Synesthesia strengthens sound-symbolic cross-modal correspondences

4 authors, including:

Simon Lacey
Emory University

39 PUBLICATIONS   **746** CITATIONS

Kelly McCormick
Emory University

5 PUBLICATIONS   **137** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   Multisensory systems and sound to meaning mappings in the brain. View project

Synesthesia strengthens sound-symbolic cross-modal correspondences

Simon Lacey,[1] Margaret Martinez,[1] Kelly McCormick[1,2] and K. Sathian[1,2,3,4]

[1]Department of Neurology, Emory University, Atlanta, GA 30322, USA
[2]Department of Psychology, Emory University, Atlanta, GA 30322, USA
[3]Department of Rehabilitation Medicine, Emory University, Atlanta, GA 30322, USA
[4]Center for Visual and Neurocognitive Rehabilitation, Atlanta VAMC, Decatur, GA 30033, USA

Abstract

Synesthesia is a phenomenon in which an experience in one domain is accompanied by an involuntary secondary experience in another, unrelated domain; in classical synesthesia, these associations are arbitrary and idiosyncratic. Cross-modal correspondences refer to universal associations between seemingly unrelated sensory features, e.g., auditory pitch and visual size. Some argue that these phenomena form a continuum, with classical synesthesia being an exaggeration of universal cross-modal correspondences, whereas others contend that the two are quite different, since cross-modal correspondences are non-arbitrary, non-idiosyncratic, and do not involve secondary experiences. Here, we used the implicit association test to compare synesthetes' and non-synesthetes' sensitivity to cross-modal correspondences. We tested the associations between auditory pitch and visual elevation, auditory pitch and visual size, and sound-symbolic correspondences between auditory pseudowords and visual shapes. Synesthetes were more sensitive than non-synesthetes to cross-modal correspondences involving sound-symbolic, but not low-level sensory, associations. We conclude that synesthesia heightens universally experienced cross-modal correspondences, but only when these involve sound symbolism. This is only partly consistent with the idea of a continuum between synesthesia and cross-modal correspondences, but accords with the idea that synesthesia is a high-level, post-perceptual phenomenon, with spillover of the abilities of synesthetes into domains outside their synesthesias. To our knowledge, this is the first demonstration that synesthetes, relative to non-synesthetes, experience stronger cross-modal correspondences outside their synesthetic domains.

Introduction

Synesthesia is characterized by the accompaniment of an experience in one domain by an involuntary, secondary experience in another, unrelated domain (Eagleman *et al.*, 2007; Ward, 2013). In classical synesthesia, these associations are both arbitrary and idiosyncratic (Deroy & Spence, 2013; Ward, 2013). Cross-modal correspondences refer to near-universally experienced associations between seemingly unrelated sensory features; for example, auditory stimuli of high and low pitch match with visual stimuli of high and low spatial elevation, respectively (Spence, 2011). A special class of cross-modal correspondences is termed sound-symbolic; for example, when matching auditorily-presented pseudowords with visual shapes, the pseudowords 'takete' and 'maluma' are associated with pointed and rounded shapes, respectively (Köhler, 1929, 1947).

Some hold that classical synesthesia represents one end of a continuum of the strength of cross-modal correspondences (Martino & Marks, 2001; Marks & Mulvenna, 2013) while others argue against

such a continuum since cross-modal correspondences are non-arbitrary, non-idiosyncratic, and do not involve secondary experiences (Deroy & Spence, 2013, 2016). Reports that non-synesthetes make non-random grapheme-color associations, some of which accord with those observed in synesthetes (Simner *et al.*, 2005), and that specific grapheme-color associations can be acquired by non-synesthetes through learning (Colizoli *et al.*, 2012; Bor *et al.*, 2014), support the continuum concept. On the other hand, evidence that synesthesia has a genetic basis (Barnett *et al.*, 2008; Mitchell, 2011; Newell & Mitchell, 2016) implies that synesthetic experiences may have a different basis from the near-universal cross-modal correspondences experienced by synesthetes and non-synesthetes alike. Previous research using cross-modal illusions is inconclusive with respect to this debate, with one study showing that synesthetes and non-synesthetes are equally susceptible (Whittingham *et al.*, 2014); another that synesthetes are less susceptible, that is, they demonstrated reduced audio-visual integration (Neufeld *et al.*, 2012); and a third finding that synesthetes were more susceptible (Brang *et al.*, 2012).

