

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/225278414>

Beyond colour perception: Auditory-visual synaesthesia induces experiences of geometric objects in specific locations

Article in *Cortex* · April 2012

DOI: 10.1016/j.cortex.2012.04.006 · Source: PubMed

CITATIONS

16

READS

64

3 authors, including:



Anina N Rich

Macquarie University

59 PUBLICATIONS **1,549** CITATIONS

SEE PROFILE



Research report

Beyond colour perception: Auditory–visual synaesthesia induces experiences of geometric objects in specific locations

Rocco Chiou*, Marleen Stelter and Anina N. Rich*

Department of Cognitive Science, and ARC Centre of Excellence in Cognition and its Disorders, Macquarie University, Australia

ARTICLE INFO

Article history:

Received 2 August 2011

Reviewed 19 September 2011

Revised 10 December 2011

Accepted 15 April 2012

Action editor Henry Buchtel

Published online 27 April 2012

Keywords:

Synaesthesia

Audio–visual interactions

Feature-based attention

Feature binding

Cross-modal correspondences

ABSTRACT

Our brain constantly integrates signals across different senses. Auditory–visual synaesthesia is an unusual form of cross-modal integration in which sounds evoke involuntary visual experiences. Previous research primarily focuses on synaesthetic colour, but little is known about non-colour synaesthetic visual features. Here we studied a group of synaesthetes for whom sounds elicit consistent visual experiences of coloured ‘geometric objects’ located at specific spatial location. Changes in auditory pitch alter the brightness, size, and spatial height of synaesthetic experiences in a systematic manner resembling the cross-modal correspondences of non-synaesthetes, implying synaesthesia may recruit cognitive/neural mechanisms for ‘normal’ cross-modal processes. To objectively assess the impact of synaesthetic objects on behaviour, we devised a multi-feature cross-modal synaesthetic congruency paradigm and asked participants to perform speeded colour or shape discrimination. We found irrelevant sounds influenced performance, as quantified by congruency effects, demonstrating that synaesthetes were not able to suppress their synaesthetic experiences even when these were irrelevant for the task. Furthermore, we found some evidence for task-specific effects consistent with feature-based attention acting on the constituent features of synaesthetic objects: synaesthetic colours appeared to have a stronger impact on performance than synaesthetic shapes when synaesthetes attended to colour, and vice versa when they attended to shape. We provide the first objective evidence that visual synaesthetic experience can involve multiple features forming object-like percepts and suggest that each feature can be selected by attention despite it being internally generated. These findings suggest theories of the brain mechanisms of synaesthesia need to incorporate a broader neural network underpinning multiple visual features, perceptual knowledge, and feature integration, rather than solely focussing on colour-sensitive areas.

© 2012 Elsevier Ltd. Open access under [CC BY-NC-ND license](http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Our brains are constantly bombarded with signals from different sensory modalities. Although vision is usually considered the dominant modality, other senses, particularly

audition, interact closely with vision to create a coherent representation of our surroundings (Shimojo and Shams, 2001). Some atypical forms of cross-modal interactions, such as synaesthesia, result in percepts that do not represent events in the external world. Synaesthesia is an unusual

* Corresponding authors. ARC Centre of Excellence in Cognition and its Disorders, Macquarie University, NSW 2019, Australia.

E-mail addresses: roccochiou@gmail.com (R. Chiou), anina.rich@mq.edu.au (A.N. Rich).

0010-9452 © 2012 Elsevier Ltd. Open access under [CC BY-NC-ND license](http://creativecommons.org/licenses/by-nc-nd/4.0/).

doi:10.1016/j.cortex.2012.04.006

phenomenon in which stimulation in one sensory modality elicits additional anomalous experiences. These additional experiences can occur in the same modality (e.g., seeing colours when viewing achromatic letters: grapheme–colour synaesthesia) or in a different modality (e.g., seeing colours when listening to music: sound–colour synaesthesia). The prevalence of synaesthesia is relatively low, with estimates ranging from .5% (Baron-Cohen et al., 1996; Rich et al., 2005) to 5% (Simner et al., 2006) of the population. Synaesthesia has drawn much scientific attention in recent years due both to the interest inherent in anomalous brain phenomena, and to the insights these phenomena can give into normal mechanisms of perception and cognition.

There are two major hypotheses regarding the neural mechanisms that give rise to synaesthesia. The first view, generally termed the cross-activation hypothesis, suggests that excessive neural connections between adjacent cortical areas underlie synaesthetic experiences. Originally, this view postulated that grapheme–colour synaesthesia occurs as a result of excessive neural connections between colour-selective area V4 and the posterior temporal grapheme area (Hubbard and Ramachandran, 2005). More recently, these authors further proposed that the parietal lobe mediates the binding of synaesthetic colour and visual word form, presumably again through excessive connections with the temporal lobe (Hubbard, 2007; Hubbard et al., 2011). The idea that synaesthesia involves an anomalous form of feature binding, which implicates the parietal lobe, has also been raised by others, although not necessarily specifying excessive connections (Esterman et al., 2006; Mattingley et al., 2001; Robertson, 2003). The second view, generally called the disinhibited-feedback hypothesis, suggests that synaesthesia results from a ‘malfunctioning’ mechanism that fails to inhibit the crosstalk between brain areas normally inhibited in non-synaesthetic brain. According to different versions of this view, the disinhibition may occur in the feedback from multi-modal regions (e.g., superior temporal sulcus: Grossenbacher and Lovelace, 2001) or from areas involved in executive control (e.g., prefrontal cortex: Cohen Kadosh et al., 2009) to unimodal areas. These two mechanisms have been primarily proposed to explain how grapheme–colour and sound–colour synaesthesia might occur in the brain and have led to a number of behavioural and brain-imaging studies (e.g., Cohen Kadosh et al., 2009; Rouw and Scholte, 2007; Ward et al., 2006).

The two hypotheses differ in explaining how synaesthesia arises in the brain. Both, however, focus primarily on colour and V4 to explain the neural bases of synaesthesia. A few recent studies do report synaesthetic experiences other than colour (e.g., seeing another person being touched induced tactile sensation: Banissy and Ward, 2007; Fitzgibbon et al., 2011; perceiving music induces tastes: Beeli et al., 2005; seeing visual flashes induces auditory experiences: Saenz and Koch, 2008; reading words induces taste: Ward and Simner, 2003). However, such experiences occur in modalities other than vision, and it is currently not clear whether the proposed mechanisms for synaesthetic visual percepts are applicable to these forms of synaesthesia. When researching synaesthetic visual experiences, the majority of studies focus on synaesthetic colour. This seems to be due to two factors: first, grapheme–colour

synaesthesia is one of the most common and widely recognised subtypes (Novich et al., 2011; Rich et al., 2005; Simner et al., 2006), assisting recruitment of participants. Second, it is relatively easy to get estimates of synaesthetic colours, which makes it more conducive to objective measurement. For example, one can manipulate the congruency between physical and synaesthetic colours, and look at effects on colour naming time (e.g., Mattingley et al., 2001). This focus on colour is echoed in the major theories of synaesthesia, which do not place much emphasis, if any, on non-colour synaesthetic visual experiences. To construct a theory comprehensive enough to explain broader aspects of synaesthetic experience, it is therefore important to assess objectively the characteristics of non-colour synaesthetic features and their impacts on behaviour.

Eagleman and Goodale (2009) recently documented subjective reports of grapheme–colour and auditory–visual synaesthetes that suggest, in addition to colour, synaesthetic experiences can also have surface textures (e.g., i looks metallic). Based on the descriptions from synaesthetes, Eagleman and Goodale propose that, in addition to V4, synaesthesia may recruit other brain regions in the medial ventral stream, such as the areas involved in texture processing. There is so far no study reporting objective measure of non-colour synaesthetic visual features and quantifying their effects on behaviour.

Here we present an investigation of seven auditory–visual synaesthetes, each reporting visual experiences in response to sounds. Their auditorily-induced visual experiences appear as geometric objects, consisting of colour and shape (and sometimes texture), which appear in a particular location. In an initial session, we asked synaesthetes to illustrate their synaesthetic experiences. Visual experiences induced by different instrument sounds were consistent over time, and systematically varied in colour, shape, and spatial location in response to changes in auditory pitch and timbre. Specifically, we observed a consistent pattern across all synaesthetes for synaesthetic ‘objects’ to become smaller in size, brighter in colour, and higher in space as the auditory pitch got higher, analogous to the trends in implicit cross-modal correspondences observed in non-synaesthetes (Spence, 2011).

To objectively examine the impacts of the synaesthetic concurrents (in this case we call them ‘synaesthetic objects’ to emphasise the multidimensional nature) on behaviour, we devised a multi-feature version of the cross-modal synaesthetic congruency paradigm used by Ward et al. (2006). Synaesthetes and non-synaesthetic controls performed colour and shape discrimination tasks on visual targets. Prior to the target displays, we presented task-irrelevant sounds that elicited synaesthetic visual percepts that either matched (congruent) or mismatched (incongruent) the target images in colour and shape (Experiment 1), or in one of these features and spatial location (Experiment 2). We had two specific predictions. First, synaesthetes’ performance should be significantly influenced by the congruency between auditorily-induced synaesthetic features and displayed features. Despite controls presumably having implicit cross-modal correspondences between audition and vision, we would not expect similarly strong effects for controls, due to their lack of consciously perceived synaesthetic images, although it is possible that there may be subtle effects. Second, previous

research has demonstrated that task-relevant features of an irrelevant object can cause stronger distraction in visual search tasks relative to other task-irrelevant features of the same object (e.g., [Olivers et al., 2006](#)). Based on such feature-based modulatory effects, we expected the focus of the task to modulate the strength of the congruency effect such that when attending to the colour, synaesthetic colours should cause a stronger congruency effect than synaesthetic shapes, and vice versa when attending to shape.

