Journal of Neuropsychology (2011), 5, 353–371

The British Psychological Society

www.wileyonlinelibrary.com

Is synaesthesia one condition or many? A large-scale analysis reveals subgroups

Scott Novich[1,2], Sherry Cheng[1] and David M. Eagleman[1,3]*

[1]Department of Neuroscience, Baylor College of Medicine, Houston, Texas, USA
[2]Department of Electrical and Computer Engineering, Rice University, Houston, Texas, USA
[3]Department of Psychiatry, Baylor College of Medicine, Houston, Texas, USA

Synaesthesia is a broadly defined neural phenomenon in which stimulation of a sense or concept triggers a second perception not normally associated with the stimulus. For example, letters or numbers may trigger a colour experience, sounds may trigger a taste sensation, or tastes may trigger a feeling of touch. Dozens of forms of synaesthesia have been reported, but the relationship between the different forms has not been studied: is someone with a particular form of synaesthesia likely to possess other types? If so, which ones? As an inroad to illuminating underlying mechanisms, we here examine which different synaesthesia types tend to co-occur. We analyzed reports of the forms of synaesthesia experienced by 19,133 participants who completed the Synaesthesia Battery (Eagleman, Kagan, Nelson, Sagaram, & Sarma, 2007), using correlation analysis, exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and multidimensional scaling (MDS). Our analyses converged on the finding of five distinct groupings of synaesthesia forms. We label these coloured sequence synaesthesias (CSSs), coloured music synaesthesias, non-visual sequela synaesthesias, spatial sequence synaesthesia (SSS), and coloured sensation synaesthesias. Collectively, our findings reveal that synaesthesia is an umbrella term that encompasses several distinct groups with independent probabilities of expression, and this may in turn suggest distinct underlying mechanisms and the possibility of different genetic bases.

In synaesthesia, ordinary stimuli elicit anomalous perceptual experiences (Cytowic, 2002; Cytowic & Eagleman, 2009; Hubbard & Ramachandran, 2005; Robertson & Sagiv, 2004). For example, some synaesthetes experience a specific colour in response to letters, numbers, weekdays, months (Baron-Cohen, Burt, Smith-Laittan, Harrison, & Bolton, 1996; Eagleman & Goodale, 2009; Hubbard, Arman, Ramachandran, & Boyman, 2005; Palmeri, Blake, Marois, Flanery, & Whetsell, 2002; Rouw & Scholte, 2007; Smilek, Dixon, Cudahy, & Merikle, 2002), sounds (Baron-Cohen, Harrison, Goldstein, & Wyke, 1993; Cytowic & Eagleman, 2009; Walsh, 1996), or tastes (Downey, 1911). Other

*Correspondence should be addressed to David M. Eagleman, Baylor College of Medicine, 1 Baylor Plaza, Houston, TX 77030, USA (e-mail: david@eaglemanlab.net).

DOI:10.1111/j.1748-6653.2011.02015.x

synaesthetes experience tastes in response to words (Jones *et al.*, 2011; Simner & Ward, 2006; Ward & Simner, 2003), tactile sensations in response to tastes (Cytowic, 2002), spatial locations associated with numberlines, weekdays, and years (Eagleman, 2009; Hubbard, Ranzini, Piazza, & Dehaene, 2009; Tang, Ward, Butterworth, 2008), or dozens of other forms (Cytowic & Eagleman, 2009; Day, 2005). Through random population studies, the prevalence of some synaesthesia (in any form) has been estimated at 1–4% of the population (Simner *et al.*, 2006).

Synaesthesia debuted in the scientific literature over a century ago (Baratoux, 1887; Calkins, 1895; Clavière, 1898; De Rochas, 1885; Giradeau, 1885; Krohn, 1892; Philippe, 1893; Smith, 1905; Suarez de Mendoza, 1890); however, because of a lack of tests to verify the phenomenon, its study did not gain momentum until the development of tests for verifying and quantifying the phenomenon (Asher *et al.*, 2006; Baron-Cohen *et al.*, 1996; Eagleman, Kagan, Nelson, Sagaram, & Sarma, 2007). Many forms of synaesthesia can now be rigorously phenotyped in a high-throughput manner via automated diagnostic tests developed in the Synaesthesia Battery (www.synaesthete.org; Eagleman *et al.*, 2007). Such tests pivot on the consistency of synaesthetic associations recorded over multiple trials (Baron-Cohen, Harrison, Goldstein, & Wyke, 1993; Eagleman *et al.*, 2007; Ward & Sumner, 2005).