Here, we compared synesthetes' and non-synesthetes' sensitivity to cross-modal correspondences using the implicit association test

Correspondence: Professor K. Sathian, [1]Department of Neurology, as above.
E-mail: krish.sathian@emory.edu

(IAT: Greenwald *et al.*, 1998; Parise & Spence, 2012) to test the hypothesis that synesthesia represents one end of a continuum of the strength of cross-modal correspondences. Originally devised as a test of social attitudes (Greenwald *et al.*, 1998), the IAT has been successfully used to test the very different associations involved in cross-modal correspondences (Parise & Spence, 2012). The underlying principle is the same: response times (RTs) are faster if the stimuli assigned to a particular key are congruent and slower if they are incongruent (Greenwald *et al.*, 1998; Parise & Spence, 2012). The advantage of the IAT for testing cross-modal correspondences is that presenting each stimulus in isolation rules out the possibility that RTs are slower for incongruent pairings because of selective attention effects (Parise & Spence, 2012). If synesthesia and cross-modal correspondences are fundamentally different (Deroy & Spence, 2013, 2016), then congruency effects should be similar in synesthetes and non-synesthetes. If, however, synesthesia and cross-modal correspondences form a continuum, then synesthetes should exhibit larger congruency effects than non-synesthetes, that is, synesthetes should respond faster on congruent trials and/or slower on incongruent trials.

Methods

Participants

Thirty-eight people took part in response to flyers posted around the Emory University campus soliciting both people who knew, or thought, they were synesthetic and those who knew they were not; respondents were predominantly students. Seventeen participants (five male, 12 female; mean age 25 years, 2 months) were confirmed as synesthetes using the online Synesthesia Battery (SB: Eagleman *et al.*, 2007). The SB essentially tests the consistency of associations for a number of synesthesias (to the extent possible online): scores < 1.0 are considered to indicate synesthesia while scores > 2.0 indicate the absence of synesthesia. SB scores between 1.0 and 2.0 are considered inconclusive; therefore, three individuals (two male, one female) with intermediate scores in this range were excluded from analysis. The SB responses showed that the primary synesthesias in the synesthete group were grapheme-color ($n = 6$), involving letters (including one for the Greek alphabet), numbers, or both; month-color ($n = 5$); musical instrument-color ($n = 3$); sequence-space ($n = 2$) for months, days, and a number line; and personality-color ($n = 1$). Fourteen synesthetes were associators, who experience their concurrents in the 'mind's eye', and three were projectors, who experience their concurrents in external space (Dixon *et al.*, 2004). Nine individuals reported secondary synesthesias. Non-synesthetic controls were also recruited, including individuals who self-identified as synesthetic but who had SB scores > 2.0 ($n = 3$), giving a control group of 18 (nine male, nine female; mean age 22 years, 11 months. The control and synesthete groups were not significantly different in age ($t_{1,32} = 0.6$, $P = 0.5$). All participants provided written informed consent and were compensated for their time. All procedures were approved by the Emory University Institutional Review Board.

Stimuli

We tested the cross-modal correspondences between high/low pitch and high/low visual spatial position, high/low pitch and small/large visual size, and a form of sound symbolism in which the two-syllable (auditorily presented) pseudowords 'lomo' and 'kike' (/lomo/;/kɛke/) were associated with (visually presented) rounded and pointed shapes, respectively (Fig. 1). The pseudowords we used, as

well as the visual shapes they were paired with, lie near the ends of the pointed-rounded dimension in each modality based on empirical ratings (McCormick *et al.*, 2015; S. List, unpublished thesis: Emory University). The pseudowords each contained a repetition of either a consonant- or vowel-sound but not both.

The pseudowords 'lomo' and 'kike' (/lomo/;/kɛke/) were digitally recorded (female voice) using Audacity v2.0.1 (Audacity Team, 2012) using a SHURE 5115D microphone and an EMU 0202 USB external sound card at a 44.1 kHz sampling rate. These recordings were then processed in Audacity, using standard tools and default settings for sound amplification, removal of noise, and finally, high- and low-pass filtering to remove ambient noise. Stimulus duration was 0.5 s for 'kike' and 0.6 s for 'lomo'. The rounded and pointed shapes were presented as black line drawings in a grey square on an otherwise black background (Fig. 1). The shapes each subtended approximately $4.5° \times 4.5°$ of visual angle (all visual angle measurements are for a viewing distance of approximately 60 cm).

The high- (1440 Hz) and low- (180 Hz) pitched tones were generated using Audacity v2.0.1 and each lasted 1 s. The size stimuli were two white circles presented on a black background; the small circle subtended 1°, and the large circle 2.5°, of visual angle (Fig. 1). The visual elevation stimulus consisted of two white squares of 1° side, diagonally offset and flipping horizontally at 4 Hz to give a 'checkerboard' effect, presented on a black background at 5° above or below a central fixation cross. Some studies used dynamic pitch-elevation stimuli, that is, ascending/descending visual dots and auditory scales (e.g. Walker *et al.*, 2010; Lewkowicz & Minar, 2014) because these made high/low information directly available in the stimulus. However, these are not necessary to demonstrate correspondences and, in common with other studies (e.g., Bernstein & Edelstein, 1971; O'Leary & Rhodes, 1984; Melara & O'Brien, 1987; Melara, 1989; Miller, 1991; Ben-Artzi & Marks, 1995; Patching & Quinlan, 2002; Evans & Treisman, 2010), our single tones and dots were static. While this meant that high/low pitch/elevation and large/small size stimuli were all relative to each other, the use of practice trials (see below) enabled participants to clearly differentiate between high/low and large/small.

