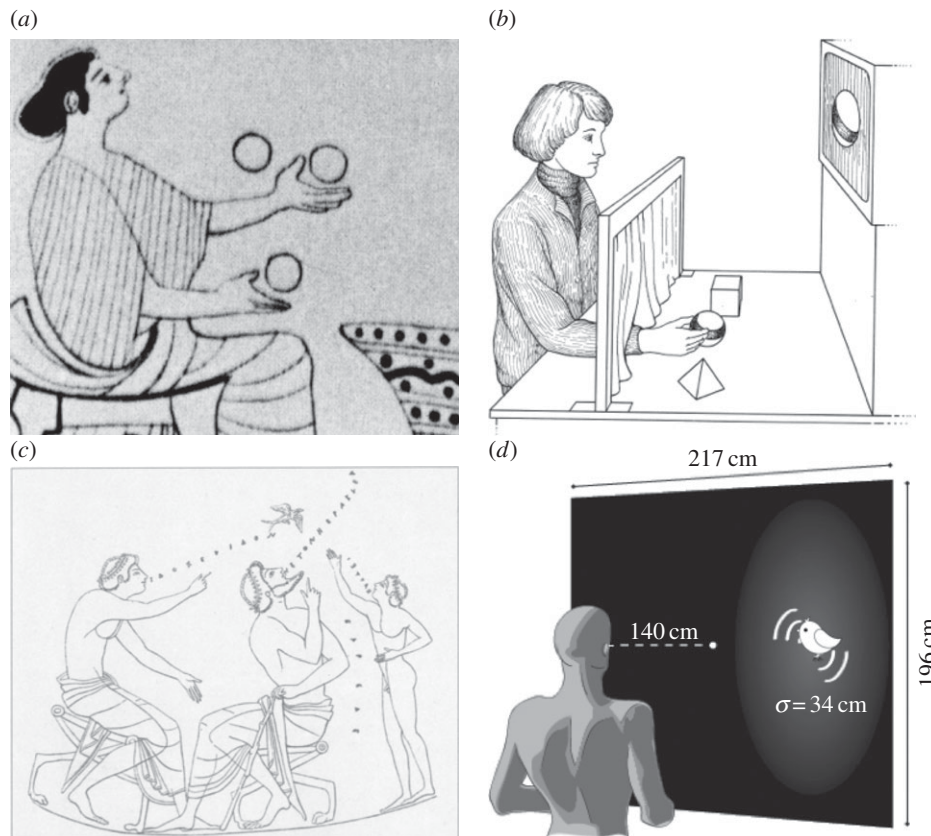
A fundamental aspect of our sensory experience is that the information from different modalities is often seamlessly integrated into a unified percept. Angelaki *et al.* [2, p. 452].

In our daily lives we perceive each multisensory event as an integral of multiple unisensory signals by using not only vision but also audition and the sense of touch. Fujisaki *et al.* [3, p. 301].

The richness of perceptual experience, as well as its usefulness for guiding behaviour, depends on the synthesis of information across multiple senses. Fetsch *et al.* [4, p. 429].

Multisensory awareness, however, seems to have fallen into the blind spot of current research. To date, theories of perceptual awareness have typically considered the phenomenon as if it were a unisensory, mostly visual, construct [5–7]. Over the past few years, there has also been a growth of interest in the study of awareness in several of the other senses as well (such as in touch, see [8,9]; audition, see [10]; and even olfaction, see [11]) but without a major change: perceptual awareness remains studied on a modality-by-modality, or unisensory, basis.

This unimodal perspective is abandoned in multisensory research, but the focus, then, is not so much on awareness *per se*, but rather on 'multisensory processing that is, how the brain deals simultaneously with information from



**Figure 1.** Multisensory awareness and its empirical study classically encompass two kinds of cases. In (a), the same common sensible or putatively ‘amodal’ property (e.g. shape) is present in consciousness via two different modes of awareness (e.g. visual and tactile) leading to a visuotactile experience of shape; experimental designs similar to that shown in (b) show that the visual awareness of the shape about which the participants have to report in a case of crossmodal matching includes tactile elements. In (c), two modally specific properties (e.g. the colour/look of the bird and its song) are bound into a single audio-visual object (that is, a singing bird), and this integration crucially depends upon the synchrony (or, more likely, the statistical correlation [19]) between the component unisensory signals, as studied in protocols such as (d). (Figure (c) adapted from [http://www.pc.rhul.ac.uk/staff/J.Zanker/PS1061/L10/PS1061\\_10.htm](http://www.pc.rhul.ac.uk/staff/J.Zanker/PS1061/L10/PS1061_10.htm); and (d) from [19], with modifications.)

different sensory systems’ [12, p. 3]. The main argument against the unisensory perspective is, in fact, that situations of multisensory stimulation often lead to a behavioural or neural response that ‘is significantly different (e.g. larger, smaller) from the best component response’ obtained during unisensory stimulation (e.g. [13, p. 1717]; see also [14]). The integration of information from different sensory modalities is, then, often supposed to be matched at the level of awareness by a genuinely multisensory experience:

To perceive the external environment our brain uses multiple sources of sensory information derived from several different modalities, including vision, touch and audition. All these different sources of information have to be efficiently merged to form a coherent and robust percept. Ernst & Bühlhoff [15, p. 161].

In addition to processing information on a sense-by-sense basis, the brain is responsible for assembling the rich *mélange* of information from the various sensory modalities into a coherent and meaningful ‘view’ of the world. Wallace *et al.* [16, p. 252].

It is important to stress that the kind of multisensory awareness mentioned here does not consist of unimodal experiences being jointly present in awareness, while keeping their individual characters. One can be aware of the sound of the bell while watching one’s computer screen, for instance, but the auditory awareness of the sound and the visual awareness of the screen do not necessarily mix. They merely seem to coexist and, according to a philosophical distinction, they constitute cases of ‘phenomenal unity’ [17]. Cases of genuinely multisensory awareness are different, as they imply that the

information coming from the different senses merges into a single unitary experience.

One kind of multisensory awareness typically includes cases in which the information concerning a single property of an object coming from different sensory channels is integrated into what people then want to label as an *amodal* representation of that property (though see [18] for a critical view of the notion of amodality). For instance, our awareness of the ball that we may be holding in our hands will be both visual and tactile; our awareness of our hands is itself presumably both visual and proprioceptive (figure 1*a,b*). Another kind of multisensory experience corresponds to cases where modality-specific features, such as coloured shapes and sounds are experienced together as features of one and the same object or event (figure 1*c,d*). One is, then, aware of a multisensory object or event—for instance, a yellow bird singing. These two kinds of cases of multisensory awareness can look different in the sense that the first concerns the integration of information into a single property, and the second the integration of parts into a single object (or whole). The processes underlying them are then sometimes distinguished as being processes of either ‘multisensory integration’ or ‘crossmodal binding’. However, following the literature on visual binding, it is common to see them treated together, as if they were instances of the same general process—and the distinction between integration and binding starts to blur. ‘Binding’ and ‘integration’ are then interchangeably used to refer to the process by which information concerning distinct sensory features

(such as tactile shape and visual shape, or sounds and colour) is bundled together as information concerning a common perceptible item. Although stated for vision, Treisman's point below is often applied to multisensory cases

(i) the binding of properties—the integration of different object properties, such as the color red and the shape + to form a red cross; (ii) the binding of parts—the integration in a structured relationship of the separately codable parts of a single object, such as the nose, eyes, and mouth in a face... The first two seem to me to be closely related and to depend on the same binding mechanism. Treisman [20, p. 98].

Research on multisensory cases does not only follow the model developed for visual research regarding the assimilation of property and object integration. Just as for vision, the operation of binding is seen as a necessary condition for consciousness [20–27].