Two main problems have hampered synaesthesia research. First, the majority of the literature on synaesthesia has involved small subject populations (Amin *et al.*, 2011; Barnet *et al.* 2008; Baron-Cohen, Wyke, & Binnie, 1987; Hubbard, Arman, Ramachandran, & Boyman, 2005; Nunn *et al.*, 2002; Simner & Ward, 2006; Simner *et al.* 2006; Ward & Simner, 2003; Ward, Tsakanikos, & Bray 2006; Witthoft & Winawer, 2006). As a result, it is not always clear how the findings generalize to the larger population of synaesthetes with the same subtype. Second, almost every synaesthesia research paper in the literature studies a single, particular form (e.g., letter-colour, or sound-taste, but not both), as opposed to several forms and how they correlate. A pervasive and untested assumption is that the conclusions drawn in one study will apply to the condition of synaesthesia more broadly. However, no extant evidence allows us to assume that the variety of different forms of synaesthesia all represent the same underlying neural phenomenon.

The diverse range of synaesthetic experiences presents one of the most fascinating, yet puzzling, aspects of synaesthesia, and it raises a question that has so far gone unasked: is it possible that the term 'synaesthesia' encompasses more than one type of neural phenomenon, all of which lead to a similar phenomenological endpoint of perceptual crosstalk? For example, this could result from different genetic bases, or differentially localized expression of the same genes. By way of analogy, researchers searching for 'the gene' for deafness have so far implicated over 200 genetic loci (Willems, 2004), unsurprising in light of the fact that many genetic changes can interfere with the complex machinery of audition. Similarly, schizophrenia (Andreasen, 2000) and autism (Volkmar, Lord, Bailey, Schultz, & Klin, 2004) are now generally considered terms that encompass not just one, but a variety of syndromes. Is it possible, then, that 'synaesthesia' is an umbrella term that embraces multiple underlying phenomena?

We reason that if the term synaesthesia represents different mechanisms with a common phenotype, we might be able to tease the mechanisms apart—and thereby shed light on the underlying biology—by investigating which forms tend to co-occur. For example, if a synaesthete has coloured weekdays, is she more likely to have coloured numbers than would be expected by chance occurrence? And would her coloured

Table 1. Types of synaesthesia asked about and/or tested on the Synaesthesia Battery (www.synaesthete.org; Eagleman *et al.*, 2007)

Chords→Colour
Emotion→Colour
Instrument→Colour
Letters→Colour
Months→Colour
Music Pitch →Colour
Numbers→Colour
Orgasm→Colour
Pain→Colour
Personality→Colour
Smell→Colour
Sound→Smell
Sound→Taste
Sound→Touch
Spatialized sequences
Taste→Colour
Temperature→Colour
Touch→Colour
Vision→Smell
Vision→Sound
Vision→Taste
Weekdays→Colour

weekdays give her a tendency to have touch–smell synaesthesia, or would possession of one be uncorrelated with the other?

To determine whether different synaesthesia forms cluster into groups, we investigate tests and reports from 19,133 self-reported synaesthetic participants whose data were collected over a period of 3.5 years. The aim of this analysis is to focus and constrain models of the neurobiology of synaesthetic experience.

Methods

Data source

Data were drawn from the online Synaesthesia Battery (Eagleman *et al.*, 2007). Participants were presented with a list of 22 synaesthesia types (Table 1) and asked which types they experienced.

If participants indicated types that could be tested on the battery, they were automatically routed to those tests. Currently, testable types in the Synaesthesia Battery include Numbers-Colour, Letters-Colour, Weekdays-Colour, Months-Colour, Music Pitch-Colour, Chords-Colour, and Instrument-Colour. For each of these tests, subjects chose their synaesthetic colours for each stimulus using a computerized colour palette (Eagleman *et al.*, 2007). RGB coordinates for the colour results were converted into CIE-Lab colour space for analysis. CIE-Lab is a perceptually uniform colour space in which a change in colour value in CIE units produces an equivalent perceptual change. In this three-dimensional space, 'L' represents lightness, 'a' represents the balance between red (positive) versus green (negative), and 'b' represents yellow (positive) versus blue (negative) (Hunter, 1948).