Procedure

Before the main experiment started, each participant listened to the high-pitched tone at a range of amplitudes and selected the loudest tone that was still comfortable. This high-pitched tone was then compared to a range of low-pitched tones similarly varying in amplitude; participants selected the low tone that they perceived as matching the high tone in loudness. The main IAT experiment was presented via Presentation software (Neurobehavioral Systems Inc., Albany, CA, USA) which also recorded RTs. In the IAT, participants were instructed to associate pairs of stimuli with one of two response keys

		Congruent		Incongruent	
Response key		←	→	←	→
Pitch-Elevation	Auditory	ꟺ	∿	ꟺ	∿
	Visual				
Pitch-Size	Auditory	ꟺ	∿	ꟺ	∿
	Visual				
Pseudoword-Shape	Auditory	"lomo"	"kike"	"lomo"	"kike"
	Visual				

Fig. 1. Schematic of the cross-modal correspondences and response key pairings used in the IAT experiment.

(the 'left' and 'right' arrow keys on a normal US 'QWERTY' keyboard). The pairs always consisted of one auditory and one visual stimulus and, in separate blocks of trials, were either congruent or incongruent. For example, Fig. 1 shows that for the pseudoword-shape correspondence, the congruent pairs were 'lomo'/rounded shape (both to be associated with the left arrow key) and 'kike'/pointed shape (associated with the right arrow key) and the incongruent pairs were 'lomo'/pointed shape (left arrow key) and 'kike'/rounded shape (right arrow key). However, each stimulus was presented in isolation, that is, a trial consisted either of an auditory presentation ('lomo' or 'kike') or a visual presentation (rounded or pointed shape) and participants were asked to respond as quickly as possible. The pitch-elevation and pitch-size stimuli were similarly paired (see Fig. 1).

Each cross-modal correspondence was tested in two runs, each run comprising 96 trials divided into a block of 48 congruent trials followed by a block of 48 incongruent trials, or vice versa, for a total of 192 trials across the two runs. Each block of a run was preceded by an instruction screen describing the response key associations to be used and by 12 practice trials (not included in the analysis) with on-screen feedback as to accuracy (for the practice trials only). Within each run, half the trials were auditory (high/low pitch or lomo/kike) and half were visual (high/low position, small/large size, or rounded/pointed shape), split evenly between congruent and incongruent trials. Trials consisted of a blank 1000 ms followed by either a visual or an auditory stimulus for 1000 ms (except for the pseudowords which occupied 500 ms or 600 ms, see above) and were terminated either by the participant pressing a response key or automatically 3500 ms after stimulus onset. RTs were measured from stimulus onset. The length of each active block thus varied between participants but was a maximum of 330 s. All participants were tested on both runs for all correspondences. The order in which the correspondences were presented and the order of the two runs were counterbalanced as far as possible across participants; the two runs for each correspondence were completed before moving on to the next. Following completion of the IAT, all participants completed the SB (Eagleman *et al.*, 2007) to assess their synesthetic status. We chose this order because, if the continuum hypothesis were correct, running the SB first might have primed performance on the IAT.

Data from one control participant were excluded from analysis because of a high proportion (> 24%) of incorrect responses for each of the cross-modal correspondences. Analyses were based only on correct responses (95.8% of the total responses) and excluding RTs more than three standard deviations from the individual means for each run (1.4% of the total correct responses). The congruency effect was calculated as:

$$((\text{incongruent RT} - \text{congruent RT})/\text{incongruent RT}) \times 100 \quad (1)$$

We used this as a measure of participants' sensitivity to cross-modal correspondences, that is, how much faster they responded to congruent trials compared to incongruent trials. Data were initially analyzed in a global repeated-measures ANOVA (RM-ANOVA) with the congruency effect as the dependent variable; between-group factor: presence/absence of synesthesia (synesthetes, controls) and within-group factors: correspondence type (pitch-size, pitch-elevation, pseudoword-shape) and modality (auditory, visual).

Results

A global RM-ANOVA showed that synesthetes had larger congruency effects than non-synesthetes ($F_{1,32} = 6.9$, $P = 0.013$, $\eta^2 = 0.2$).