Can we simply take the current theories and protocols used to try and understand unisensory cases and then import them into the field of multisensory research? This is the approach that we wish to question here. As argued below, shifting to multisensory cases is not cost-free for the study of perceptual awareness. It introduces both methodological and theoretical pressures—what we broadly call here 'constraints'. In §2, we focus on those methodological constraints presented by multisensory cases and review the limitations on the canonical protocols that are used to study unisensory awareness when considered in a multisensory setting. In our recent search through the literature, we have been unable to find a crucial experiment in which the results unambiguously demonstrate that a genuine case of multisensory awareness has occurred. If there is no doubt that subjective reports will refer to multisensory objects, or that the responses obtained in multisensory cases differ from a simple addition of the unisensory stimuli, it is very difficult to find an objective measure to confirm that awareness is genuinely multisensory. In §3, we review the reasons that might lead one to dissociate multisensory integration or binding from multisensory awareness. In §4, we turn to the theoretical constraints, showing why the shift to multisensory cases raises fundamental questions about the unified or dis-unified nature of consciousness, and the role of attention in determining our conscious experience of the objects and events in the world around us. Here, we introduce two alternative models that are capable of accounting for the experience of multisensory objects without claiming that awareness itself becomes multisensory. According to these two alternative models, awareness remains unimodal.

## 2. Methodological constraints

The past decade or so of studies on the neural basis of awareness, combined with decades of research on the topic of implicit or unconscious processing, have helped lead to the design of robust empirical protocols to evidence cases where awareness occurs from those cases where it does not. The most famous examples here might be Owen *et al.*'s [28] studies of those most unfortunate of patients who find themselves stuck in a vegetative state, some of whom demonstrate responsiveness to verbal instructions—which, in turn, has been taken as evidence that the contents of the latter were consciously perceived (but see [29,30] for discussion). Multisensory cases raise quite a different challenge, as what needs to be evidenced is not

the occurrence of a conscious episode, but rather the occurrence of a certain mode of awareness (the 'merged' or multisensory percept mentioned in §1). The evidence collected, therefore, needs to be incompatible with other interpretations, such as the occurrence of merely unisensory modes of awareness. How successful are the various protocols used to study perceptual awareness in a unisensory setting in terms of addressing this major question?

As reviewed below, drawing in part on our recent interest and work to investigate this question, experiments can successfully demonstrate a crossmodal modulation of participants' unisensory awareness. The majority of experimental studies have involved the presentation of a stimulus that is ambiguous or difficult to identify in one sensory modality (e.g. vision); at the same time, a clear and unambiguous stimulus is presented in another modality (e.g. audition). Presumably, or so the logic goes, if the sensory inputs from the two modalities are integrated, the clear and unambiguous one should bias the interpretation of the less clear, or more ambiguous, one, thus giving rise to a clearer perception of the ambiguous stimulus. These important results, however, fall short of locating the facilitatory effect in multisensory awareness. Most of these protocols, at most, merely demonstrate that the information presented in one sensory modality will boost the occurrence of perceptual awareness of stimuli presented in another modality.

### (a) Ambiguous displays

One sort of evidence that would support the multisensory awareness view would be if any perceptual alternations in the interpretation of an ambiguous display presented in one sensory modality were to be influenced by (or somehow linked to) the perceptual alternations taking place in another modality. However, to date, there is surprisingly little evidence in support of such a claim (see [29,30] for empirical evidence; and [31,32] for reviews). The only study that we are aware of that comes close to what we are after here was reported by Sato *et al.* [33]. These researchers investigated the auditory and visual verbal transformation effect. In the auditory version of this phenomenon [34], a participant listens to a speech stimulus that is presented repeatedly, for example the syllable '/psə/'. After a number of repetitions, the sound will be perceived to alternate and the observer likely will hear it as '/səp/' instead. As time passes by, the percept then alternates back and forth between the two possible interpretations.

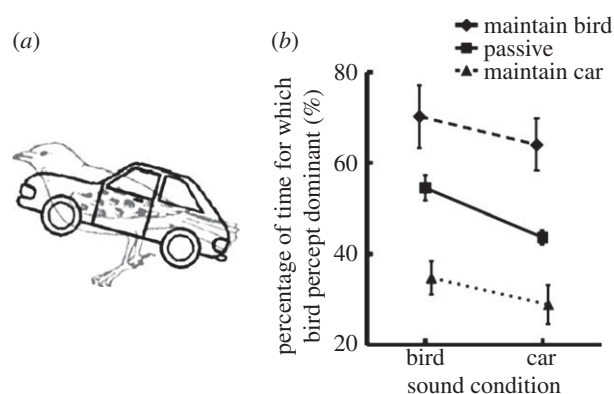
Sato *et al.* [33] discovered that the same thing happens if an observer looks at moving lips that happen to be repeatedly uttering the same syllable (this is known as the visual transformation effect). In their study, Sato and his colleagues presented auditory alone, visual alone and audio-visual stimulus combinations that were either congruent or incongruent (i.e. an auditory '/psə/' paired with a visual '/səp/'). The participants were instructed to report their initial auditory 'percept', and then to report whenever it changed over the course of the 90 s of each trial. The incongruent audio-visual condition, in which the visual stimulus alternated between being congruent and incongruent with what was heard, resulted in a higher rate of perceptual alternations than any of the other three conditions (though see [35] for the suggestion that speech may constitute a unique case when it comes to multisensory multistability).

## (b) Binocular rivalry

The phenomenon of binocular rivalry can be seen as providing a fascinating window into visual awareness in humans [36]. Binocular rivalry occurs when two dissimilar figures are presented to corresponding regions of each eye. Observers perceive one of the figures as dominant (while often being unaware of the other). After a while, the dominance of the figures may well reverse and then keep alternating over time. This perceptual alternation has been attributed to the fact that the visual system receives ambiguous information and hence tries to find a unique perceptual solution, and therefore the two images compete for control of the current conscious percept (see [37] for a review). Several studies have tried to understand how visual awareness emerges in the binocular rivalry situation. According to Helmholtz's [38] early view, the alternation of perceptual dominance is under voluntary attentional control. Later researchers, however, have argued that the phenomenon occurs as a result of the competition between two monocular channels [39], or else between two pattern representations, one presented to either eye [40]. More recent models [41] have suggested that the mechanism that underlies binocular rivalry includes not only competition at multiple levels of information processing (for a review, see [42]), but also some form of excitatory connections that facilitate the perceptual grouping of visual stimuli [43], as well as top-down feedback, for example attentional control [44–46]. That said, the underlying mechanisms giving rise to conscious perception in the binocular rivalry situation likely involve a variety of neural structures throughout the visual processing hierarchy.

van Ee *et al.* [47] conducted the first study to demonstrate that the information from another sensory modality, in this case either audition or touch, can modulate people's visual awareness in the binocular rivalry situation. In their study, the participants were presented with a looming pattern to one eye and a rotating pattern to the other eye. The dominance duration of the looming (or rotating) pattern was shown to be enhanced when the rate at which the pattern happened to move was synchronized with a series of pure tones or vibrotactile stimuli (or their combination). Nevertheless, it is important to note that this crossmodal modulatory effect was only observed when the participants endogenously attended to the visual pattern that was consistent with the rate of stimulus presentation in the other modality. Therefore, on the basis of this crossmodal modulation of participants' awareness, one cannot exclude the possibility that the information presented to the other sensory modality was modulating visual awareness via an attentional process.

To date, a growing number of studies have documented the multisensory modulation on the perception of binocular rivalry in terms of various crossmodal correspondences. For example, an auditory modulation of binocular rivalry has been shown to result from coherent moving direction (e.g. motion on the left/right axis [48]; or motion on the looming/receding axis [49]). Crossmodal modulations have also been observed in terms of the semantic coherence of natural/artificial objects (such as the sound of a bird-chirping or a car-engine revving, see [50]). Meanwhile, the tactile modulation of binocular rivalry occurs when the grating in the visual and tactile modalities has the same orientation [51,52]. Finally, the odour of rose/marker pen has been shown to enhance the dominance duration of the corresponding visual percept



**Figure 2.** (a) An example of the dichoptic stimulus pairs used in Chen *et al.*'s [50] study. The black and grey represent red and cyan (or vice versa), respectively. (b) The results demonstrated that the percentage of time that the bird percept was dominant was higher when the bird-chirping soundtrack was presented in the background than when the sound of a car-revving was played instead. Note that this crossmodal congruency effect was significant and the magnitude was similar in all three attention conditions.