2. Method

2.1. Participants

Fourteen individuals reporting auditory synaesthesia participated in the initial subjective session, in which we asked them to depict their synaesthetic experiences in response to sounds and evaluated their level of consistency across repetition of sounds. Six did not give consistent responses (details specified in the [Procedure](#) section), so we did not include them in subsequent experiments. Although inconsistent responses do not necessarily mean their synaesthesia is not genuine ([Simner, 2012](#)), it is not possible to test them using our paradigm, which relies on replicable responses. One additional individual did not participate because she experienced consistent colour and texture but no experiences of shape and location. Thus, seven individuals with consistent colour and non-colour synaesthetic experiences (two males; mean age (\pm SD): 32.7 ± 11.6 years; range: 21–50 years) participated in the subsequent assessments and experiments. They reported vivid visual experiences in response to auditory stimuli (voices, music, and ambient sounds). These visual experiences predominately resembled simple geometric objects (e.g., cube, sphere, or wavy line), and changes in auditory characteristics (pitch, timbre, and melody) altered the described hue, brightness, shape, and spatial location. All reported also seeing colours induced by graphemes. Five of them had musical training (one is a professional musician), but none reported having perfect pitch.¹ All seven synaesthetes were right-handed. We also tested seven sex-, age-, and handedness-matched non-synaesthetic controls (mean age (\pm SD): 32.5 ± 12.2 years; range: 21–50 years) for comparison in the main experiments. As controls do not have any kind of synaesthesia (criteria for inclusion in the control group), they did not participate in the subjective session. Four of the controls had music training (none had perfect pitch).

2.2. Stimuli

The auditory stimuli comprised 30 different instrument sounds, each of 2 sec duration. All sound clips were 16-bit stereo files at the sampling frequency of 44.1 kHz and 65 dB. The 30 sounds consisted of 10 flute notes, 10 piano notes, and 10 violin notes. The instrument notes were computer-synthesised, matched for frequency of the fundamental, and consisted of notes from C1 (33 Hz) up to Eb6 (1245 Hz), separated by intervals of musical fifths (i.e., 700 cents). Thus, the

following notes were used: C1, G1, D2, A2, E3, B3, F#4, Db5, Ab5, and Eb6.

2.3. Procedure

2.3.1. Subjective session

We mapped out the characteristics of responses to instrument sounds to see whether they varied systematically with timbre and pitch and whether there was any coherent pattern across synaesthetes. We also used the images generated in this session to construct stimuli to assess the specificity of the synaesthetic experiences and for our experimental manipulations.

We presented 60 sounds (30 different notes \times two repetitions) in a randomised order. After listening to each sound, the synaesthetes were asked to select their synaesthetic colour using the graphics software Gimp (<http://www.gimp.org>). If their synaesthetic percepts involved more than one colour or visual features other than colour, we asked them to draw their synaesthetic image using Gimp or pastels. We also asked them to provide as much additional description as possible. After drawing their synaesthetic experience for each sound, they were asked to rate how well their image matched their synaesthesia on a five-point scale, with ‘one’ being ‘poor match’ and ‘five’ being ‘perfect match’. They were encouraged to continue modifying the image until it was at least ‘three’ (‘good match’). We also administered a questionnaire probing the subjective locus of their synaesthetic experience, specifically asking whether their sound-induced synaesthetic images were perceived *internally* (in mind’s eye) or *externally* (out in space). The questionnaire also asked similar questions about mental imagery (e.g., picturing a familiar object in mind). They were encouraged to add descriptions if neither of the two options precisely depicted their experiences.

2.3.2. Consistency assessment

The aim of the consistency assessment was to evaluate the consistency of the reported synaesthetic experiences across two repetitions of sounds. Two independent raters evaluated consistency by comparing drawings and descriptions between the repetitions of the same sound. The evaluations were made based on the three prominent features in the synaesthetic experiences: (1) whether the chosen colours were similar in hue and saturation; (2) whether the reported objects were similar in shape and size; (3) whether the reported locations were similar in on-screen position and in their verbal descriptions of location. The raters used a binary scale (consistent/inconsistent) to rate the consistency of each feature (colour, shape, and location) associated with each sound. Responses were considered consistent only if all three dimensions were rated consistent. Based on these criteria, seven of the 14 synaesthetes were judged to give consistent reports in more than 90% of the pairs.

To ensure that the level of consistency of the seven synaesthetes was reliably higher than a level that would occur by chance, we randomly shuffled the pairings between images within each synaesthete, resulting in 30 random pairs for each synaesthete. We had a third independent rater, who was naïve to our research aim and had not seen the images from the subjective session before, judge the consistency of those random pairs, as well as that of the original pairs from the

¹ The effects were evident in all synaesthetic participants, regardless of musical background.

subjective session (presented in an intermingled order). This rater was instructed to use identical criteria to those adopted by the first two raters (i.e., a pair should only be deemed consistent when colour, shape, and location were all rated consistent) and the same binary scale (consistent vs inconsistent). The average rating given to random pairs was 19% [standard deviation (SD) = .10], providing us with a measure of how high a consistency level would be by chance alone. This was then compared to the drawings created by the synaesthetes, which were rated by this third rater as significantly higher than this chance level [71%, SD = .21; $t(6) = 10.74$, $p < .001$].

2.3.3. Specificity assessment

The aims of this test were to examine the specificity of the experiences and to test the consistency of the synaesthetes' reports over a longer period of time. It was conducted approximately 2 months after the initial session. We selected 10 sounds yielding 'very good' or 'perfect' match ratings during the subjective session for each of the seven participants with consistent synaesthetic experiences of geometric objects. For each participant, 40 individually tailored images (4 images for each of the 10 sounds) were created using Photoshop. Based on each individual's descriptions and illustrations of their synaesthetic experiences, one image for each of 10 sounds was constructed to replicate their experience (based on their hand-drawings, computer graphics, and verbal descriptions). We then made subtle variations in colour, shape, or location from the original images to create three 'foils' for each sound (see Fig. 1 for examples). In each trial, the synaesthete was presented with an instrument sound (2 sec) followed by an image (until response). The image could either be the one that represented their synaesthetic object or one of the three foils for that sound. They were asked to evaluate how well each image matched their synaesthetic experience on the same five-point

scale. Responses were considered consistent if they gave a rating of 'four' ('very good match') or 'five' ('perfect match') to the images that was generated to match their synaesthetic experience and a lower rating to the foils. The foils were highly similar to the original images. Thus, relative to our earlier consistency test in which the ratings were performed by independent raters, this specificity test provides a more rigorous examination of consistency and specificity. If the synaesthetic percepts were consistent over time and specific in their features, we would expect synaesthetes to give more ratings of 'very good match' or 'perfect match' to images created to replicate their synaesthetic objects, relative to foils that look very similar but differ subtly in one or two features. The assessment contained 40 trials. Stimulus presentation and response collection were controlled by E-Prime.

The mean percentage of re-rating the original images as 'very good/perfect match' was 88% (SD = .13), significantly greater than for foil images [67%; SD = .21; $t(6) = 3.41$, $p < .05$]. Note we expect some positive response to the foil images, as they were consistent in at least one of the three features we measured, but our synaesthetes' experiences were specific and consistent enough to identify the matching images over the highly similar foils.

2.3.4. Experiment 1

We developed a multi-feature version of a synaesthetic congruency paradigm to objectively measure the impact of synaesthetic colour and shape on behavioural performance. For each individual, we selected four sound–image pairs rated as 'very good match' or 'perfect match' in the test for feature specificity that had clearly distinguishable colours, shapes, and locations. We constructed a unique set of stimuli for each synaesthete by independently altering colour and shape of the images. An age-, gender- and handedness-matched non-

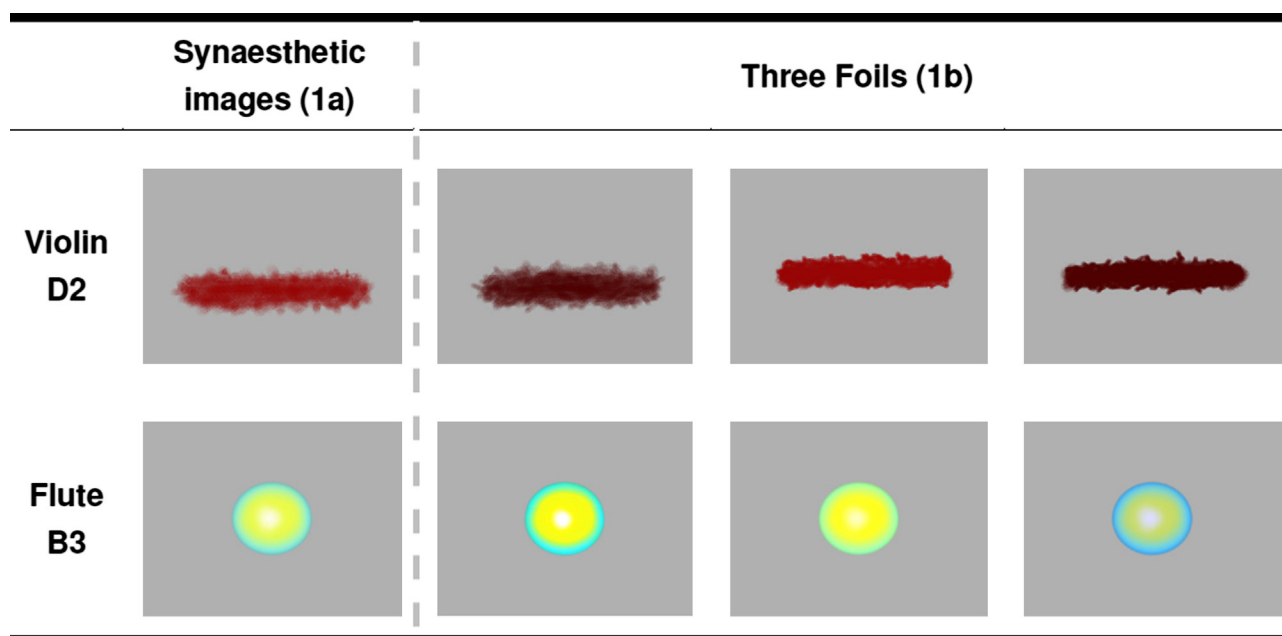


Fig. 1 – Examples of image stimuli used in the consistency assessment. (a) Images constructed based on a synaesthete's original drawings and descriptions, to replicate the synaesthetic experiences induced by the sounds of violin D2 and flute B3. (b) Three 'foils' developed to differ subtly from the synaesthetic images in colour, shape, or location.

synaesthetic control used the identical stimulus set as each synaesthete.