There was a main effect of correspondence type ($F_{2,64} = 3.3$, $P = 0.04$, $\eta^2 = 0.09$), with larger congruency effects for pitch-elevation (mean $\pm$ SEM 17.9 ± 2.5) and pseudoword-shape (15.4 ± 3.0) than for pitch-size (9.3 ± 2.9) although *post hoc t*-tests showed that none of these differences was significant (Bonferroni-corrected $\alpha = 0.016$: all $t_{33} > -1.6$, all $P > 0.02$). There was a significant interaction between synesthesia and correspondence type ($F_{2,64} = 3.3$, $P = 0.04$, $\eta^2 = 0.1$). Analysis of this interaction using separate RM-ANOVAs for each correspondence (Bonferroni-corrected $\alpha = 0.016$) showed that synesthetes were more sensitive than non-synesthetes to the pseudoword-shape correspondence ($F_{1,32} = 15.3$, $P = 0.001$, $\eta^2 = 0.3$; but not the pitch-size or pitch-elevation correspondences (both $F_{1,32} < 0.7$, $P > 0.4$) (Fig. 2). Inspection of the absolute data showed that RTs were faster for congruent than incongruent trials in both modalities and for all three correspondences for both synesthetes and non-synesthetes. However, for the non-synesthetes, this difference was only significant for pitch-elevation (Bonferroni-corrected $\alpha = 0.008$: auditory $t_{16} = -4.0$, $P = 0.001$; visual $t_{16} = -3.2$, $P = 0.006$), while for synesthetes this difference was significant for both pitch-elevation (auditory $t_{16} = -3.4$, $P = 0.004$; visual $t_{16} = -3.8$, $P = 0.002$) and pseudoword-shape (auditory $t_{16} = -6.4$, $P < 0.001$; visual $t_{16} = -5.1$, $P < 0.001$) – the differences in pitch-size for synesthetes were only significant at an uncorrected $\alpha = 0.05$ (Fig. S1). (Both groups contained individuals who had faster RTs to incongruent trials, though there were slightly more of these in the non-synesthete group, see Table S1).

As there were more grapheme- and month-color synesthetes than any other kind, we tested whether either of these sub-groups was driving the main result. However, congruency effects were not significantly different between grapheme-color and other synesthetes ($F_{1,15} = 1.2$, $P = 0.3$) and there was no interaction with correspondence type ($F_{2,30} = 0.8$, $P = 0.5$). In view of the small and unequal group sizes, these results were confirmed with non-parametric Mann–Whitney tests. There was also no significant difference between month-color and other synesthetes ($F_{1,15} = 0.3$, $P = 0.6$), but there was an interaction with correspondence type ($F_{2,30} = 5.1$, $P = 0.01$) in which month-color synesthetes had larger congruency effects for pitch-size (23.5 ± 5.0 vs. 6.3 ± 5.4) and smaller effects for pitch-elevation (12.1 ± 3.7 vs. 23.2 ± 4.5). However, these differences were not significant in either Bonferroni-corrected *post hoc t*-tests or Mann–Whitney tests and, importantly, the mean congruency effects for the pseudoword-shape correspondence were very similar between month-color and other synesthetes (27.7 ± 4.0 vs. 24.2 ± 3.4). We therefore conclude that neither of these sub-groups was driving the main synesthesia $\times$ correspondence interaction.

For the synesthetic group, there was no relationship between congruency effects for any of the three correspondences and (i) the strength of their synesthetic associations as measured by SB scores (all $r < 0.43$, all $P > 0.13$), (ii) their projector/associator scores (all $r < 0.17$, all $P > 0.5$), or (iii) their vividness of visual imagery (VVIQ-2: Marks (1995) incorporated into the SB) scores (all $r < 0.39$, all $P > 0.15$). Furthermore, SB, projector/associator, and VVIQ-2 scores were also uncorrelated with one another (all $r < 0.5$, $P > 0.07$). VVIQ-2 scores and congruency effects were also uncorrelated in the non-synesthetic control group (all $r < -0.27$, all $P > 0.3$). Although the synesthete and control groups were not matched for gender, we did not find any effect of gender in either the synesthete or non-synesthete groups, nor overall (all $F < 0.07$, all $P > 0.4$).