[53,54] or vice versa (i.e. visual modulation on binocular rivalry, see [55]).

In all of the cases reported above, the multisensory inputs can, to a certain extent at least, be argued to be coherent. The possibility therefore remains that the information in another sensory modality provides a cue that simply directs a participant's attention toward the congruent percept.<sup>1</sup> Chen *et al.* [50] therefore designed an experiment in order to try and tease apart the auditory semantic and the attentional modulation of binocular rivalry (figure 2): they manipulated the target of participants' selective attention over the dichoptic figure independently from the meaning of the sound. That is, their participants were instructed to maintain the bird percept, to maintain the car percept or else to view the figures passively in the control condition [45,46]; meanwhile, the participants either heard the sound of birds chirping or else an engine revving. Their results demonstrated that both the auditory semantic context and the participant's voluntary attention influenced the participants' dominance percept and that these two factors appeared to work in an additive fashion. A further analysis of dominance duration of the bird/car percept (that was not reported in the published paper) revealed that the presentation of the auditory semantic context was more likely to shorten the duration of the semantically incongruent percept than to prolong the duration of the semantically congruent one. Note that this effect is different from the modulatory effect of attention on binocular rivalry that resulted from a prolonging of the dominance duration of the attended percept [44]. In summary, Chen *et al.*'s results suggest that the modulation of binocular rivalry by crossmodal congruency and attention can, to a certain extent, be dissociated, and hence that the former should not always be thought of as relying on the latter.

## (c) Bistable figure perception

Bistable figure perception constitutes another phenomenon whereby the presentation of an ambiguous static figure can elicit a dynamically alternating percept: when an observer looks at an ambiguous visual figure that affords two possible interpretations (such as the Necker cube or the Rubin face/

vase figure), they may perceive either one of the interpretations initially, and then become aware of the other interpretation a short while thereafter. Subsequently, as time goes by, the observer's perception will likely alternate spontaneously between the two possible interpretations. The critical difference between binocular rivalry and bistable figure perception lies in the fact that in the latter condition, observers watch the same ambiguous figure with both eyes; that is, the visual stimulus cannot be segregated into two visual patterns early in visual information processing. This difference certainly leads to the suggestion that the level at which the visual competition occurs should be higher for the case of bistable figure perception than in the case of binocular rivalry. As a result, it has been suggested that the perception of bistable figures may be more susceptible to modulation by attention than binocular rivalry [45,46].

Crossmodal modulations of bistable figure perception have been reported using, for example, the face/vase figure modulated by speech stimuli [58] and the 'my wife or my mother-in-law' bistable figure [59] modulated by female voices of recognizably different ages [60]. In Hsiao *et al.*'s [60] work, the modulatory effects of auditory semantic congruency and attention were further compared. In particular, the sound that the participants heard (the voice of either the old woman or the young lady) and the participants' attention over the bistable figure (when either instructed to try and maintain the view of the old woman, to maintain the view of the young lady or to view the stimulus passively) were manipulated independently. Interestingly, the results demonstrated that even though the effect of auditory semantic congruency was significant in all three attention conditions, its magnitude was significantly less pronounced when the participants were attending voluntarily to either the old woman or young lady views (as compared to the passive viewing condition). That said, the modulations attributable to auditory semantic congruency and attention in the bistable figure perception case are difficult to dissociate completely. In addition, combining the results of research on the auditory modulation of binocular rivalry and bistable figure perception can be taken to suggest that crossmodal modulations likely represent a heterogeneous set of empirical phenomena (or, put another way, they are determined by different mechanisms) in various situations (see also [61]).

#### (d) Continuous flash suppression

Continuous flash suppression (CFS) is the name given to a recently developed method that uses the mechanism of interocular suppression in order to maintain a visual stimulus in a subliminal state for a few seconds [62,63]. In the typical experimental setting, an observer is presented with a low-contrast stimulus to one eye and a dynamically changing high-contrast colourful Mondrian-type pattern to the other eye in the overlapping visual field of each eye. This presentation protocol often results in the low-contrast visual stimulus being suppressed and hence being unavailable for conscious report by the observer. If the contrast of the suppressed stimulus is gradually increased, however, it eventually breaks through to consciousness and hence can then be detected (and reported on) by the observer. In this paradigm, the response time (RT) for participants to detect the suppressed stimulus is usually used as the behavioural index of awareness.

Alsius & Munhall [64] used CFS to demonstrate that RTs were shorter when a suppressed talking face was paired with

a matched auditory speech stream, as compared with an incongruent one (though note that Palmer & Ramsey [65] have demonstrated that the unaware, or suppressed, face does not modulate auditory speech identification). Similarly, Yang & Yeh [66] have reported that RTs to detect a static face are modulated by various face-voice correspondences that include timbral features (e.g. mom's or baby's laughter) and emotional prosody (e.g. happy or angry). Using CFS can, then, be used by researchers to provide robust evidence that audio-visual integration can occur at an unconscious level and that it can, in turn, boost the subliminal visual stimulus to the conscious level.

#### (e) The attentional blink

When two different letters serving as targets are embedded within a string of digits and all of these stimuli are presented one-by-one in a rapid serial visual presentation (RSVP) stream, the participants frequently miss the second target (T2) after having detected or identified the first (T1), especially when T2 is presented close in time to T1 (typically between 100 and 500 ms later). This phenomenon, known as the attentional blink (AB), was first reported by Raymond *et al.* [67]. Note that even though T2 may not be consciously detected or reported by the observer, it can nevertheless still elicit a priming effect to the stimulus presented afterward [68]. That said, T2 is processed while suppressed below the level of conscious awareness.

Olivers & van der Burg [69] demonstrated that when a tone is presented at the same time as T2, participants are more likely to report it correctly. However, if the tone is presented with the item that precedes T2 instead, no such crossmodal facilitation is observed; that said, this crossmodal facilitatory effect was not attributable to the tone alerting the participants regarding the imminent arrival of T2 (see also [70] for a review of the literature in this area). Interestingly, the magnitude of the AB can be increased (i.e. the accuracy of T2 given T1 was correctly reported) when T1 is a letter paired with a different rather than same letter name presented auditorily [71]. That said, the crossmodal modulation of the AB is closely linked to participants' devoting attentional resources to the processing of T1.

#### (f) Repetition blindness

Repetition blindness (RB) refers to the phenomenon whereby when an item is presented twice within an RSVP stream, observers may miss its second occurrence ([72]; note that RB occurs within rapid serial auditory presentation streams as well; see [73]). It has been suggested that RB reflects a failure by the observer to individuate the two occurrences of the same item; that is, the type of that item has been recognized while people were unaware of its second occurrence [74].

Chen & Yeh [75] presented two tones that were simultaneous with the onset of the two repeated items. Participants were more likely to correctly report the repeated item twice (i.e. the RB effect was reduced), compared with when the tones were not presented. In a follow-up study, Chen & Yeh [76] manipulated the two tones so that they had either the same or different pitch. Once again, the presentation of the tones effectively reduced the magnitude of the RB effect to a similar extent, which suggests that temporal synchrony is the critical factor in the audio-visual facilitation in the RB paradigm.

### (g) Backward masking

In the backward masking paradigm, two stimuli are presented, one after the other, from more or less the same spatial location, with the first serving as the target and the second as the mask. Backward masking has frequently been used in previous cross-modal studies in order to control the visibility of a visual target by manipulating the stimulus onset asynchrony between target and mask. Within a certain range of task difficulties, the presentation of a simultaneous (or asynchronous within a range of about  $\pm 200$  ms) sound has been shown to enhance the accuracy with which participants detect or identify the visual target. Such crossmodal facilitation effects have been reported under various conditions of backward masking. So, for example, when the perception of the target light emitting diode (LED) is degraded by the subsequent presentation of four surrounding LEDs (e.g. [77,78]; see also [79]), or when a target letter is followed by a masking letter [80].