Participants performed two separate tasks on identical stimuli. In half of the blocks, they identified the colour of each item, using a four-alternative keypress. In the other half, they identified the shape of each item, again using a four-alternative keypress. The order of colour and shape tasks was counterbalanced across participants. In Experiment 1 (Fig. 2), we manipulated image colour and shape while keeping the on-screen location of the object congruent with the synaesthetic location elicited by the sound. On incongruent trials, the sound elicited a synaesthetic colour or shape that mismatched either colour or shape (or both) of the displayed image (a single incongruent colour and shape was selected for each sound based on the synaesthetic object elicited by another sound in the set; see Fig. 2). Thus, the synaesthetic colour and shape induced by sounds could match (congruent) or mismatch (incongruent) the colour and shape of the target, resulting in four different congruency conditions: (1) both colour and shape congruent; (2) colour congruent, shape incongruent; (3) colour incongruent, shape congruent; and (4) both colour and shape incongruent (see Fig. 2a–d). We therefore define congruency as having four levels, consistent with our conceptualisation that the ‘mixed’ congruency conditions (e.g., colour congruent/shape incongruent) are ‘partially incongruent’ conditions (for precedent, see Rich and Mattingley, 2003). In Supplementary Materials, we also provide the results of alternative analyses of both experiments in which each synaesthetic feature is treated as

an individual congruency factor. The results of the alternative analyses are consistent with those reported in the main article and enable us to make the same conclusion.

Prior to each task (colour or shape), participants completed 160 training trials on the mappings between the four keys and the stimulus features (colours or shapes). For training, we used centrally presented coloured squares or achromatic shapes, respectively, to avoid any hints about associations between the features. Each task consisted of a practice block of 24 trials and four experimental blocks of 48 trials, giving 48 trials in each congruency condition. The four conditions were randomly intermingled within a block, and each colour and shape was equally likely to appear in each of the four conditions. Throughout the experiment, they were told to respond to the task-relevant visual feature on the screen and ignore sounds and irrelevant visual dimensions. The experiment was controlled by MATLAB with Psychophysics Toolbox (Brainard, 1997; Pelli, 1997).

Each trial began with a black fixation dot on a grey background [RGB triplet = (176 176 176); 500 msec], followed by an instrumental sound presented for 2 sec before the onset of the target image. The sounds came from loudspeakers positioned to the left and right of the monitor. After the sound, a target image was presented for a maximum of 4 sec or until response. Participants had to press the designated keys with the index and middle fingers of their left and right hands to indicate image colour/shape once they saw the image. We asked them to respond as quickly and accurately as possible. Participants received feedback on accuracy on each trial (750 msec).

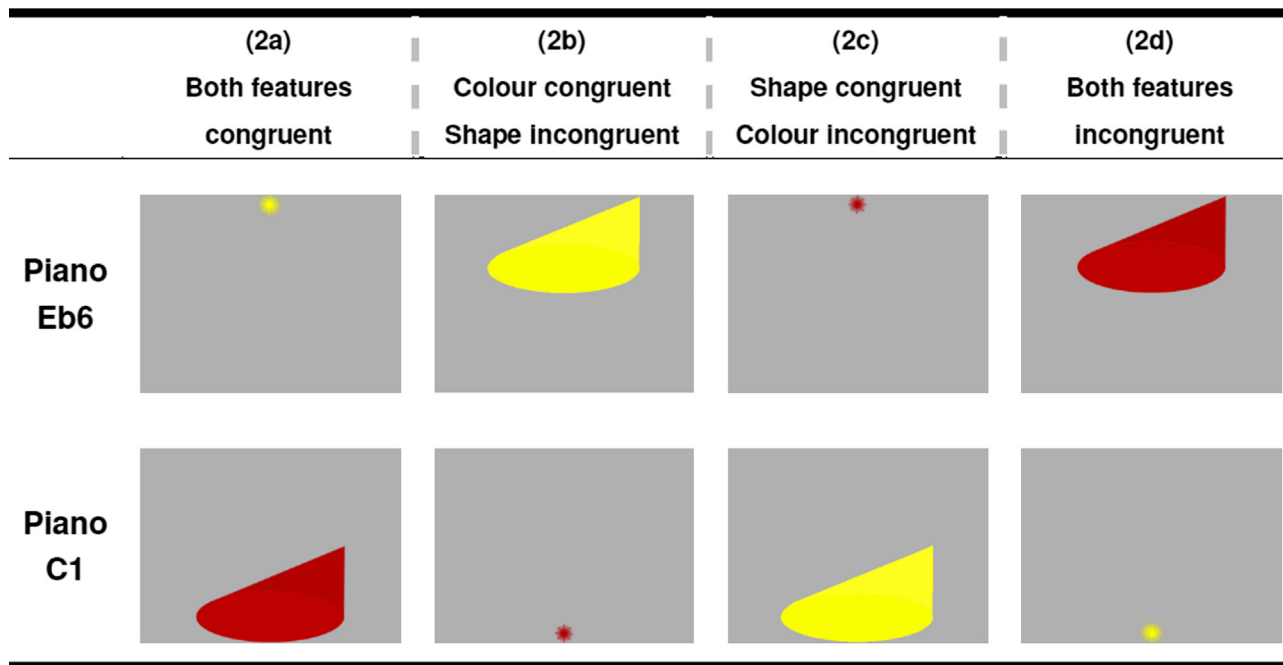


Fig. 2 – Examples of the four congruency conditions for Experiment 1 for one synaesthete, for whom the sound of piano Eb6 elicits a ‘small, glittering, golden star-shape object located high’ and piano C1 elicits a ‘big, squashed, oval, dark red cone located low’. (a) *Both features congruent*: the sound induces a colour and shape that both match the display features. (b) *Colour congruent, shape incongruent*: the synaesthetic colour matches image colour, whereas the synaesthetic shape differs from the image shape. (c) *Shape congruent, colour incongruent*: the synaesthetic shape matches the image shape, whereas the synaesthetic colour differs from the image colour. (d) *Both features incongruent*: the synaesthetic colour and shape both differ from the image features. In Experiment 1, the on-screen location of the stimulus was always consistent with the synaesthetic location.

2.3.5. Experiment 2

The aim of Experiment 2 was to examine the impact of spatial location in synaesthetic experience. We tested this by manipulating the on-screen position of targets. The spatial congruency was defined by where the target was positioned on the computer screen in relation to where synaesthetes positioned their drawing on the screen or paper. For each synaesthete, we used the same set of four sound–image pairs as those in Experiment 1 such that the images were manifestly distinct from each other in colour, shape, and location. The design, procedure, and instructions of Experiment 2 were identical to Experiment 1, with the exception that we manipulated the on-screen position of targets, while keeping one of the other visual features constant. In the *colour* task, the image colour and on-screen location were either congruent or incongruent with the synaesthetic colour and location while the synaesthetic shape induced by the sound was always consistent with the image shape. Conversely, in the *shape* task, shape and location were independently manipulated while synaesthetic colour was always consistent with image colour. As a result, two different versions of the stimuli were used in the colour and shape tasks. There were four conditions for each task: (1) both features congruent; (2) location incongruent; (3) colour or shape incongruent (in the colour / shape task, respectively); and (4) both features incongruent.

3. Results

3.1. Subjective session

Although the reported experiences initially seem idiosyncratic and variable across synaesthetes, there is a systematic relationship between auditory pitch and visual features: in all seven synaesthetes, high-pitched sounds induce visual experiences that are brighter in colour, smaller in size, and higher in space, relative to low-pitched sounds. Fig. 3 illustrates the pattern of the synaesthetic experiences from two representative participants. Such a pattern bears similarity to previous research on the way non-synaesthetes map auditory pitch to visual features (Spence, 2011), and is also consistent with Ward et al. (2006) who reported similarities between synaesthetes and non-synaesthetes in auditory–visual mappings.

To quantify the phenomenological relationship between auditory pitch and the size, brightness, and location of synaesthetic objects, we performed correlation analyses: for each of the seven synaesthetes, we calculated the size (number of pixels) of the synaesthetic object and brightness of the selected colour (in Hue-Saturation-Brightness colour coordinates, ranging from 0 to 100) using Photoshop (hand-drawings were scanned and converted into JPG files). If multiple colours were present in an image, we used the colour that occupied the most area. As some chose to draw their experiences using computer graphics and others did it on papers with pastels, the different spatial frames in the two situations led to difficulties in extracting precise variations in space. In addition, most synaesthetes expressed difficulty in precisely locating the synaesthetic object in space or transferring its location onto a two-dimensional (2D) image (often they provided

generic descriptions like ‘it is low down’ or ‘it is in the middle’). Therefore, we categorised their descriptions about the spatial components of synaesthetic experiences into three main types (low, middle, and high) and coded them as an ordinal variable. After obtaining the data of number of pixels, brightness values, and location codings for each person, the results were averaged across three instruments, giving us 20 data-points (10 notes \times two repetitions) per synaesthete. The data were then averaged across synaesthetes and submitted to correlation analyses, relating auditory pitch (in Hz) to size, brightness, and spatial location.