Pseudoword - Shape
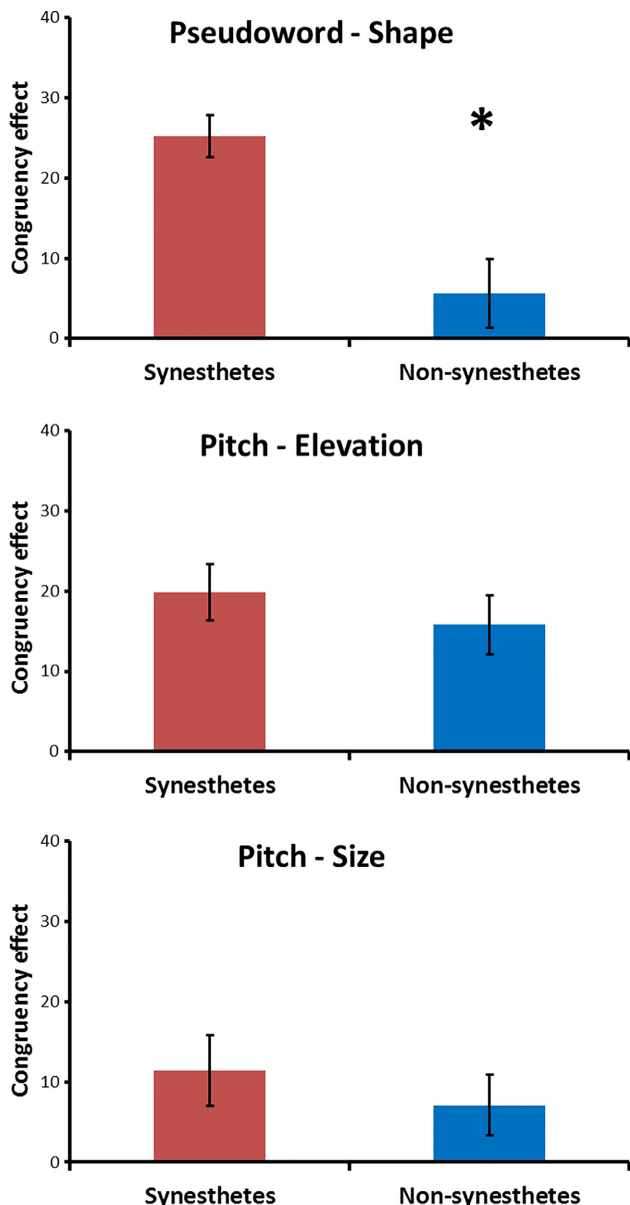

Pitch - Elevation

Pitch - Size

FIG. 2. Synesthetes were more sensitive (had larger congruency effects) than non-synesthetes to cross-modal correspondences involving sound-symbolic (pseudoword-shape), but not low-level sensory (pitch-elevation/pitch-size) associations (*$P = 0.001$). Although the global RM-ANOVA showed that visual congruency effects were larger than auditory effects ($F_{1,32} = 6.7$, $P = 0.01$, $\eta^2 = 0.2$), modality did not interact with either correspondence type or synesthetic status (both $F_{2,64} < 1.0$, $P > 0.4$) and, in the correspondence-specific RM-ANOVAs, modality was only significant for pitch-size ($F_{1,32} = 4.7$, $P = 0.04$, $\eta^2 = 0.1$), therefore, the data displayed above are collapsed across modalities.

Discussion

Our study indicates that synesthetes, compared to non-synesthetes, demonstrate stronger congruency effects for a high-level sound-symbolic association between auditory pseudowords and visual shapes but not for the low-level sensory associations of auditory pitch with visual size or elevation. This finding is in keeping with the view of synesthesia as a high-level post-perceptual phenomenon (Ward *et al.*, 2006; Chiou & Rich, 2014; see also Brang *et al.*, 2011; who argue for an interaction between perceptual and conceptual factors).

Synesthetes in this study also exhibited a tendency for stronger congruency effects, compared to non-synesthetes, for pitch-elevation and pitch-size pairings, although the between-group differences were not significant. Thus, while we cannot completely rule out stronger congruency effects for these pairings as well, the significant group-by-correspondence interaction indicates that if between-group differences in the pitch-size/elevation correspondences exist, they are substantially smaller than for the sound-symbolic association. The lack of correlation between congruency effects in synesthetes and their projector–associator scores suggests that our findings may be independent of whether a synesthete is a projector or associator, although a caveat is that our sample of synesthetes only contained three projectors, so a larger group of projectors should be studied. Consistent with previous studies of cross-modal illusions that found no differences between synesthetic sub-types (Neufeld *et al.*, 2012; Whittingham *et al.*, 2014), we also found that the present results were not driven by the grapheme- or month-color synesthetic subgroups.

Although the raw RT data showed that neither group exhibited significant differences between congruent and incongruent trials for the pitch-size correspondence, and that non-synesthetes were additionally insensitive to the pseudoword-shape correspondence, RTs were faster for congruent trials in all cases, as would be expected. One reason for the lack of significant differences in our data may be that that the response key pairings and instructions were changed much less often compared to Parise & Spence (2012), perhaps enabling participants to learn the associations and settle into a response set. However, it is important to remember that absolute RTs do not take account of individual RT variability, which is why we used Eqn 1 to calculate the magnitude of congruency effects.

Superficially, the finding of a stronger sound-symbolic congruency effect for synesthetes is consistent with the continuum hypothesis (Martino & Marks, 2001; Marks & Mulvenna, 2013) especially since the sound-symbolic pseudoword-shape association is not part of the classical synesthesias experienced by our synesthetic participants. However, the enhancement for synesthetes was limited to the sound-symbolic correspondence, consistent with the view of synesthesia as a conceptual, post-perceptual phenomenon (Ward *et al.*, 2006; Chiou & Rich, 2014), whereas the continuum hypothesis predicts a general enhancement across all types of cross-modal correspondence (Martino & Marks, 2001; Marks & Mulvenna, 2013). Thus, our findings only partially support the continuum hypothesis. The results of this study are also consistent with recent findings that synesthetes are better than non-synesthetes at guessing the meaning of sound-symbolic words in a foreign language (Bankieris & Simner, 2015; Lockwood *et al.*, 2016). Our findings in no way negate the unique aspects of synesthetic experience (Deroy & Spence, 2013, 2016) or the genetic characteristics of synesthetes (Barnett *et al.*, 2008; Mitchell, 2011; Newell & Mitchell, 2016; see also Gregersen *et al.*, 2013), both of which suggest that synesthesia is a distinct entity.