### (h) Interim summary

This far, we have reviewed the latest research to have emerged from those studies that have used a number of different experimental paradigms to demonstrate that visual stimuli are processed despite having been suppressed in an unconscious state. By presenting a stimulus in another sensory modality (typically audition) that is coherent, or congruent, with the visual stimulus in a certain regard (such as in terms of its temporal synchrony or semantic content), multisensory integration may occur, or the other sensory input may guide attention to the coherent (or congruent) visual stimulus [65].

## 3. The limits of the binding-awareness link

It is unquestionably the case that individuals report experiencing objects or features across two or more sensory modalities. They talk about singing birds, warm blue cups or whistling kettles. The assumption, though, is that these reports correspond to an awareness that is genuinely multisensory in nature and that this maps onto multisensory processes of binding or integration. We want to point out, however, that the existence of this mapping is put under pressure by various lines of evidence.

### (a) Evidence of multisensory interactions does not constitute evidence of multisensory awareness

It is fair to say that most of the evidence that has been collected about multisensory interactions has not focused on the relation between these processes and consciousness. At the level of human behaviour, the speed and accuracy of a participant's performance in a multisensory task setting has often been shown to be significantly different from their baseline level of performance obtained under conditions of unisensory stimulation. Performance with auditory and visual targets presented together will, for instance, be better than that with the auditory or visual target presented alone. Whether participants perform better because they are aware of a single audio-visual target or because of the influence of one type of sensory information on the processing of another is, however, not demonstrated (see [81] for discussion).

At the neurological level, the occurrence of multisensory integration is evidenced by the fact that the responses of single neurons (or else populations of neurons, or even specific

voxels or groups of voxels in neuroimaging studies) are significantly different in a specific multisensory setting from the sum of the responses that have been observed in a unisensory setting [13,82]. An alternative index of multisensory integration at a whole-brain level may be revealed by studying the synchronized oscillation of neural activity. That is, when visual and auditory stimuli are presented close together in space and/or time, or semantically congruent, neural activities coherent in phase in the gamma-band frequency range can often be observed [83,84]. This, or other forms of activation, though, fail to say anything about the nature of an observer's conscious experience, and whether participants (or the anaesthetized animals on which the majority of the early multisensory integration research was conducted) experience a unified object or event is simply not evidenced by such measures. This said, neurological studies might provide evidence of whether multisensory integration occurred in the brain and whether it is the focus of attention; see for instance, the study by Fairhall & Macaluso [85] which reported that multisensory integration (proved by subcortical activation in the superior colliculus) grabs attentional resources (as shown by activations in high-order fronto-parietal areas) even when entirely task-irrelevant.

### (b) Subjective reports do not establish the occurrence of multisensory awareness

The method that gets closest to investigating the occurrence of multisensory awareness consists in looking for contrast cases at the level of subjective report. The principle is to establish when subjective reports are consistent with the experience of a single object integrating the features processed in two (or more) modalities and to contrast them with reports of two separate features or events in two modalities. This technique has been widely used in the field of multisensory research. For instance, individuals have frequently been asked which syllable has been uttered when presented with incongruent speech sounds and clips of lip movements (for instance, speech sounds corresponding to the syllable 'ba' and lips articulating a 'ga', see [86]). In some conditions, participants report experiencing a syllable ('da'; this is known as the McGurk effect) which is inconsistent with what they should have experienced if they relied only on audition and suggest that their experience has been influenced by the information that has been presented visually. In another famous case, Shams *et al.* [87] had their participants say whether they were aware of one or two flashes when jointly presented with a single visual flash and two brief sounds.

Here, though, the perceptual reports can indicate the occurrence of audio-visual awareness, but are also compatible with the experience itself being in one sensory modality (auditory for the speech stimulus in the McGurk effect, visual for the individuation of flashes in the two-flash illusion): the experience in one modality might well have been influenced by the second source of information presented in another modality which does not, however, appear as such to form a unified multisensory conscious episode. The fact that participants only report experiencing a multisensory object when the component stimuli are presented within a certain temporal and/or spatial window, could mean that spatio-temporal proximity is necessary for participants to judge that the two pieces of information refer to one and the same object (e.g. [88]; though see also [89]).

We are actually sceptical that any contrast case will help to decide between cases where information coming from two sensory modalities are genuinely experienced together and cases where they are experienced one after the other and merely judged to have come from the same source and to belong to the same object. That said, let us be clear that this does not mean that we do not see a conceptual difference between the two [81].

In order to make some progress here, the study of multisensory consciousness should not be limited to the study of perceptual qualitative reports about objects and events. Take all those incidental reports concerning accompanying feelings—such as feelings of unity or disunity, and of incongruence and congruence [90] that have been noted in studies resting on the presentation of stimuli in different sensory modalities. Are they not conscious manifestations as well? Anyone who has experienced the McGurk effect can try it for themselves and see that, when the auditory and visual stimuli start to fall too far apart in time [91], or when a male voice is paired with a synchronized female face, one may feel slightly ill at ease (see also [92,93]). Jackson [94] presented his participants with the sound of a whistling kettle coming from one of seven different locations. He then asked them to determine from which of seven visible kettles (placed behind a Perspex screen) by which they were surrounded the sound was coming from. As Jackson himself admitted, the participants reported a mild feeling or impression of ‘intermodal conflict’ when they received an auditory stimulus and a simultaneous visual cue separately, although they supposed and perceived that it was from one and the same object. Many of the participants maintained that they ‘must be imagining’, or that they had experienced a ‘detached feeling’, or even nausea when they received an auditory stimulus and a simultaneous visual stimulus separately, although they supposed and perceived that it was from one and the same object. This might signal a gradual breach in the experience of a single object even in what presumably count as successful cases of multisensory integration or another kind of conscious manifestation.

Metacognitive feelings—of which one is also aware—are then an obvious category to investigate in relation with the awareness that manifests itself in multisensory situations, but which seem to have disappeared from experimental reports since the 1970s (just see [90,95] for a couple of engaging early examples). Researchers no longer seem to want to ask their participants what they felt about an experiment or set-up or stimulus display, perhaps because of the difficulties associated with the treatment of their participants’ (admittedly sometimes strange) replies. However, asking participants how confident they felt about their perception of a single object, or else starting to pay attention to individual differences in subjective reports is a desirable change that should be imported from unisensory to multisensory studies of consciousness.

### (c) Dissociating multisensory integration and multisensory awareness

Confirming our initial diagnosis, the investigation of multisensory awareness is not central in many studies documenting multisensory interactions, and establishing whether and when awareness is multisensory rather than unisensory is problematic. Several recent studies have pushed things further

and advocate a dissociation between the occurrences of multisensory integration and multisensory awareness.

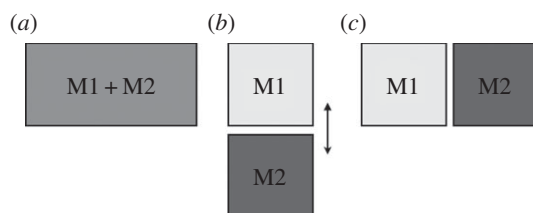
The most recent demonstration comes from a study, using a paradigm used for visual binding, showing that reports of experiencing a single audio-visual object were not perfectly aligned with the independent evidence of multisensory integration [96]. The authors concluded that ‘feature binding was entirely unaffected by conscious experience: features were bound whether they were experienced as occurring together or as belonging to a separate events, suggesting that the conscious experience of unity is not a prerequisite for, or a direct consequence of binding’ ([96], p. 586). Meanwhile, research on the Colavita visual dominance effect provides evidence of multisensory integration in the absence of awareness of one of the component stimuli (normally it is the visual stimulus that extinguishes awareness, or at least report, of some proportion of the auditory stimuli on bimodal target trials [97]).