The results of the correlations are consistent with the apparent patterns from looking at the images: as Fig. 4a illustrates, the size of synaesthetic objects decreases when auditory pitch gets higher, as indexed by a significant negative correlation (Pearson’s $r = -.79$, $p < .001$). Fig. 4b shows a significant positive correlation that the brightness of synaesthetic colour gradually becomes greater as auditory pitch gets higher (Pearson’s $r = .76$, $p < .001$). Finally, Fig. 4c shows that the location of synaesthetic objects elevates as pitch gets higher (Kandall’s $\tau = .84$, $p < .001$).

In the questionnaire probing the subjective locus of synaesthetic experience, one of the seven synaesthetes indicated that her synaesthetic percepts appeared out in space. This individual also described seeing objects she was voluntarily imagining as ‘out in space’, rather than ‘in mind’s eye’. The other six synaesthetes reported seeing their synaesthetic objects in the mind’s eye. One of these six people reported seeing imagined objects ‘out in space’, another reported them as both in space and in mind’s eye, and the rest described imagined objects as appearing only in mind’s eye. Interestingly, although the six individuals chose ‘in the mind’s eye’ over ‘out in space’ for auditorily-induced synaesthetic images in the binary question, some of their descriptions raise questions about the appropriateness of the categorisation of ‘in the mind’s eye’ versus ‘out in space’. For example, one synaesthete added a description about his grapheme–colour synaesthesia suggesting it may be experienced in external space: ‘When I read texts, it’s projected over the letter or sort of floating just above the text.’, and two synaesthetes described their sound-induced synaesthetic images as ‘it’s like something in front of me’ and ‘it’s in my mind’s eye but with a strong spatial sense’. This implies that their synaesthetic percepts may not entirely be situated only in mind’s eye, and illustrate the difficulty in describing such an experience spatially. Taken together, these subjective reports hint that, although the vividness of synaesthetic percepts certainly varies among individuals, the dichotomy of ‘in mind’s eye’ versus ‘out in space’ may be confounded by the way in which synaesthetes choose to describe their experiences. For example, a person reporting seeing grapheme-induced synaesthetic colour appearing on the page may describe his sound-induced images in mind’s eye because there is no external visual stimulus for it to be ‘pinned’ onto spatially, leading to contradictory categorisations. Given the difficulty in describing the spatial location of an internally generated experience, subjective reports may be affected by how the questions are framed and how the options are interpreted. (For related discussion in grapheme–colour synaesthesia, often referred to as ‘associator vs projector’ distinction, see

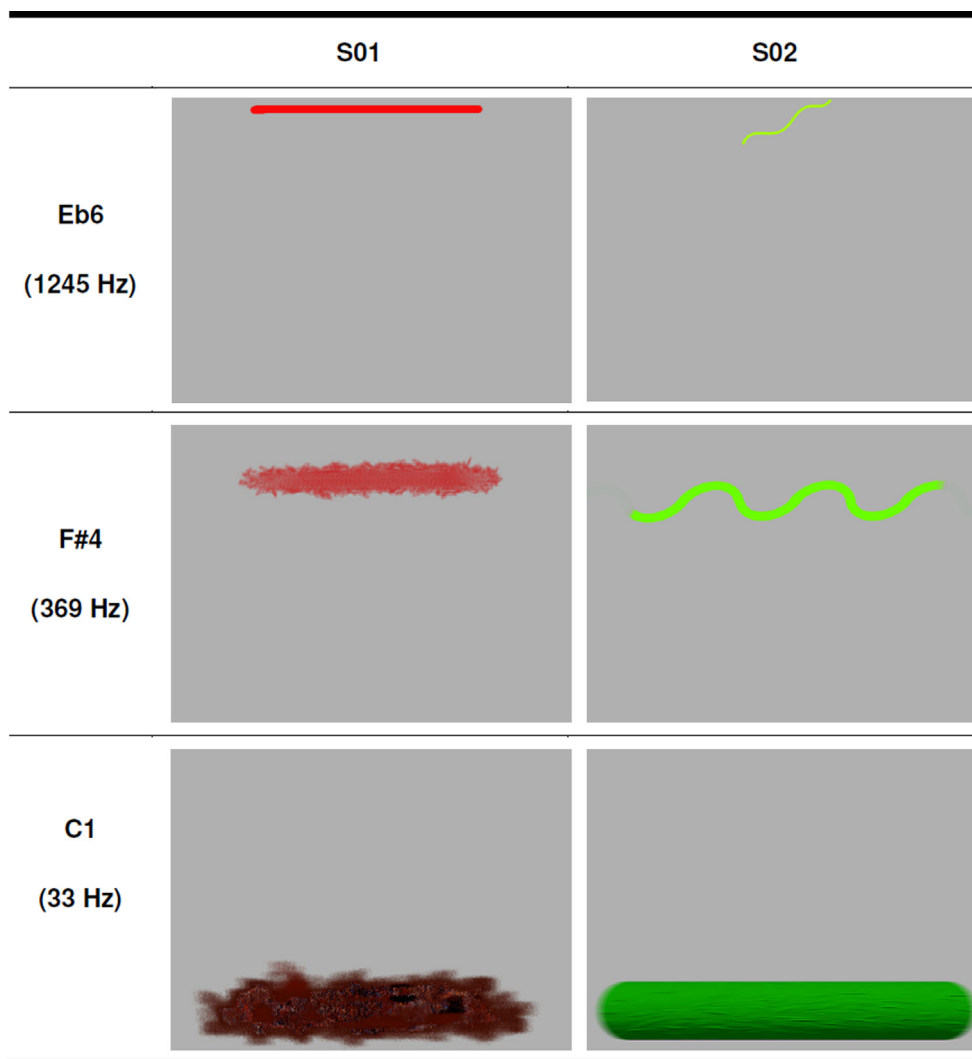


Fig. 3 – Examples of the synaesthetic percepts induced by the sounds of violin C1, F#4, and Eb6 for synaesthetes S01 and S02. Note that both show an obvious trend of the cross-modal mapping between auditory pitch and visual features (brightness, size, and spatial height).

Dixon et al., 2004; Edquist et al., 2006; Ward et al., 2007; Karstoft and Rich, submitted for publication).

3.2. Experiment 1

For all participants, erroneous responses (2.5%) and outliers (defined as responses < 100 msec and > 3000 msec; .1%) were excluded from further analyses. Fig. 5a shows the mean correct RT and repeated-measures standard error (SE) of each condition for synaesthetes and controls. Table 1 shows the mean error rate of each condition. We analysed correct RTs and error rates using a mixed design analysis of variance (ANOVA) with a between-subject factor of group (*synaesthetes vs controls*), and within-subject factors of task (*colour vs shape*) and congruency (*both features congruent, shape incongruent, colour incongruent, and both features incongruent*). In all statistics reported in the present study, we used the Greenhouse–Geisser adjustment to adjust violations of sphericity where necessary, and the Bonferroni correction to control for family-wise error rates in all post-hoc multiple comparisons.

The results of the ANOVA show no significant main effect of group [$F < 1.0$, *n.s.*] and significant main effects of task [$F(1, 12) = 9.02$, $p = .01$, $\eta^2 = .42$] and congruency [$F(1.93, 23.22) = 6.65$, $p = .006$, $\eta^2 = .35$]. These main effects are modified by a significant task \times congruency interaction [$F(1.66, 19.93) = 4.49$, $p = .03$, $\eta^2 = .27$], as well as a significant group \times congruency interaction [$F(3, 36) = 5.52$, $p = .003$, $\eta^2 = .31$; see Fig. 5b]. The three-way interaction of group \times task \times congruency is not significant [$F(1.66, 19.93) = 1.19$, $p = .31$].

Based on the significant group \times congruency interaction, we conducted post-hoc pair-wise comparisons (Bonferroni corrected α -level: $.05/6 = .008$, with .05 being the conventional α -level of statistical tests and six being the number of pair-wise comparisons) to explore how the congruency effect affected the two groups differently. This interaction is illustrated in Fig. 5b, where the results are collapsed across task. The analyses show that synaesthetes are significantly slower in the *shape incongruent, colour incongruent, and both features incongruent* conditions than in the *both features congruent* condition (all $ps < .004$; see Fig. 5b), but the former three incongruent

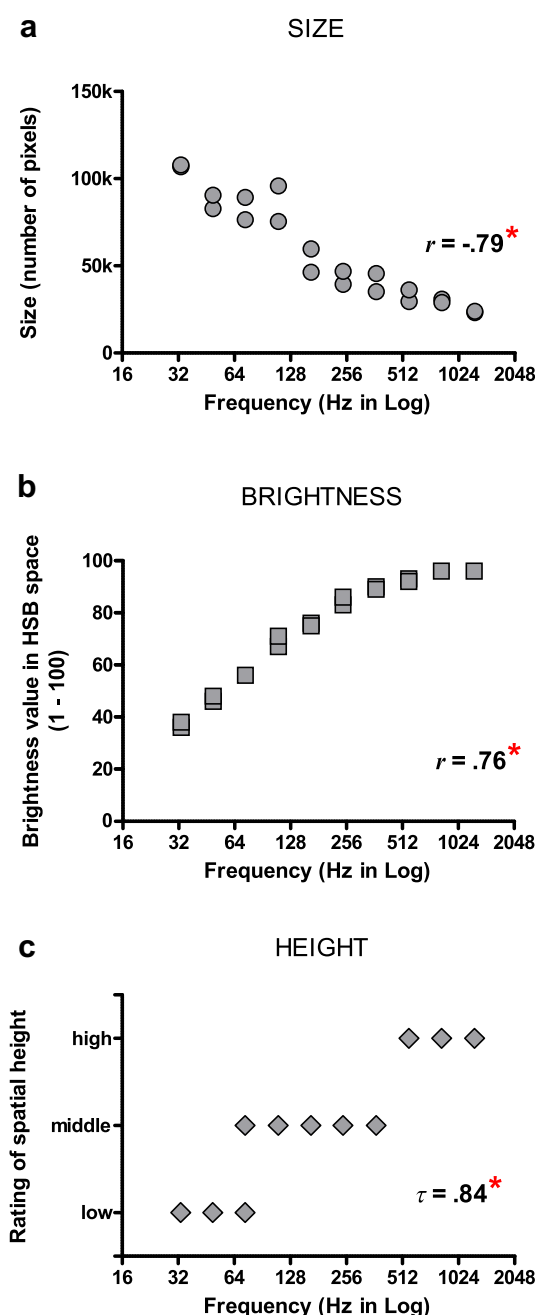


Fig. 4 – (a) The relationship between pitch and the size of synaesthetic objects as measured by the number of pixels in the drawings of synaesthetes. (b) The relationship between pitch and the brightness of synaesthetic objects indicated by brightness value (1–100) of their chosen colour. (c) The relationship between auditory and the location of synaesthetic objects indicated by categorical codings of synaesthetes’ description. Asterisks indicate a statistically significant correlation ($p < .05$). Note that there are 20 data-points (2 repetitions by 10 notes) in each figure, but some data-points are concealed due to auditory pitch (Hz) being compressed in logarithmic scale.