Our findings fit within a context of superior performance of synesthetes across a variety of domains. Some of these reflect the general nature of synesthesia as a high-level conceptual phenomenon, for example, the finding that synesthetes may have higher creativity scores (Ward *et al.*, 2008; Chun & Hupé, 2016) and better memory (Rothen *et al.*, 2012; Ward *et al.*, 2013). Others relate to more specific aspects: For example, while synesthesia and mental imagery are considered discrete rather than continuous phenomena (Craver-Lemley & Reeves, 2013; Spence & Deroy, 2013), synesthetes report more vivid mental imagery (Barnett & Newell, 2008; Spiller *et al.*, 2015) but only in those sensory modalities involved in

their particular synesthesia (Spiller *et al.*, 2015). Several studies show enhanced mental rotation ability in sequence-space synesthesia (Brang *et al.*, 2013; Havlik *et al.*, 2015) with performance enhanced still further in projectors compared to associators (Havlik *et al.*, 2015). These observations may reflect genetic factors determining only a general predisposition to synesthesia, with the particular type, specific associations, and individual differences, including those related to projector/associator sub-groups, being modulated by other factors (Barnett *et al.*, 2008; Mitchell, 2011; Newell & Mitchell, 2016). This line of thinking would be consistent with our finding of stronger cross-modal correspondences for a heterogeneous group of synesthetes with no influence of any particular type of synesthesia.

Our findings also raise questions about causal mechanisms for these instances of enhanced synesthetic performance on non-synesthetic tasks. One explanation may be that such performance 'comes for free' on the back of structural differences in the synesthetic brain. For example, it is well-known, principally from studies of grapheme-color synesthesia, that synesthetes exhibit hyperconnectivity between cortical areas potentially related to their synesthetic associations (e.g., Dovern *et al.*, 2012; Sinke *et al.*, 2012; Tomson *et al.*, 2013; see Zamm *et al.*, 2013 for colored-hearing synesthesia) as well as more global differences compared to non-synesthetes (e.g., Bargary & Mitchell, 2008; Dovern *et al.*, 2012). Such hyperconnectivity might give rise not only to a particular form of synesthesia but also result in enhanced performance for any task, or class of stimuli, relying on the same pathways. Thus, greater white matter integrity for synesthetes in the inferior fronto-occipital fasciculus, which connects visual and auditory association cortex to frontal cortex, could support not only colored-hearing synesthesia (Zamm *et al.*, 2013), but also more vivid visual and auditory imagery (Spiller *et al.*, 2015). Similarly, hyperconnectivity in cortical areas involved in sequence-space synesthetes might also result in enhanced spatial imagery (Brang *et al.*, 2013; Havlik *et al.*, 2015). Such studies will continue to be important because connectivity patterns differentiate between projector and associator synesthetes (Rouw & Scholte, 2007; van Leeuwen *et al.*, 2011) and also correlate with behavioral measures of synesthetic associations (Dovern *et al.*, 2012; Zamm *et al.*, 2013).

Conclusion

We provide the first empirical evidence that, synesthetes, relative to non-synesthetes, demonstrate tighter cross-modal correspondences outside their synesthetic domains. However, this was limited to the sound-symbolic pseudoword-shape correspondence and did not extend to low-level sensory correspondences. This is only partially consistent with the idea that synesthetic and more universally experienced cross-modal correspondences form a continuum, but fits well with the notion of synesthesia as a high-level, post-perceptual phenomenon, and with genetic evidence suggesting that synesthetes are distinct from non-synesthetes. Further inquiry is necessary to establish whether these findings extend to high-level correspondences other than those involved in sound symbolism, and the underlying mechanisms.

Supporting Information

Additional supporting information can be found in the online version of this article:
Fig. S1. Raw response times (RTs) for congruent and incongruent trials in each modality for each cross-modal correspondence (correct trials only and excluding RTs more than three standard deviations from the individual means).
Table S1. Number of participants in synesthete and non-synesthete groups who were sensitive to each cross-modal correspondence (i.e., RTs were faster for congruent trials than incongruent trials) in each modality.

References

Audacity Team. (2012) Audacity v2.0.1 [Computer program]. Available http://audacity.sourceforge.net/. Audacity® software is copyright © 1999–2014 Audacity Team.

Bankieris, K. & Simner, J. (2015) What is the link between synaesthesia and sound symbolism? *Cognition*, **136**, 186–195.

Bargary, G. & Mitchell, K.J. (2008) Synesthesia and cortical connectivity. *Trends Neurosci.*, **31**, 335–342.

Barnett, K.J. & Newell, F.N. (2008) Synesthesia is associated with enhanced, self-rated visual imagery. *Conscious. Cogn.*, **17**, 1032–1039.

Barnett, K.J., Finucane, C., Asher, J.E., Bargary, G., Corvin, A.P., Newell, F.N. & Mitchell, K.J. (2008) Familial patterns and the origins of individual differences in synesthesia. *Cognition*, **106**, 871–893.

Ben-Artzi, E. & Marks, L.E. (1995) Visual-auditory interaction in speeded classification: role of stimulus difference. *Percept. Psychophys.*, **57**, 1151–1162.