More pressure also comes from the claim that even certain reasonably well-established examples of multisensory interaction (and assumed cases of multisensory awareness) may not actually necessitate there being any kind of integration in the first place. Over the years, many studies have, for instance, taken violations of the Miller inequality under conditions of simultaneous audio-visual stimulation as evidence for multisensory integration. The redundant target effect shows that responses are faster when visual and auditory target stimuli requiring a speeded detection response are presented together [98] than when they are presented separately. However, as argued recently by Otto & Mamassian [99], the results of such studies do not necessarily show that sensory evidence is integrated multisensorially prior to the participant making a perceptual decision, but rather may have been accumulated separately for each signal (see also [100]). In addition, Chen & Spence’s [101] study demonstrating that simultaneously presented semantically congruent pictures and sounds can modulate a participant’s response criterion without necessarily impacting on their perceptual sensitivity (as indexed by  $d'$ ) in a picture-detection task provides another good example here. One thing to investigate is whether individuals would then be happy to report having been aware of multisensory objects or events. If not, these subjective reports of integrated object would not necessarily correspond to a process of multisensory integration.

In summary, contrary to the visual paradigm, binding and awareness come apart in multisensory cases. They even present a form of double dissociation: awareness is not necessary for all cases of multisensory integration; multisensory integration is not necessary for the attribution of features to a single object across sensory modality.

## 4. Theoretical challenges: choosing among three competing models

Does the fact that several sensory modalities cooperate to build our awareness of the objects and events in the environment around us necessarily mean that awareness (or what philosophers might prefer to call ‘phenomenal consciousness<sup>2</sup>’) is itself multisensory? How can we be sure that a situation of multisensory stimulation leads to the awareness of a unified multisensory object (figure 3a) rather than merely to a sequence, or juxtaposition, of dis-unified unisensory states of awareness (figure 3b,c)?



**Figure 3.** What kind of awareness occurs in multisensory situations where sensory modalities S1 and S2 (let us take audition and vision) are stimulated simultaneously? The figure presents three options, with M1 and M2 corresponding to the kind of awareness posited for unisensory stimulation of senses S1 and S2, respectively: the model (a) posits a multisensory (audio-visual) percept where the content and the phenomenal characters of M1 and M2 combine; the second option (b) (as defended by Spence & Bayne [104], see also [105]) considers that only a single unisensory percept is conscious at any given time and that awareness switches rapidly (albeit with some delay) back and forth between M1 and M2; the third option (c) accepts that M1 and M2 are both conscious at each time but that awareness is not necessarily unified ('dis-unified awareness'). Here, attention might help to make one state of awareness more salient than the other, or else gain preferential access to awareness [106].

Whereas the aforementioned evidence is indeed compatible with a unified multisensory awareness (for instance, an audio-visual conscious episode), it is also compatible with at least two other distinct possibilities which maintain that each conscious experience remains unisensory. The two alternative models resort to a fast alternation between episodes of unisensory awareness so that, at any one moment in time, one is only aware of one object or feature in one sensory modality. What explains the behavioural results, and the report of experiencing a single object with features or aspects that belong to different modalities, is the fast and unnoted attentional shifts between the distinct unisensory experiences. It might be, then, that the unattended experience is simply not conscious, but brought in consciousness when attention shifts to the corresponding modality (figure 3*b*). Awareness, being the outcome of where (and on what) attention is focused, switches between the two features depending on an observer's voluntary attentional set or based on the pattern of external (or exogenous) stimulation that happens to be impinging on the sensory end organs at any given time [104]. Alternatively, however, it might be that the unattended experience is conscious, but simply not salient (figure 3*c*). In both cases, the main claim is that for the brain to deal with simultaneous information from different sensory modalities might not always, or necessarily, lead to the relevant information being integrated into unified multisensory object representations that are then made available to awareness (see also [81,107] for discussion).

One reason to resist the view that awareness is just a 'seamless percept' is the observation that cases of multisensory stimulation have never been shown to confuse observers with regard to the sensory modalities in which they have had an experience: with the exception of the confusing case of flavour, which might itself anyway have to count as a single sense [108,109], individuals can still tell that they saw a flash, heard a beep and felt a stroke of their hand. In the rubber-hand illusion, the result of seeing the rubber hand being stroked while being synchronously stroked is felt as a change in one's proprioceptive awareness [110]. If the percept is so unified, and if modes of awareness

are fused, it seems difficult to address the question of how one recovers the modal signature of each component feature (unless this is done outside of consciousness, see [111]). Thinking about awareness keeping a form of inner distinction or being rather like a juxtaposition of modal awareness, which could be part visual and part auditory, is difficult to reconcile with the view that awareness is just multisensory through and through. What this shows is that theories of multisensory awareness need to go beyond the generic or merely metaphorical talk of 'seamless percepts' or 'unity' and explain the mechanisms at stake in the sharing.

One thing we feel that the multisensory awareness approach explains more elegantly than the other approaches is the attribution of features belonging to distinct sensory modalities to one and the same object. As we said earlier, this is certainly part of everyday phenomenological reports and it seems simpler to think that awareness just mirrors these reports. However, the possibility cannot be excluded that reports of experiencing multisensory objects or features could come from a rapid oscillation between unisensory experiences and the unity of the object could be simply part of the cognitive treatment of this shifting of information. If the solution is plausible, though, it certainly needs to be developed further: how does the switch from, say, the awareness of a sound to the awareness of a shape tell us that this sound is of this shape or that the sound and the shape refer to one and the same object? Is spatial co-location preserved in the switch, and is it sufficient to give what one expects from the attribution to a single object or event [89]? Various replies could be developed here: the unity of objects across sensory modalities could be a cognitive or conceptual construct, and not directly experienced. There is no reason why unisensory and multisensory objects should necessarily be explained in the same way, and binding could, for example, explain the first and not the second. Recently, several authors (see [112] for an overview) have also started to look back into early Gestalt principles of connectedness and proximity, as perhaps more relevant ways to think about the relations between the unisensory experiences than those offered by the multisensory awareness model (see also [32,113]).

## 5. Conclusion

In this article, we have attempted to stress a gap between the observation that paradigmatic cases of awareness occur in multisensory settings and the fact that awareness remains studied in a unisensory framework. The recycling of unisensory protocols is unlikely to provide good ways to study multisensory awareness, if there is indeed such a thing. One additional element not to be missed is the possible role of attention in multisensory cases (see [114,115] for reviews, and §2*b,c*). If there is much to expect from the growing body of research in this area, we would like to suggest that other possible manifestations of multisensory interactions—such as metacognitive feelings of congruence—should also become a central focus.<sup>3</sup>

Our main conclusion is that multisensory cases of perceptual awareness do not just challenge empirical methods of investigating awareness, but raise fundamental questions regarding our theories of awareness—including the prospect of unveiling its neural correlates. The natural assumption when it comes to cases of multisensory stimulation is that the different modes of awareness generated in unisensory



cases can all homogeneously be integrated into the multisensory awareness of a multisensory object or event. So, for instance, the awareness of two yellow balls noisily bouncing against one another is treated as an audio-visual percept that integrates both the visual awareness of the yellow balls and the auditory awareness of the bouncing sounds (e.g. [116]; though see [117]). The view here clearly seems to be that awareness is of the same kind regardless of the particular sense that happens to be under consideration: that is, there would appear to be nothing different, as far as our theories of conscious awareness are concerned, between being aware of a sound and being aware of a patch of colour, or say a taste. These are all just modes of perceptual awareness, but not different kinds of consciousness—the reason, then, why people never talk about a plurality of consciousness. The quest for the neural basis of awareness, in this sense, has been made a general project that spans across the various sensory modalities and one does not expect different replies to arise for each of the senses. However, simple and elegant, this homogeneity assumption remains to be tested, along with the principle of unified multisensory awareness.