conditions do not differ from each other (all $ps > .12$). By contrast, there are no significant effects for controls (all $ps > .32$; see Fig. 5b). The exact p -values of all post-hoc comparisons for this critical interaction are reported in [Supplementary Materials](#).

The significant task \times congruency interaction in the omnibus ANOVA indicates that the congruency effect is modulated by task-related attentional set: synaesthetic congruency affected performance differently when participants attended to the colour versus shape dimensions in the two tasks. Post-hoc comparisons revealed the source of the two-way interaction: in the colour task, the *both features congruent* condition is marginally different from the *shape incongruent* condition ($p = .009$) and significantly different from the *colour incongruent* condition ($p < .0001$). The two partially incongruent conditions also significantly differ from each other ($p = .008$). In the shape task, however, there are no significant differences among the conditions (all $ps > .05$, except 3 contrasts: *both features congruent vs shape incongruent* and *colour incongruent vs both features incongruent*, both $ps = .03$; *shape incongruent vs colour incongruent*, $p = .02$; note these are not significant after correction for multiple comparisons). Notice that, in this task \times congruency interaction, data are collapsed across synaesthetes and controls, which implies that controls show a similar pattern to that of synaesthetes (albeit numerically much less evident, see Fig. 5a). Nonetheless, this pattern needs to be interpreted with caution, because the significant group \times congruency interaction and subsequent analyses indicated that only synaesthetes, not controls, were affected by synaesthetic congruency. Unfortunately we lack the statistical power to pull out the three-way interaction (which would show that task-related attentional set modulates the effects of synaesthetic colour and shape differently in synaesthetes and in controls), due to the difficulty in recruiting individuals with this relatively rare form of synaesthesia. If we look at the pattern for the partially incongruent conditions in Fig. 5a, it appears that for synaesthetes, in the colour task, the impact of incongruent colours is greater than incongruent shapes [compare the two grey bars in Fig. 5a - COLOUR] whereas the two conditions with identical stimuli show an inverse pattern in the shape task, such that incongruent shapes appear to interfere more than incongruent colours [the two grey bars in Fig. 5a - SHAPE]. This pattern fits our *a priori* hypothesis that a task-relevant feature should have a stronger impact than a task-irrelevant one despite them being integrated to form an object-like percept, albeit not strong enough to come out in a three-way interaction with our sample size. Thus, the comparisons for the partially incongruent conditions are crucial as they allow us to evaluate the effect of each feature under different attentional sets, but with identical stimuli. If we do planned comparisons on these data, the difference between the two partially incongruent conditions is significant in the colour task [$t(6) = -3.32$, $p = .01$; *colour incongruent > shape incongruent*], and a trend in the shape task [$t(6) = 2.04$, $p = .08$; *shape incongruent > colour incongruent*²], with this pattern also evident in all synaesthetes individually.

² Note the variance for the shape incongruent condition is $\sim 2\times$ larger than the other partially incongruent conditions for both tasks; hence the statistics reveal a trend whereas the graph implies a stronger effect.

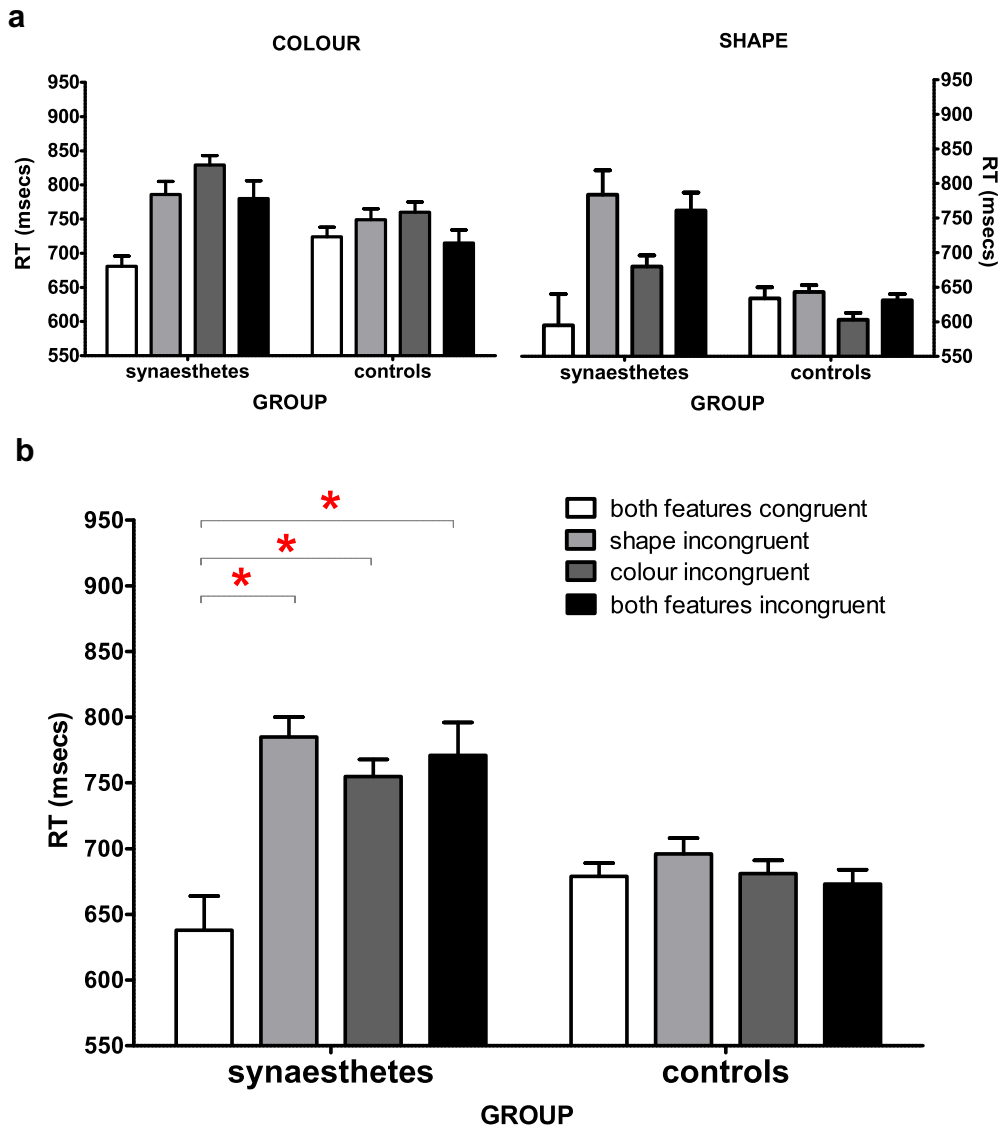


Fig. 5 – (a) Mean reaction times (+1 repeated-measures SE) for all conditions in Experiment 1, plotted as a function of group, task, and congruency. (b) Mean reaction times (+1 repeated-measures SE) for synaesthetes and controls in Experiment 1, plotted as a function of group and congruency, collapsed across task, to illustrate the interaction. Asterisks indicate a statistically significant difference (corrected for multiple comparisons).

The identical analysis on control data from these conditions show no reliable difference in the colour task [$t(6) = -.97$, $p = .36$] and a reliable difference in the shape task [$t(6) = 2.39$, $p = .05$; *shape incongruent* > *colour incongruent*]. In [Supplementary Materials](#), we report an alternative exploratory analysis, which treats each feature as an individual congruency factor, to test how task-related attentional set modulates the respective impact of synaesthetic colour and shape. The results are consistent with the planned comparisons, such that, for synaesthetes only, the impact of synaesthetic colour is more powerful in the colour than in the shape task and, conversely, the impact of synaesthetic shape is stronger in the shape than in the colour task.

The same analyses on the error rate of each condition reveal a significant main effect of congruency [$F(2, 24) = 4.15$, $p = .02$, $\eta^2 = .25$], with no post-hoc tests being significant

(all $ps > .10$). No other statistics reached significance (all $ps > .12$).

3.3. Experiment 2

Errors (2.5%) and outliers (.2%) were excluded from further analyses. [Fig. 6](#) shows the mean correct RT and repeated-measures SE of each condition for synaesthetes and controls. The mean error rate of each condition is reported in [Table 2](#). Note that in Experiment 2 we used different image sets in the colour and shape task to control for the effects of the third feature (shape or colour in different tasks). The displayed shape was always congruent with the synaesthetic shape in the colour task and vice versa for the colour in the shape task, while the other feature and location were manipulated. Therefore, we conducted separate analyses for

Table 1 – The mean error rates (%) of each condition in Experiment 1.