Bernstein, I.H. & Edelstein, B.A. (1971) Effects of some variations in auditory input upon visual choice reaction time. *J. Exp. Psychol.*, **87**, 241–247.

Bor, D., Rothen, N., Schwartzman, D.J., Clayton, S. & Seth, A.K. (2014) Adults can be trained to acquire synesthetic experiences. *Sci. Rep.*, **4**, 7089.

Brang, D., Rouw, R., Ramachandran, V.S. & Coulson, S. (2011) Similarly shaped letters evoke similar colors in grapheme-color synesthesia. *Neuropsychologia*, **49**, 1355–1358.

Brang, D., Williams, L.E. & Ramachandran, V.S. (2012) Grapheme-color synesthetes show enhanced crossmodal processing between auditory and visual modalities. *Cortex*, **48**, 630–637.

Brang, D., Miller, L.E., McQuire, M., Ramachandran, V.S. & Coulson, S. (2013) Enhanced mental rotation ability in time-space synesthesia. *Cogn. Process.*, **14**, 429–434.

Chiou, R. & Rich, A.N. (2014) The role of conceptual knowledge in understanding synesthesia: evaluating contemporary findings from a 'hub-and-spokes' perspective. *Front. Psychol.*, **5**, 105.

Chun, C.A. & Hupé, J.-M. (2016) Are synesthetes exceptional beyond their synesthetic associations? A systematic comparison of creativity, personality, cognition, and mental imagery in synesthetes and controls. *Brit. J. Psychol.*, **107**, 397–418.

Colizoli, O., Murre, J.M.J. & Rouw, R. (2012) Pseudo-synesthesia through reading books with colored letters. *PLoS One*, **7**, e39799.

Craver-Lemley, C. & Reeves, A. (2013) Is synesthesia a form of mental imagery? In Lacey, S. & Lawson, R. (Eds), *Multisensory Imagery*. Springer, New York, USA, pp. 185–206.

Deroy, O. & Spence, C. (2013) Why we are not all synesthetes (not even weakly so). *Psychon. B. Rev.*, **20**, 643–664.

Deroy, O. & Spence, C. (2016) Lessons of synaesthesia for consciousness: learning from the exception, rather than the general. *Neuropsychologia*, **88**, 49–57.

Dixon, M.J., Smilek, D. & Merikle, P.M. (2004) Not all synaesthetes are created equal: projector versus associator synesthetes. *Cogn. Affect. Behav. Ne.*, **4**, 335–343.

Dovern, A., Fink, G.R., Fromme, C.B., Wohlschläger, A.M., Weiss, P.H. & Riedl, V. (2012) Intrinsic network connectivity reflects consistency of synesthetic experiences. *J. Neurosci.*, **32**, 7614–7621.

Eagleman, D.M., Kagan, A.D., Nelson, S.S., Sagaram, D. & Sarma, A.K. (2007) A standardized test battery for the study of synesthesia. *J. Neurosci. Meth.*, **159**, 139–145.

Evans, K.K. & Treisman, A. (2010) Natural cross-modal mappings between visual and auditory features. *J. Vision*, **10**, 6.

Greenwald, A.G., McGhee, D.E. & Schwarz, J.L.K. (1998) Measuring individual differences in implicit cognition: the implicit association test. *J. Pers. Soc. Psychol.*, **74**, 1464–1480.

Gregersen, P.K., Kowalsky, E., Lee, A., Baron-Cohen, S., Fisher, S.E., Asher, J.E., Ballard, D., Freudenberg, J. *et al.* (2013) Absolute pitch exhibits phenotypic and genetic overlap with synesthesia. *Hum. Mol. Genet.*, **22**, 2097–2104.

Havlik, A.M., Carmichael, D.A. & Simner, J. (2015) Do sequence-space synaesthetes have better spatial imagery skills? Yes, but there are individual differences. *Cogn. Process.*, **16**, 245–253.

Köhler, W. (1929) *Gestalt Psychology.* Liveright, New York.

Köhler, W. (1947) *Gestalt Psychology: An Introduction to New Concepts in Modern Psychology.* Liveright, New York.

van Leeuwen, T.M., den Ouden, H.E.M. & Hagoort, P. (2011) Effective connectivity determines the nature of subjective experience in grapheme-color synesthesia. *J. Neurosci.*, **31**, 9879–9884.

Lewkowicz, D.J. & Minar, N.J. (2014) Infants are not sensitive to synesthetic cross-modality correspodences: a comment of Walker et al. (2010) *Psychol. Sci.*, **25**, 832–834.

Lockwood, G., van Leeuwen, T., Drijvers, L. & Dingemanse, M. (2016) *Synaesthesia and sound symbolism*. Poster, Synaesthesia and Cross-Modal Perception, Dublin, April 21–23, 2016.

Marks, D.F. (1995) New directions for mental imagery research. *J. Ment. Imagery*, **19**, 153–167.

Marks, L.E. & Mulvenna, C.M. (2013) Synesthesia at and near its borders. *Front. Psychol.*, **4**, 651.