## References

- Marcel AJ. 2003 Introspective report. *J. Conscious. Stud.* **10**, 167–186.
- Angelaki DE, Gu Y, DeAngelis GC. 2009 Multisensory integration: psychophysics, neurophysiology, and computation. *Curr. Opin. Neurobiol.* **19**, 452–458. (doi:10.1016/j.conb.2009.06.008)
- Fujisaki W, Kitazawa S, Nishida S. 2012 Multisensory timing. In *The new handbook of multisensory processes* (ed. BE Stein), pp. 301–317. Cambridge, MA: MIT Press.
- Fetsch CR, DeAngelis GC, Angelaki DE. 2013 Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. *Nat. Rev. Neurosci.* **14**, 429–442. (doi:10.1038/nrn3503)
- Baars B. 1997 *In the theater of consciousness: the workspace of the mind*. Oxford, UK: Oxford University Press.
- Singer W. 1998 Consciousness and the structure of neuronal representations. *Phil. Trans. R. Soc. Lond. B* **353**, 1829–1840. (doi:10.1098/rstb.1998.0335)
- VanRullen R, Koch C. 2003 Is perception discrete or continuous? *Trends Cogn. Sci.* **7**, 201–213. (doi:10.1016/S1364-6613(03)00095-0)
- Gallace A, Spence C. 2010 Touch and the body: the role of the somatosensory cortex in tactile awareness. *Psyche* **16**, 30–67 ([http://www.theassc.org/journal\\_psyche/archive/vol\\_16\\_no\\_1\\_2010](http://www.theassc.org/journal_psyche/archive/vol_16_no_1_2010)).
- Gallace A, Spence C. 2014 *In touch with the future: the sense of touch from cognitive neuroscience to virtual reality*. Oxford, UK: Oxford University Press.
- Snyder JS, Gregg MK, Weintraub DM, Alain C. 2012 Attention, awareness, and the perception of auditory scenes. *Front. Psychol.* **3**, 15. (doi:10.3389/fpsyg.2012.00015)
- Stevenson RJ. 2009 Phenomenal and access consciousness in olfaction. *Conscious. Cogn.* **18**, 1004–1017. (doi:10.1016/j.concog.2009.09.005)
- Meredith MA. 2012 Commentary: multisensory convergence: where it all begins. In *The handbook of multisensory processing* (ed. BE Stein), pp. 3–11. Cambridge, MA: MIT Press.
- Stein BE *et al.* 2010 Semantic confusion regarding the development of multisensory integration: a practical solution. *Eur. J. Neurosci.* **31**, 1713–1720. (doi:10.1111/j.1460-9568.2010.07206.x)
- Holmes NP, Spence C. 2005 Multisensory integration: space, time, and superadditivity. *Curr. Biol.* **15**, R762–R764. (doi:10.1016/j.cub.2005.08.058)
- Ernst MO, Bühlhoff HH. 2004 Merging the senses into a robust percept. *Trends Cogn. Sci.* **8**, 162–169. (doi:10.1016/j.tics.2004.02.002)
- Wallace MT, Robertson GE, Hairston WD, Stein BE, Vaughan JW, Schirillo JA. 2004 Unifying multisensory signals across time and space. *Exp. Brain Res.* **158**, 252–258. (doi:10.1007/s00221-004-1899-9)
- Bayne T. 2010 *The unity of consciousness*. Oxford, UK: Oxford University Press.
- Spence C, Deroy O, Bremner A. 2013 Questioning the utility of the concept of amodality: towards a revised framework for understanding crossmodal relations. *Multisens. Res.* **26**, 57. (doi:10.1163/22134808-00050036)
- Parise CV, Spence C, Ernst M. 2012 When correlation implies causation in multisensory integration. *Curr. Biol.* **22**, 46–49. (doi:10.1016/j.cub.2011.11.039)
- Treisman A. 2003 Consciousness and perceptual binding. In *The unity of consciousness: binding, integration, and dissociation* (ed. A Cleeremans), pp. 95–113. Oxford, UK: Oxford University Press.
- Crick F, Koch C. 1990 Towards a neurobiological theory of consciousness. *Sem. Neurosci.* **2**, 263–275.
- Engel AK, Singer W. 2001 Temporal binding and the neural correlates of sensory awareness. *Trends Cogn. Sci.* **5**, 16–25. (doi:10.1016/S1364-6613(00)01568-0)
- Treisman A. 1986 Features and objects in visual processing. *Sci. Am.* **255**, 114–125. (doi:10.1038/scientificamerican1186-1148)
- Treisman A. 1988 Features and objects: the fourteenth Bartlett memorial lecture. *Q. J. Exp. Psychol.* **40A**, 201–237. (doi:10.1080/02724988843000104)
- Treisman A. 1996 The binding problem. *Curr. Opin. Neurobiol.* **6**, 171–178. (doi:10.1016/S0959-4388(96)80070-5)
- Treisman A. 1998 Feature binding, attention and object perception. *Phil. Trans. R. Soc. Lond. B* **353**, 1295–1306. (doi:10.1098/rstb.1998.0284)
- van Leeuwen C. 2007 Synchrony, binding, and consciousness: how are they related? *Theory Psychol.* **17**, 779–790. (doi:10.1177/0959354307083493)
- Owen AM, Coleman MR, Davis MH, Boly M, Laureys S, Pickard JD. 2006 Detecting awareness in the vegetative state. *Science* **313**, 1402. (doi:10.1126/science.1130197)
- Celesia GG. 2013 Conscious awareness in patients in vegetative states: myth or reality? *Curr. Neurol. Neurosci. Rep.* **13**, 1–9. (doi:10.1007/s11910-013-0395-7)
- Hupé JM, Joffoa LM, Pressnitzer D. 2008 Bistability for audiovisual stimuli: perceptual decision is modality specific. *J. Vis.* **8**, 1–15. (doi:10.1167/8.7.1)
- Schwartz JL, Grimalt N, Hupé JM, Moore BC, Pressnitzer D. 2012 Multistability in perception:

**Funding statement.** O. Deroy and C. Spence are funded by the RTS-AHRC grant, within the ‘Science in Culture’ scheme.

## Endnotes

<sup>1</sup>One could legitimately wonder whether any mental imagery that is elicited by the information presented in another sensory modality [56] can serve as a top-down cue that may modulate perception in the binocular rivalry situation [57].

<sup>2</sup>Phenomenal consciousness refers to the fact that the world seems or appears a certain way from an observer’s point of view; it is sometimes distinguished from access consciousness (see [102] for a recent statement). The notion is close to Nagel’s [103] idea that there is something ‘it is like’ to perceive the world from a subjective point of view.

<sup>3</sup>Finally, what about other conscious activities, such as mental imagery or memory? Arguably, if an experience is multisensory, then the memory or re-enactment of that experience will also be multisensory—but does it really work so simply? On the one hand, it is currently unclear whether we have imagery for complex sensory experiences, for example, flavours. On the other hand, the rapid switching of awareness between unisensory episodes of imagery or memory also needs to be compared to explanations in term of multimodal episodes [56].