Synaesthetes								Controls							
Colour task				Shape task				Colour task				Shape task			
CC–SC	CC–SI	CI–SC	CI–SI	CC–SC	CC–SI	CI–SC	CI–SI	CC–SC	CC–SI	CI–SC	CI–SI	CC–SC	CC–SI	CI–SC	CI–SI
0	5.06	4.46	1.49	0	4.46	2.68	2.08	0	3.87	4.46	2.38	2.38	1.49	1.49	1.79

Abbreviations: CC–SC (colour congruent, shape congruent); CC–SI (colour congruent, shape incongruent); CI–SC (colour incongruent, shape congruent); CI–SI (colour incongruent, shape incongruent)

the colour and shape tasks. All other aspects of the analyses matched Experiment 1.

For the colour task, we carried out a mixed design ANOVA with a between-participant factor of group (*synaesthetes* vs *controls*) and a within-participant factor of congruency (*both features congruent*, *location incongruent*, *colour incongruent*, and *both features incongruent*). Consistent with the pattern we found in Experiment 1, *synaesthetes* showed effects of *synaesthetic congruency* that were not present in *controls*. The ANOVA revealed no significant main effect of group ($F < 1.0$, *n.s.*), a significant main effect of congruency [$F(1.57, 18.92) = 10.10$,

$p = .002$, $\eta^2 = .45$], and a significant group \times congruency interaction [$F(3, 36) = 5.47$, $p = .003$, $\eta^2 = .31$; see Fig. 6a]. Post-hoc tests (the Bonferroni corrected α -level: .008) showed that, in *synaesthetes*, RTs were slower in the *location incongruent*, *colour incongruent*, and *both features incongruent* conditions than the *both features congruent* condition (all $ps < .002$, except the difference between *location incongruent* and *both features congruent*, which was a strong trend, $p = .01$, not significant after correction for multiple comparisons). In addition, the three incongruent conditions did not differ from one another (all $ps > .06$), except for the *both incongruent* condition being significantly slower than the *location incongruent* condition ($p < .0001$). By contrast, *controls* showed no effect of congruency (all $ps > .07$). The exact p -values of all post-hoc comparisons for this critical interaction are reported in [Supplementary Materials](#).

For the shape task, we conducted the identical analysis with a between-participant factor of group (*synaesthetes* vs *controls*) and a within-participant factor of congruency (*both features congruent*, *location incongruent*, *shape incongruent*, and *both features incongruent*). The results revealed no significant main effect of group ($F < 1.0$, *n.s.*), a significant main effect of congruency [$F(1.28, 15.44) = 4.47$, $p = .04$, $\eta^2 = .27$], and a significant group \times congruency interaction [$F(3, 36) = 3.95$, $p = .01$, $\eta^2 = .24$; see Fig. 6b]. Post-hoc comparisons (Bonferroni corrected α -level: .008) showed that *synaesthetes* were significantly slower in the *location incongruent*, *shape incongruent*, and *both features incongruent* conditions than the *both features congruent* condition (all $ps \leq .008$). No other comparisons in the *synaesthete* group achieved significance (all $ps > .05$; except for *location incongruent* vs *shape incongruent*, $p = .03$, not significant after correction for multiple comparisons). Consistent with the colour task, *controls* show no effect of congruency (all $ps > .4$, except *both congruent* vs *location incongruent*, $p = .048$, not significant after correction for multiple comparisons). The exact p -values are reported in [Supplementary Materials](#).

The same analyses on the error rate reveal, in the colour task, a significant main effect of congruency [$F(2.13, 25.67) = 4.21$, $p = .02$, $\eta^2 = .26$]. Post-hoc tests show that error rate is significantly higher in the *location incongruent* condition (1.48%, $p = .01$) and marginally higher in the *both features incongruent* condition (3.42%, $p = .08$) than in the *both features congruent* condition (0%). In the shape task, there were no significant effects (all $ps > .18$).

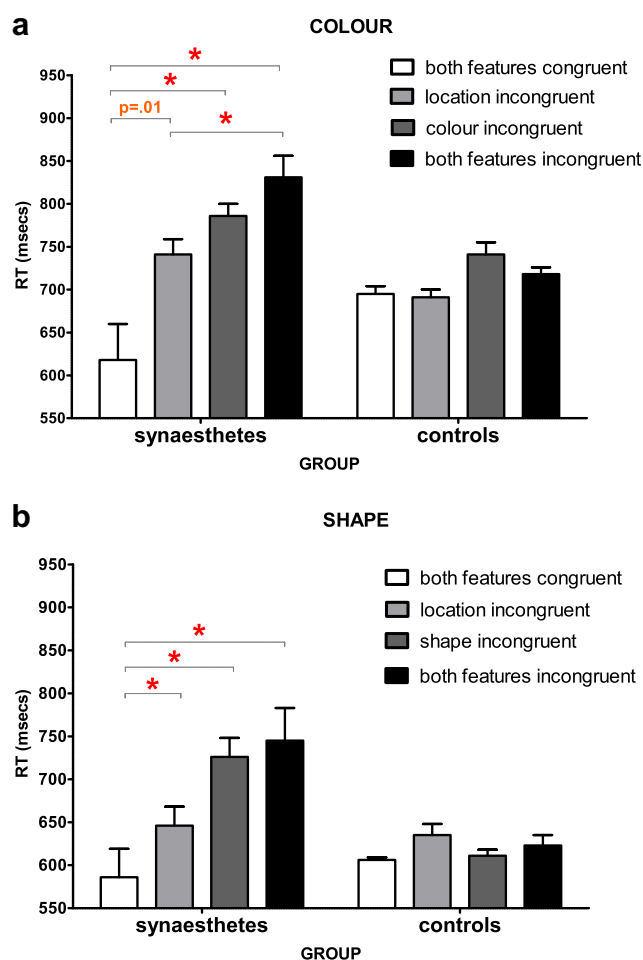


Fig. 6 – Mean reaction times (+1 repeated-measures SE) for the colour (a) and shape (b) discrimination tasks in Experiment 2, plotted as a function of group and congruency. Asterisks indicate a statistically significant difference (corrected for multiple comparisons).

4. Discussion

Auditory–visual *synaesthesia*, an unusual phenomenon in which sounds elicit visual experiences, is often mentioned

Table 2 – The mean error rates (%) of each condition in Experiment 2.

Synaesthetes								Controls							
Colour task				Shape task				Colour task				Shape task			
CC–LC	CC–LI	CI–LC	CI–LI	SC–LC	SC–LI	SI–LC	SI–LI	CC–LC	CC–LI	CI–LC	CI–LI	SC–LC	SC–LI	SI–LC	SI–LI
0	.30	2.68	3.27	.60	.89	1.49	2.68	0	2.68	2.38	3.57	1.49	2.08	3.87	3.27

Abbreviations: CC–LC (colour congruent, location congruent); CC–LI (colour congruent, location incongruent); CI–LC (colour incongruent, location congruent); CI–LI (colour incongruent, location incongruent); SC–LC (shape congruent, location congruent); SC–LI (shape congruent, location incongruent); SI–LC (shape incongruent, location congruent); SI–LI (shape incongruent, location incongruent).

anecdotally in scientific literature but has rarely been studied experimentally. The few studies that use objective measures focus on the reported colour experience (e.g., Goller et al., 2009; Ward et al., 2006). In the present study, we studied seven synaesthetes with consistent visual experiences of coloured geometric objects in space when listening to sounds. These synaesthetic object features were specific and consistent over time, as evident in specificity and consistency tests, although synaesthetes often expressed difficulty in precisely locating the object-like percepts in space. Synaesthetes' drawings in response to sounds showed systematic trends between auditory pitch and synaesthetic experience, which follow the same rules as the implicit cross-modal mappings in non-synaesthetes. These patterns show up as significant correlations between increasing pitch and increase in brightness, reduction in size, and elevation in spatial location. The experimental results show that the visual experience of coloured shapes in specific spatial locations affects the behavioural performance of synaesthetes on both colour and shape judgements, despite it being irrelevant to the task.³ This is consistent with previous reports on other forms of synaesthesia that synaesthetes are unable to effectively suppress their unusual experiences once they perceive the inducing stimuli (e.g., grapheme–colour synaesthesia: Mattingley et al., 2001; sound–colour synaesthesia: Ward et al., 2006). Although it was not as strong as these overall effects, we also observed modulations by feature-based attention. Specifically, in Experiment 1, when synaesthetes attended to colour, a mismatch between the displayed colour and the synaesthetic colour caused a stronger congruency effect than a mismatch of shape, and *vice versa* when they attended to shape. Although this effect was not strong enough to survive the three-way interaction, it was evident in both planned comparisons (based on our *a priori* prediction) and in the alternative exploratory analyses (see [Supplementary Materials](#)). These results suggest that *after* synaesthetic percepts of coloured objects are elicited, feature-based attention acts on these objects to select and prioritise

relevant features, which, in turn, modulates their behavioural impact.

These congruency effects suggest both colour and non-colour features can be integral components of the unusual experience and should be considered in theories for synaesthesia. In addition, we need further studies to examine the mechanisms that underlie these phenomena. The perceptual characteristics and neural underpinnings of synaesthetic colour have been extensively studied, which point the way for future research on non-colour synaesthetic features. At the psychophysical level, the majority of evidence suggests that synaesthetic colour does not 'behave' like real colour (e.g., it shows no chromatic adaptation: Hong and Blake, 2008; it shows no pre-attentive pop-out: Ward et al., 2010; Edquist et al., 2006; Sagiv et al., 2006; Nijboer et al., 2011; Karstoft and Rich (submitted for publication), although see Ramachandran and Hubbard (2001), as well as Kim and Blake (2005), for synaesthetic colour showing properties like real colour). This is consistent with the idea that synaesthetic colour experiences arise at a late stage in the hierarchy of visual processing.