Martino, G. & Marks, L.E. (2001) Synesthesia: strong and weak. *Curr. Dir. Psychol. Sci.*, **10**, 61–65.

McCormick, K., Kim, J.Y., List, S. & Nygaard, L.C. (2015) Sound to meaning mappings in the bouba-kiki effect. *Proceedings 37th Annual Meeting Cognitive Science Society*, 1565–1570.

Melara, R.D. (1989) Similarity relations among synesthetic stimuli and their attributes. *J. Exp. Psychol. Human*, **15**, 212–231.

Melara, R.D. & O'Brien, T.P. (1987) Interaction between synesthetically corresponding dimensions. *J. Exp. Psychol. Gen.*, **116**, 323–336.

Miller, J. (1991) Channel interaction and the redundant targets effect in bimodal divided attention. *J. Exp. Psychol. Human*, **17**, 160–169.

Mitchell, K.J. (2011) Curiouser and curiouser: genetic disorders of cortical specialization. *Curr. Opin. Genet. Dev.*, **21**, 271–277.

Neufeld, J., Sinke, C., Zedler, M., Emrich, H.M. & Szycik, G.R. (2012) Reduced audio-visual integration in synaesthetes indicated by the double-flash illusion. *Brain Res.*, **1473**, 78–86.

Newell, F.N. & Mitchell, K.J. (2016) Multisensory integration and cross-modal learning in synaesthesia: a unifying model. *Neuropsychologia*, **88**, 140–150.

O'Leary, A. & Rhodes, G. (1984) Cross-modal effects on visual and auditory object perception. *Percept. Psychophys.*, **35**, 565–569.

Parise, C.V. & Spence, C. (2012) Audiovisual cross-modal correspondences and sound symbolism: a study using the implicit association test. *Exp. Brain Res.*, **220**, 319–333.

Patching, G.R. & Quinlan, P.T. (2002) Garner and congruence effects in the speeded classification of bimodal signals. *J. Exp. Psychol. Human*, **28**, 755–775.

Rothen, N., Meier, B. & Ward, J. (2012) Enhanced memory ability: insights from synaesthesia. *Neurosci. Biobehav. R.*, **36**, 1952–1963.

Rouw, R. & Scholte, H.S. (2007) Increased structural connectivity in grapheme-color synesthesia. *Nat. Neurosci.*, **10**, 792–797.

Simner, J., Ward, J., Lanz, M., Jansari, A., Noonan, K., Glover, L. & Oakley, D.A. (2005) Non-random associations of graphemes to colours in synaesthetic and non-synaesthetic populations. *Cogn. Neuropsychol.*, **22**, 1069–1085.

Sinke, C., Neufeld, J., Emrich, H.M., Dillo, W., Bleich, S., Zedler, M. & Szycik, G.R. (2012) Inside a synesthete's head: a functional connectivity analysis with grapheme-color synesthetes. *Neuropsychologia*, **50**, 3363–3369.

Spence, C. (2011) Cross-modal correspondences: a tutorial review. *Atten. Percept. Psycho.*, **73**, 971–995.

Spence, C. & Deroy, O. (2013) Cross-modal mental imagery. In Lacey, S. & Lawson, R. (Eds), *Multisensory Imagery*. Springer, New York, USA, pp. 157–183.

Spiller, M.J., Jonas, C.N., Simner, J. & Jansari, A. (2015) Beyond visual imagery: how modality-specific is enhanced mental imagery in synesthesia? *Conscious. Cogn.*, **31**, 73–85.

Tomson, S.N., Narayan, M., Allen, G.I. & Eagleman, D.M. (2013) Neural networks of colored sequence synesthesia. *J. Neurosci.*, **33**, 14098–14106.

Walker, P., Bremner, J.G., Mason, U., Spring, J., Mattock, K., Slater, A. & Johnson, S.P. (2010) Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. *Psychol. Sci.*, **21**, 21–25.

Ward, J. (2013) Synesthesia. *Annu. Rev. Psychol.*, **64**, 49–75.

Ward, J., Huckstep, B. & Tsakanikos, E. (2006) Sound-colour synaesthesia: to what extent does it use cross-modal mechanisms common to us all? *Cortex*, **42**, 264–280.

Ward, J., Thompson-Lake, D., Ely, R. & Kaminski, F. (2008) Synaesthesia, creativity and art: what is the link? *Brit. J. Psychol.*, **99**, 127–141.

Ward, J., Hovard, P., Jones, A. & Rothen, N. (2013) Enhanced recognition memory in grapheme-color synaesthesia for different categories of visual stimuli. *Front. Psychol.*, **4**, 762.

Whittingham, K.M., McDonald, J.S. & Clifford, C.W.G. (2014) Synesthetes show normal sound-induced flash fission and fusion illusions. *Vision. Res.*, **105**, 1–9.

Zamm, A., Schlaug, G., Eagleman, D.M. & Loui, P. (2013) Pathways to seeing music: enhanced structural connectivity in colored-music synesthesia. *NeuroImage*, **74**, 359–366.