- binding sensory modalities, an overview. *Phil. Trans. R. Soc. B* **367**, 896–905. (doi:10.1098/rstb.2011.0254)
32. Spence C. In press. Cross-modal perceptual organization. In *The Oxford handbook of perceptual organization* (ed. J Wagemans). Oxford, UK: Oxford University Press.
  33. Sato M, Basirat A, Schwartz J. 2007 Visual contribution to the multistable perception of speech. *Percept. Psychophys.* **69**, 1360–1372. (doi:10.3758/BF03192952)
  34. Warren RM, Gregory RL. 1958 An auditory analogue of the visual reversible figure. *Am. J. Psychol.* **71**, 612–613. (doi:10.2307/1420267)
  35. Kubovy M, Yu M. 2012 Multistability, cross-modal binding and the additivity of conjoint grouping principles. *Phil. Trans. R. Soc. B* **367**, 954–964. (doi:10.1098/rstb.2011.0365)
  36. Crick F. 1996 Visual perception: rivalry and consciousness. *Nature* **379**, 485–486. (doi:10.1038/379485a0)
  37. Alais D, Blake R (eds). 2005 *Binocular rivalry*. Cambridge, MA: MIT Press.
  38. Helmholtz HV. 1962 *Helmholtz's treatise on physiological optics* (ed. JPC Southall). New York, NY: Dover Publications. (Original work published 1866).
  39. Levelt WJM. 1965 *On binocular rivalry*. Assen, The Netherlands: Royal VanGorcum.
  40. Leopold DA, Logothetis NK. 1996 Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature* **379**, 549–553. (doi:10.1038/379549a0)
  41. Tong F, Meng M, Blake R. 2006 Neural bases of binocular rivalry. *Trends Cogn. Sci.* **10**, 502–511. (doi:10.1016/j.tics.2006.09.003)
  42. Blake R, Logothetis NK. 2002 Visual competition. *Nat. Rev. Neurosci.* **3**, 1–11. (doi:10.1038/nrn701)
  43. Alais D, Blake R. 1999 Grouping visual features during binocular rivalry. *Vis. Res.* **39**, 4341–4353. (doi:10.1016/S0042-6989(99)00146-7)
  44. Chong SC, Tadin D, Blake R. 2005 Endogenous attention prolongs dominance durations in binocular rivalry. *J. Vis.* **5**, 1004–1012. (doi:10.1167/5.11.6)
  45. Meng M, Tong F. 2004 Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *J. Vis.* **4**, 539–551. (doi:10.1167/4.7.2)
  46. van Ee R, van Dam LCJ, Brouwer GJ. 2005 Voluntary control and the dynamics of perceptual bi-stability. *Vis. Res.* **45**, 41–55. (doi:10.1016/j.visres.2004.07.030)
  47. van Ee R, van Boxtel JJA, Parker AL, Alais D. 2009 Multimodal congruency as a mechanism for willful control over perceptual awareness. *J. Neurosci.* **29**, 11 641–11 649. (doi:10.1523/JNEUROSCI.0873-09.2009)
  48. Conrad V, Bartels A, Kleiner M, Noppeney U. 2010 Audiovisual interactions in binocular rivalry. *J. Vis.* **10**, 1–15. (doi:10.1167/10.10.27)
  49. Conrad V, Kleiner M, Bartels A, Hartcher-O'Brien J, Bülthoff HH, Noppeney U. 2013 Naturalistic stimulus structure determines the integration of audiovisual looming signals in binocular rivalry. *PLoS ONE* **8**, e70710. (doi:10.1371/journal.pone.0070710)
  50. Chen Y-C, Yeh S-L, Spence C. 2011 Crossmodal constraints on human visual awareness: can auditory semantic context modulate binocular rivalry? *Front. Percept. Sci.* **2**, 212. (doi:10.3389/fpsyg.2011.00212)
  51. Lungchi C, Binda P, Morrone MC. 2010 Touch disambiguates rivalrous perception at early stages of visual analysis. *Curr. Biol.* **20**, R143–R144. (doi:10.1016/j.cub.2009.12.015)
  52. Lungchi C, Alais D. 2013 Touch interacts with vision during binocular rivalry with a tight orientation tuning. *PLoS ONE* **8**, e58754. (doi:10.1371/journal.pone.0058754)
  53. Zhou W, Jiang Y, He S, Chen D. 2010 Olfaction modulates visual perception in binocular rivalry. *Curr. Biol.* **20**, 1356–1358. (doi:10.1016/j.cub.2010.05.059)
  54. Zhou W, Zhang X, Chen J, Wang L, Chen D. 2012 Nostril-specific olfactory modulation of visual perception in binocular rivalry. *J. Neurosci.* **32**, 17 225–17 229. (doi:10.1523/JNEUROSCI.2649-12.2012)
  55. Chen J, Zhou W, Chen D. 2012 Binaural rivalry in the presence of visual perceptual and semantic influences. *PLoS ONE* **7**, e47317. (doi:10.1371/journal.pone.0047317)
  56. Spence C, Deroy O. 2013 Crossmodal mental imagery. In *Multisensory imagery: theory and applications* (eds S Lacey, R Lawson), pp. 157–183. New York, NY: Springer.
  57. Pearson J, Clifford CWG, Tong F. 2008 The functional impact of mental imagery on conscious perception. *Curr. Biol.* **18**, 982–986. (doi:10.1016/j.cub.2008.05.048)
  58. Munhall KG, ten Hove MW, Brammer M, Paré M. 2009 Audiovisual integration of speech in a bistable illusion. *Curr. Biol.* **19**, 735–739. (doi:10.1016/j.cub.2009.03.019)
  59. Boring EG. 1930 A new ambiguous figure. *Am. J. Psychol.* **42**, 444–445. (doi:10.2307/1415447)
  60. Hsiao JY, Chen YC, Spence C, Yeh SL. 2012 Assessing the effects of audiovisual semantic congruency on the perception of a bistable figure. *Conscious. Cogn.* **21**, 775–787. (doi:10.1016/j.concog.2012.02.001)
  61. Blake R, Sobel KV, James TW. 2004 Neural synergy between kinetic vision and touch. *Psychol. Sci.* **15**, 397–402. (doi:10.1111/j.0956-7976.2004.00691.x)
  62. Fang F, He S. 2005 Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nat. Neurosci.* **8**, 1380–1385. (doi:10.1038/nn1537)
  63. Tsuchiya N, Koch C. 2005 Continuous flash suppression reduces negative afterimages. *Nat. Neurosci.* **8**, 1096–1101. (doi:10.1038/nn1500)
  64. Alsius A, Munhall K. 2013 Detection of audiovisual speech correspondences without visual awareness. *Psychol. Sci.* **24**, 423–431. (doi:10.1177/0956797612457378)
  65. Palmer TD, Ramsey AK. 2012 The function of consciousness in multisensory integration. *Cognition* **125**, 353–364. (doi:10.1016/j.cognition.2012.08.003)
  66. Yang YH, Yeh SL. 2011 Semantic congruency in audiovisual integration as revealed by the continuous flash suppression paradigm. *i-Perception*, **2**, 840.
  67. Raymond JE, Shapiro KL, Arnell KM. 1992 Temporary suppression of visual processing in an RSVP task: an attentional blink? *J. Exp. Psychol. Hum. Percept. Psychophys.* **18**, 849–860. (doi:10.1037/0096-1523.18.3.849)
  68. Shapiro K, Driver J, Ward R, Sorensen RE. 1997 Priming from the attentional blink: a failure to extract visual tokens but not visual types. *Psychol. Sci.* **8**, 95–100. (doi:10.1111/j.1467-9280.1997.tb00689.x)
  69. Olivers CNL, van der Burg E. 2008 Bleeping you out of the blink: sound saves vision from oblivion. *Brain Res.* **1242**, 191–199. (doi:10.1016/j.brainres.2008.01.070)
  70. Spence C, Ngo MK. 2012 Does attention or multisensory integration explain the crossmodal facilitation of masked visual target identification? In *The new handbook of multisensory processing* (ed. BE Stein), pp. 345–358. Cambridge, MA: MIT Press.
  71. Van der Burg E, Bredero SG, Nieuwenstein MR, Theeuwes J, Olivers CN. 2010 Audiovisual semantic interference and attention: evidence from the attentional blink paradigm. *Acta Psychol.* **134**, 198–205. (doi:10.1016/j.actpsy.2010.01.010)
  72. Kanwisher NG. 1987 Repetition blindness: type recognition without token individuation. *Cognition* **27**, 117–143. (doi:10.1016/0010-0277(87)90016-3)
  73. Soto-Faraco S, Spence C. 2002 Modality-specific auditory and visual temporal processing deficits. *Q. J. Exp. Psychol. A* **55**, 23–40. (doi:10.1080/02724980143000136)
  74. Kanwisher N. 2001 Neural events and perceptual awareness. *Cognition* **79**, 89–113. (doi:10.1016/S0010-0277(00)00125-6)
  75. Chen YC, Yeh SL. 2008 Visual events modulated by sound in repetition blindness. *Psychon. Bull. Rev.* **15**, 404–408. (doi:10.3758/PBR.15.2.404)
  76. Chen Y-C, Yeh S-L. 2009 Catch the moment: multisensory enhancement of rapid visual events by sound. *Exp. Brain Res.* **198**, 209–219. (doi:10.1007/s00221-009-1831-4)
  77. Bolognini N, Frassinetti F, Serino A, Ladavas E. 2005 'Acoustical vision' of below threshold stimuli: interaction among spatially converging audiovisual inputs. *Exp. Brain Res.* **160**, 273–282. (doi:10.1007/s00221-004-2005-z)
  78. Frassinetti F, Bolognini N, Ladavas E. 2002 Enhancement of visual perception by crossmodal visuo-auditory interaction. *Exp. Brain Res.* **147**, 332–343. (doi:10.1007/s00221-002-1262-y)
  79. McDonald JJ, Teder-Salejarvi WA, Hillyard SA. 2000 Involuntary orienting to sound improves visual perception. *Nature* **407**, 906–908. (doi:10.1038/35038085)
  80. Chen Y-C, Spence C. 2011 The crossmodal facilitation of visual object representations by sound: evidence from the backward masking