At the neural level, whether synaesthetic colour activates colour-selective area V4 has sparked heated discussion among researchers: some studies observed V4 activation induced by achromatic letters in the brains of synaesthetes (e.g., Hubbard et al., 2005; Nunn et al., 2002; Sperling et al., 2006) whereas other studies found no activation in V4 or only in areas related to colour knowledge (Hupe et al., 2011; Rich et al., 2006). In addition, Rich et al. (2006) found that voluntary colour imagery (but not synaesthetic colour) in both synaesthetes and controls activated regions around V4. Using the repetition suppression paradigm of functional magnetic resonance imaging (fMRI), which detects reduction in neural activity if repeated stimuli are represented in overlapping brain areas, a recent study found that synaesthetic colour failed to suppress the activity induced by real colour in V4, leading to the conclusion that synaesthetic colour is mediated by higher-order areas of the visual hierarchy and does not fully share neural substrates with real colour (van Leeuwen et al., 2010). These conflicting results might be due to methodological differences or limited statistical power, as suggested by a recent review (Rouw et al., 2011), or indeed over liberal criteria (Hupe et al., 2011). However, it would be premature to state at this stage that the colour-selective areas (e.g., V4) are equally involved in synaesthetic and real colour, despite them seeming phenomenally similar in subjective reports (although note that synaesthetes can clearly distinguish between their synaesthetic experiences and 'real' colours).

³ Synaesthetic experiences occur involuntarily in the sense that there is no benefit for synaesthetes to 'activate' such experiences when it is actually detrimental to task performance. However, this does not necessarily mean that synaesthetic experiences are automatic. Typical criteria for a process to be considered automatic include freedom from dual-task interference and requiring little or no attention (Moors and De Houwer, 2006), neither of which is true of synaesthesia (Mattingley et al., 2006).

In a similar vein, although the psychophysical properties and neural correlates of non-colour synaesthetic features remain to be explored, we should perhaps not assume that the shape- and location-selective areas of the visual system (e.g., lateral-occipital cortex: Kourtzi and Kanwisher, 2001) are the only regions potentially involved in such multi-feature phenomena. In addition to these brain areas specially tuned for visual features, we must look also at brain areas that lie beyond the visual cortex, such as those involved in shape/object knowledge (e.g., middle temporal and inferior frontal gyri: Pulvermüller and Hauk, 2006). We can also explore the similarities between synaesthetic form and real shapes psychophysically to see if synaesthetic shape shows similar psychophysical properties to real shape, much as comparing synaesthetic and real colour has been used to explore whether this experience involves early or late mechanisms of the visual system. For instance, shape perception is susceptible to illusions (e.g., a physically straight line can appear perceptually curved in certain surroundings: Todd, 2004), but it is unknown whether synaesthetic shapes would be affected by illusion-inducing contexts.

Relevant to the present findings, a few recent proposals suggest that brain areas not directly involved in the representation of colour may play critical roles in synaesthesia. For instance, the inferior temporal gyrus is suggested to represent the contour of spatial sequence synaesthesia, in which over-learned sequences (e.g., alphabet or numbers) are configured spatially with reliable form in the person's mind's eye (Eagleman, 2009). This phenomenon may share neural underpinnings with the spatial representation attached to the synaesthetic objects reported here. In addition, the right parietal lobule may be important in the attentional integration of different synaesthetic features, akin to the way visual features of real objects are bound (Esterman et al., 2006; Hubbard, 2007; Robertson, 2003).

The major theories for the neural bases of synaesthesia involving colour percepts (e.g., the cross-activation and disinhibited views) need to expand to incorporate a broader neural network, beyond V4. For instance, higher-order brain areas involved in the knowledge of the canonical colour and shape of objects might be possible candidate regions that represent the experience of synaesthetic objects. Additionally, previous studies have suggested that recognition of the meaning of letters/numbers plays a crucial role in grapheme–colour synaesthesia (Dixon et al., 2006). As our synaesthetes can readily recognise the instruments by their timbre and different instruments induce apparently distinct colours and shapes, brain areas involved in representing meaning (e.g., anterior temporal lobe: Pobric et al., 2007) might also play a role in this cross-modal phenomenon.

The modulatory effect of voluntary attention over synaesthetic features is consistent with previous studies demonstrating the effects of voluntary attention on grapheme–colour synaesthesia (Mattingley et al., 2006; Rich and Mattingley, 2003, 2010; Sagiv et al., 2006). These studies show that diverting attention from graphemes can reduce or eliminate the congruency effects of synaesthetic colour. Essentially, attending to the grapheme serves as a prerequisite for synaesthetic colour to be elicited, although once the

inducing stimulus is attended and recognised, the subsequent processes that elicit synaesthetic percepts seem to be relatively involuntary (for related debates about the role of attention in synaesthesia, see Edquist et al., 2006; Hubbard et al., 2005; Nijboer et al., 2011; Ramachandran and Hubbard, 2001; Ward et al., 2010). Our findings further reveal how attention modulates the perceptual representation of synaesthetic objects: first, the congruency effect caused by unattended feature (e.g., a mismatching shape when colour is attended) fits with the idea that once an object is selected, all its constituent features are processed to an extent, regardless of their relevance to the current task (Blaser et al., 2000). Second, the attended feature seems to cause larger interference than the unattended feature, suggesting that, although a synaesthetic object is generated internally and consists of tightly integrated features, the feature-based attentional set may effectively select a component feature of the internal image. This could be analogous to the effects holding an item in working memory has in guiding attention to matching features (for review, see Soto et al., 2008). Thus, setting voluntary attention to the task-relevant feature also selects the same feature in an image that is internally created in the absence of incoming visual signals, analogous to its effect on 'normal' perception when multiple features physically appear in a visual scene (Saenz et al., 2003).

Our results also show that the relationship between pitch and synaesthetic objects follow the same rules as the subtle cross-modal mappings seen in non-synaesthetes: non-synaesthetic individuals tend to map high-pitched sounds with small, bright objects located high in space. This effect in non-synaesthetes has been documented using subjective report (Eitan and Timmers, 2010; Ward et al., 2006), speeded reaction time (Ben-Artzi and Marks, 1995; Evans and Treisman, 2010; Marks, 1987), and preferential looking in infants (Walker et al., 2010). Although the implicit cross-modal correspondences in non-synaesthetes can only be measured under specific experimental settings, whereas synaesthetes have daily conscious experiences of auditorily-induced visual percepts, there are some hints in the data that controls may be subtly affected by these mappings even when we use stimuli tailored to synaesthete experiences. For example, as Fig. 5a illustrates, controls show a pattern numerically similar to that of synaesthetes across conditions, although there are no statistically significant congruency effects in their data.

Ward et al. (2006) suggest that similarities between synaesthetes and non-synaesthetes in sound–colour mappings show that synaesthesia co-opts the neural substrates for 'normal' cross-modality mappings and reveals the associations in a more explicit form. Another study reporting the similarity between synaesthetes and non-synaesthetes in their mapping between luminance and numerical quantity also fits the notion that synaesthesia builds on 'normal' mechanisms of non-synaesthetic brain (Cohen Kadosh et al., 2007). We interpret our data similarly as implying a common neural/cognitive mechanism underlying both auditory–visual synaesthesia and 'normal' cross-modal mappings.

The documentation of non-colour synaesthetic visual features is crucial for developing more comprehensive models

to explain how synaesthesia relates to general aspects of cognition. Here we provide objective evidence showing that auditorily-induced synaesthetic objects with multiple features affect behaviour, as well as that attention modulates the component features of synaesthetic objects. Our findings suggest overt synaesthetic experiences induced by sounds reflect implicit cross-modal mechanisms we all share. More broadly, they demonstrate the fundamental importance that intersensory integration and voluntary attention have on conscious experience.

Acknowledgements

We thank Alex Holcombe for helpful comments, and Bojan Neskovic for help with stimuli. RC is supported by a Macquarie University Research Excellence Scholarship & the Education Ministry of Taiwanese Government. ANR is supported by the Australian Research Council (DP0984494).

Supplementary material

Supplementary material related to this article can be found online at [doi:10.1016/j.cortex.2012.04.006](https://doi.org/10.1016/j.cortex.2012.04.006).