- paradigm. *J. Exp. Psychol. Hum. Percept. Perform.* **37**, 1784–1802. (doi:10.1037/a0025638)
81. O'Callaghan C. 2014 Perception and multimodality. In *Oxford handbook to philosophy and cognitive science* (eds E Margolis, R Samuels, S Stich), pp. 80–95. New York, NY: Oxford University Press.
  82. Stein BE, Meredith MA. 1993 *The merging of the senses*. Cambridge, MA: MIT Press.
  83. Bauer M. 2008 Multisensory integration: a functional role for inter-area synchronization? *Curr. Biol.* **18**, R709–R710. (doi:10.1016/j.cub.2008.06.051)
  84. Senkowski D, Schneider TR, Foxe JJ, Engel AK. 2008 Crossmodal binding through neural coherence: implications for multisensory processing. *Trends Neurosci.* **31**, 401–409. (doi:10.1016/j.tins.2008.05.002)
  85. Fairhall S, Macaluso E. 2009 Spatial attention can modulate audiovisual integration at multiple cortical and sub-cortical sites. *Eur. J. Neurosci.* **29**, 1247–1257. (doi:10.1111/j.1460-9568.2009.06688.x)
  86. McGurk H, MacDonald J. 1976 Hearing lips and seeing voices. *Nature* **264**, 746–748. (doi:10.1038/264746a0)
  87. Shams L, Kamitani Y, Shimojo S. 2000 What you see is what you hear: sound induced visual flashing. *Nature* **408**, 788. (doi:10.1038/35048669)
  88. Beierholm UR, Quartz SR, Shams L. 2009 Bayesian priors are encoded independently from likelihoods in human multisensory perception. *J. Vis.* **9**, 1–9. (doi:10.1167/9.5.23)
  89. Spence C. 2013 Just how important is spatial coincidence to multisensory integration? Evaluating the spatial rule. *Ann. NY Acad. Sci.* **1296**, 31–49. (doi:10.1111/nyas.12121)
  90. Holt-Hansen K. 1976 Extraordinary experiences during cross-modal perception. *Percept. Mot. Skills* **43**, 1023–1027. (doi:10.2466/pms.1976.43.3f.1023)
  91. Soto-Faraco S, Alsius A. 2009 Deconstructing the McGurk–MacDonald illusion. *J. Exp. Psychol. Hum. Percept. Perform.* **35**, 580–587. (doi:10.1037/a0013483)
  92. Navarra J. 2003 Visual speech interference in an auditory shadowing task: the dubbed movie effect. In *Proc. 15th International Congress of Phonetic Sciences, Barcelona, Spain* (ed. MJ Solé), pp. 1–4. Barcelona, Spain: ICPhS Organizing Committee.
  93. Vatakis A, Spence C. 2007 Crossmodal binding: evaluating the 'unity assumption' using audiovisual speech stimuli. *Percept. Psychophys.* **69**, 744–756. (doi:10.3758/BF03193776)
  94. Jackson CV. 1953 Visual factors in auditory localization. *Q. J. Exp. Psychol.* **5**, 52–65. (doi:10.1080/17470215308416626)
  95. Zapparoli GC, Reatto LL. 1969 The apparent movement between visual and acoustic stimulus and the problem of intermodal relations. *Acta Psychol.* **29**, 256–267. (doi:10.1016/0001-6918(69)90019-5)
  96. Zmigrod S, Hommel B. 2011 The relationship between feature binding and consciousness: evidence from asynchronous multi-modal stimuli. *Conscious. Cogn.* **20**, 586–593. (doi:10.1016/j.concog.2011.01.011)
  97. Spence C, Parise C, Chen Y-C. 2011 The Colavita visual dominance effect. In *Frontiers in the neural bases of multisensory processes* (eds MM Murray, M Wallace), pp. 523–550. Boca Raton, FL: CRC Press.
  98. Miller J. 1982 Divided attention: evidence for coactivation with redundant signals. *Cogn. Psychol.* **14**, 247–279. (doi:10.1016/0010-0285(82)90010-X)
  99. Otto TU, Mamassian P. 2012 Noise and correlations in parallel perceptual decision making. *Curr. Biol.* **22**, 1391–1396. (doi:10.1016/j.cub.2012.05.031)
  100. Harrar V, Tammam J, Pérez-Bellido A, Pitt A, Stein J, Spence C. In press. Multisensory integration and attention: differences in developmental dyslexia. *Curr. Biol.*
  101. Chen Y-C, Spence C. 2011b Crossmodal semantic priming by naturalistic sounds and spoken words enhances visual sensitivity. *J. Exp. Psychol. Hum. Percept. Perform.* **37**, 1554–1568. (doi:10.1037/a0024329)
  102. Block N. 2011 Perceptual consciousness overflows cognitive access. *Trends Cogn. Sci.* **15**, 567–575. (doi:10.1016/j.tics.2011.11.001)
  103. Nagel T. 1974 What is it like to be a bat? *Phil. Rev.* **83**, 435–456. (doi:10.2307/2183914)
  104. Spence C, Bayne T. In press. Is consciousness multisensory? In *Perception and its modalities* (eds M Matthen, S Biggs, D Stokes). Oxford, UK: Oxford University Press.
  105. Deroy O. In press. Multisensory integration and the many unities of consciousness. In *Multisensory integration and the unity of consciousness* (eds D Bennett, C Hill). Cambridge, MA: MIT Press.
  106. Spence C, Parise C. 2010 Prior entry. *Conscious. Cogn.* **19**, 364–379. (doi:10.1016/j.concog.2009.12.001)
  107. Fulkerson M. 2011 The unity of haptic touch. *Phil. Psychol.* **24**, 493–516. (doi:10.1080/09515089.2011.556610)
  108. Auvray M, Spence C. 2008 The multisensory perception of flavor. *Conscious. Cogn.* **17**, 1016–1031. (doi:10.1016/j.concog.2007.06.005)
  109. Spence C, Smith B, Auvray M. In press. Confusing tastes and flavours. In *Perception and its modalities* (eds M Matthen, S Biggs, D Stokes). Oxford, UK: Oxford University Press.
  110. Botvinick M, Cohen J. 1998 Rubber hands 'feel' touch that eyes see. *Nature* **391**, 756. (doi:10.1038/35784)
  111. Liang M, Moraux A, Hu L, Lannetti GD. 2013 Primary sensory cortices contain distinguishable spatial patterns of activity for each sense. *Nat. Commun.* **4**, 1979. (doi:10.1038/ncomms2979)
  112. Wagemans J. (ed.) In press. *The Oxford handbook of perceptual organization*. Oxford, UK: Oxford University Press.
  113. Spence C, Chen Y-C. 2012 Intramodal and crossmodal perceptual grouping. In *The new handbook of multisensory processing* (ed. BE Stein), pp. 265–282. Cambridge, MA: MIT Press.
  114. Santangelo V, Macaluso E. 2012 Spatial attention and audiovisual processing. In *The new handbook of multisensory processes* (ed. BE Stein), pp. 359–370. Cambridge, MA: The MIT Press.
  115. Talsma D, Senkowski D, Soto-Faraco S, Woldorff MG. 2010 The multifaceted interplay between attention and multisensory integration. *Trends Cogn. Sci.* **14**, 400–410. (doi:10.1016/j.tics.2010.06.008)
  116. Sekuler R, Sekuler AB, Lau R. 1997 Sound alters visual motion perception. *Nature* **385**, 308. (doi:10.1038/385308a0)
  117. Calvert G, Spence C, Stein BE. (eds) 2004 *The handbook of multisensory processing*. Cambridge, MA: MIT Press.