REFERENCES

- Banissy MJ and Ward J. Mirror-touch synesthesia is linked with empathy. *Nature Neuroscience*, 10(7): 815–816, 2007.
- Baron-Cohen S, Burt L, Smith-Laittan F, Harrison J, and Bolton P. Synaesthesia: Prevalence and familiarity. *Perception*, 25(9): 1073–1079, 1996.
- Beeli G, Esslen M, and Jancke L. Synaesthesia: When coloured sounds taste sweet. *Nature*, 434(7029): 38, 2005.
- Ben-Artzi E and Marks LE. Visual–auditory interaction in speeded classification: Role of stimulus difference. *Perception and Psychophysics*, 57(8): 1151–1162, 1995.
- Blaser E, Pylyshyn ZW, and Holcombe AO. Tracking an object through feature space. *Nature*, 408(6809): 196–199, 2000.
- Brainard DH. The psychophysics toolbox. *Spatial Vision*, 10(4): 433–436, 1997.
- Cohen Kadosh R, Henik A, and Walsh V. Small is bright and big is dark in synaesthesia. *Current Biology*, 17(19): R834–R835, 2007.
- Cohen Kadosh R, Henik A, Catena A, Walsh V, and Fuentes LJ. Induced cross-modal synaesthetic experience without abnormal neuronal connections. *Psychological Science*, 20(2): 258–265, 2009.
- Dixon MJ, Smilek D, and Merikle PM. Not all synaesthetes are created equal: Projector versus associator synaesthetes. *Cognitive, Affective, & Behavioural Neuroscience*, 4(3): 335–343, 2004.
- Dixon MJ, Smilek D, Duffy PL, Zanna MP, and Merikle PM. The role of meaning in grapheme–colour synaesthesia. *Cortex*, 42(2): 243–252, 2006.
- Eagleman DM. The objectification of overlearned sequences: A new view of spatial sequence synesthesia. *Cortex*, 45(10): 1266–1277, 2009.
- Eagleman DM and Goodale MA. Why color synesthesia involves more than color. *Trends in Cognitive Science*, 13(7): 288–292, 2009.
- Edquist J, Rich AN, Brinkman C, and Mattingley JB. Do synaesthetic colours act as unique features in visual search? *Cortex*, 42(2): 222–231, 2006.
- Eitan Z and Timmers R. Beethoven’s last piano sonata and those who follow crocodiles: Cross-domain mappings of auditory pitch in a musical context. *Cognition*, 114(3): 405–422, 2010.
- Esterman M, Verstynen T, Ivry RB, and Robertson LC. Coming unbound: Disrupting automatic integration of synesthetic color and graphemes by transcranial magnetic stimulation of the right parietal lobe. *Journal of Cognitive Neuroscience*, 18(9): 1570–1576, 2006.
- Evans KK and Treisman A. Natural cross-modal mappings between visual and auditory features. *Journal of Vision*, 10(1): 1–12, 2010.
- Fitzgibbon BM, Enticott PG, Rich AN, Giummarra MJ, Georgiou-Karistianis N, Tsao JW, et al. High incidence of ‘synaesthesia for pain’ in amputees. *Neuropsychologia*, 48(12): 3675–3678, 2011.
- Goller AI, Otten LJ, and Ward J. Seeing sounds and hearing colors: An event-related potential study of auditory–visual synesthesia. *Journal of Cognitive Neuroscience*, 21(10): 1869–1881, 2009.
- Grossenbacher PG and Lovelace CT. Mechanisms of synesthesia: Cognitive and physiological constraints. *Trends in Cognitive Science*, 5(1): 36–41, 2001.
- Hong SW and Blake R. Early visual mechanisms do not contribute to synesthetic color experience. *Vision Research*, 48(8): 1018–1026, 2008.
- Hubbard EM, Arman AC, Ramachandran VS, and Boynton GM. Individual differences among grapheme–color synesthetes: Brain–behavior correlations. *Neuron*, 45(6): 975–985, 2005.
- Hubbard EM and Ramachandran VS. Neurocognitive mechanisms of synesthesia. *Neuron*, 48(3): 509–520, 2005.
- Hubbard EM. A real red-letter day. *Nature Neuroscience*, 10(6): 671–672, 2007.
- Hubbard EM, Brang D, and Ramachandran VS. The cross-activation theory at 10. *Journal of Neuropsychology*, 5(2): 152–177, 2011.
- Hupe JM, Bordier C, and Dojat M. The neural bases of grapheme–color synesthesia are not localized in real color-sensitive areas. *Cerebral Cortex*, 2011.
- Karstoft K-I and Rich AN. Exploring the benefit of synaesthetic colours: Individual variability in grapheme–colour synaesthesia. submitted for publication.
- Kourtzi Z and Kanwisher N. Representation of perceived object shape by the human lateral occipital complex. *Science*, 293(5534): 1506–1509, 2001.
- Kim CY and Blake R. Watercolor illusion induced by synesthetic colors. *Perception*, 34(12): 1501–1507, 2005.
- Marks LE. On cross-modal similarity: Auditory–visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception and Performance*, 13(3): 384–394, 1987.
- Mattingley JB, Rich AN, Yelland G, and Bradshaw JL. Unconscious priming eliminates automatic binding of colour and alphanumeric form in synaesthesia. *Nature*, 410(6828): 580–582, 2001.
- Mattingley JB, Payne JM, and Rich AN. Attentional load attenuates synaesthetic priming effects in grapheme–colour synaesthesia. *Cortex*, 42(2): 213–221, 2006.
- Moors A and De Houwer J. Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, 132(2): 297–326, 2006.
- Nijboer TC, Satris G, and Stigchel SV. The influence of synesthesia on eye movements: No synesthetic pop-out in an oculomotor target selection task. *Consciousness and Cognition*, 20(4): 1193–1200, 2011.

- Novich S, Cheng S, and Eagleman DM. Is synaesthesia one condition or many? A large-scale analysis reveals subgroups. *Journal of Neuropsychology*, 5(2): 353–371, 2011.
- Nunn JA, Gregory LJ, Brammer M, Williams SC, Parslow DM, Morgan MJ, et al. Functional magnetic resonance imaging of synesthesia: Activation of V4/V8 by spoken words. *Nature Neuroscience*, 5(4): 371–375, 2002.
- Olivers CN, Meijer F, and Theeuwes J. Feature-based memory-driven attentional capture: Visual working memory content affects visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, 32(5): 1243–1265, 2006.
- Pelli DG. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10(4): 437–442, 1997.
- Pobric G, Jefferies E, and Ralph MA. Anterior temporal lobes mediate semantic representation: Mimicking semantic dementia by using rTMS in normal participants. *Proceedings of the National Academy of Sciences*, 104(50): 20137–20141, 2007.
- Pulvermüller F and Hauk O. Category-specific conceptual processing of color and form in left fronto-temporal cortex. *Cerebral Cortex*, 16(8): 1193–1201, 2006.
- Ramachandran VS and Hubbard EM. Psychophysical investigations into the neural basis of synaesthesia. *Proceedings of the Royal Society: Biological Science*, 268(1470): 979–983, 2001.
- Rich AN and Mattingley JB. The effects of stimulus competition and voluntary attention on colour–graphemic synaesthesia. *NeuroReport*, 14(14): 1793–1798, 2003.
- Rich AN, Bradshaw JL, and Mattingley JB. A systematic, large-scale study of synaesthesia: Implications for the role of early experience in lexical–colour associations. *Cognition*, 98(1): 53–84, 2005.
- Rich AN, Williams MA, Puce A, Syngeniotis A, Howard MA, McGlone F, et al. Neural correlates of imagined and synaesthetic colours. *Neuropsychologia*, 44(14): 2918–2925, 2006.
- Rich AN and Mattingley JB. Out of sight, out of mind: The attentional blink can eliminate synaesthetic colours. *Cognition*, 114(3): 320–328, 2010.
- Robertson LC. Binding, spatial attention and perceptual awareness. *Nature Review Neuroscience*, 4(2): 93–102, 2003.
- Rouw R and Scholte HS. Increased structural connectivity in grapheme–color synesthesia. *Nature Neuroscience*, 10(6): 792–797, 2007.
- Rouw R, Scholte HS, and Colizoli O. Brain areas involved in synaesthesia: A review. *Journal of Neuropsychology*, 5(2): 214–242, 2011.
- Saenz M, Buracas GT, and Boynton GM. Global feature-based attention for motion and color. *Vision Research*, 43(6): 629–637, 2003.
- Saenz M and Koch C. The sound of change: Visually-induced auditory synesthesia. *Current Biology*, 18(15): R650–R651, 2008.
- Sagiv N, Heer J, and Robertson L. Does binding of synesthetic color to the evoking grapheme require attention? *Cortex*, 42(2): 232–242, 2006.
- Shimojo S and Shams L. Sensory modalities are not separate modalities: Plasticity and interactions. *Current Opinion in Neurobiology*, 11(4): 505–509, 2001.
- Simner J, Mulvenna C, Sagiv N, Tsakanikos E, Witherby SA, Fraser C, et al. Synaesthesia: The prevalence of atypical cross-modal experiences. *Perception*, 35(8): 1024–1033, 2006.
- Simner J. Defining synaesthesia. *British Journal of Psychology*, 103(1): 1–15, 2012.
- Soto D, Hodsoll J, Rotshtein P, and Humphreys GW. Automatic guidance of attention from working memory. *Trends in Cognitive Science*, 12(9): 342–348, 2008.
- Spence C. Crossmodal correspondences: A tutorial review. *Attention, Perception, and Psychophysics*, 73(4): 971–995, 2011.
- Sperling JM, Prvulovic D, Linden DE, Singer W, and Stirn A. Neuronal correlates of colour–graphemic synaesthesia: A fMRI study. *Cortex*, 42(2): 295–303, 2006.
- Todd JT. The visual perception of 3D shape. *Trends in Cognitive Science*, 8(3): 115–121, 2004.
- van Leeuwen TM, Petersson KM, and Hagoort P. Synaesthetic colour in the brain: Beyond colour areas. A functional magnetic resonance imaging study of synaesthetes and matched controls. *PLoS One*, 5(8): e12074, 2010.
- Walker P, Bremner JG, Mason U, Spring J, Mattock K, Slater A, et al. Preverbal infants’ sensitivity to synaesthetic cross-modality correspondences. *Psychological Science*, 21(1): 21–25, 2010.
- Ward J and Simner J. Lexical–gustatory synaesthesia: Linguistic and conceptual factors. *Cognition*, 89(3): 237–261, 2003.
- Ward J, Huckstep B, and Tsakanikos E. Sound–colour synaesthesia: To what extent does it use cross-modal mechanisms common to us all? *Cortex*, 42(2): 264–280, 2006.
- Ward J, Li R, Salih S, and Sagiv N. Varieties of grapheme–colour synaesthesia: A new theory of phenomenological and behavioural differences. *Consciousness and Cognition*, 16(4): 913–931, 2007.
- Ward J, Jonas C, Dienes Z, and Seth A. Grapheme–colour synaesthesia improves detection of embedded shapes, but without pre-attentive ‘pop-out’ of synaesthetic colour. *Proceedings of the Royal Society: Biological Science*, 277(1684): 1021–1026, 2010.