Table 2. Frequencies of occurrence for each type of synaesthesia (sorted by the Inclusive set) as asked about and/or tested on the Synaesthesia Battery (www.synaesthete.org; Eagleman *et al.*, 2007)

	Inclusive (n = 19,133)	Stringent (n = 12,127)
Numbers→Colour	0.46	0.39
Weekdays→Colour	0.42	0.37
Months→Colour	0.41	0.34
Letters→Colour	0.36	0.29
Emotion→Colour	0.34	0.27
Personalities→Colour	0.34	0.27
Spatial	0.31	0.32
Pain→Colour	0.23	0.18
Instrument→Colour	0.18	0.05
Smell→Colour	0.18	0.13
Taste→Colour	0.18	0.13
Chords→Colour	0.16	0.01
Music Pitch→Colour	0.16	0.03
Temperature→Colour	0.16	0.11
Orgasm→Colour	0.13	0.11
Vision→Sound	0.13	0.09
Sound→Touch	0.12	0.10
Vision→Smell	0.11	0.09
Vision→Taste	0.11	0.09
Touch→Colour	0.09	0.06
Sound→Taste	0.08	0.07
Sound→Smell	0.07	0.05

For our analysis, we used two sets of data, one with a loose inclusion threshold ('Inclusive', n = 19,133 participants) and one with a more draconian threshold ('Stringent', n = 12,127 participants). The frequencies for each of the synaesthesia types are described in Table 2; both sets yielded similar results. We describe the criteria used in determining these sets now.

Data set 1: Inclusive (n = 19133)

Tens of thousands of people have come through the Synaesthesia Battery website, but not all complete the registration, the questionnaire, and any appropriate tests available for their forms, a process that takes from 5 min to over an hour (depending on the number of claimed synaesthetic forms). Not everyone who finishes this process is truly a synaesthete; however, the time to complete the process reduces the enthusiasm for most malingerers, thus reducing data-vandalism. This Inclusive data set idealistically errs on the side of including everyone who claimed to have any forms of synaesthesia in Table 1 and who appropriately completed the Battery (i.e., filled out all required fields in the questionnaire and completed all tests relevant to the types of synaesthesia that they claimed to possess). We disqualified any form of synaesthesia indicated by a participant if their results on the corresponding tests (a) revealed that answers were submitted in less than 1 second for more than two trials, which we took to belie random clicking, (b) the chosen colours had insufficient variability, as determined by the average colour distance between all the subjects' choices across the test (e.g., if all letters were assigned red, the test is disqualified), or (c) the participant assigned colours to fewer than half of the stimuli (e.g., if 12 letters were assigned colours and the remainder were labelled

as 'no colour'). Participants with more than two disqualified claims were dropped from the data set (e.g., if they claimed number- and letter-colour synaesthesia but failed both tests, then the rest of their claims were assumed false).

The broad inclusion in this data set is unsatisfactorily optimistic, but the inclusivity is demanded by the fact that we simply cannot test most of the forms of synaesthesia that we ask participants about (e.g., Pain→Colour). Therefore, if we want to elucidate at a population level which forms tend to co-occur, our best strategy is to gather unprecedented numbers of participants who claim to have synaesthesia and leverage the large numbers to magnify the signal in the noise. As will be seen in the next section, we will use standard statistical practices to extract the co-occurrences of forms, and the resulting clusters are statistically significant. If our data were overly polluted by malingerers, we would not expect to find statistically significant results.

Data set 2: Stringent (n = 12127)

To further enhance the signal in our data, we repeated our analysis with a more Stringent set of exclusion criteria, this time only including the 12,127 participants in whose responses we were the most confident. From the 19,133 in the previous data set, we first discredited any forms indicated by a subject if they failed the corresponding test (passing the threshold for a test is explained below). For example, if a participant checked the boxes for Chords-Colour, Instruments-Colour, and Letters-Colour, she would then undergo the consistency testing (Eagleman *et al.*, 2007) for all three of those forms. If, for example, she passed the threshold for the first two forms, but failed to meet the threshold for the last one, we would 'uncheck' that form for her and proceed with the assumption that she only had two valid forms of synaesthesia. Participants with more than two failed tests, or at least one failed test and no passed tests, were dropped entirely from the Stringent data set.

One of the defining characteristics of synaesthesia is that associations are consistent in repeated testing. The threshold for qualifying as synaesthetic was determined as follows: within each test, participants chose their synaesthetic colours for each stimulus presented three times in a random order. These colours were chosen from a colour palette of 16 million possible colours. For each stimulus, therefore, three colours were chosen; we converted these colour values from RGB to CIE-Lab space and calculated the distance between them to quantify the degree of consistency (Eagleman *et al.*, 2007). Even controls asked to fake the test for letter-colour synaesthesia, for example, have a difficult time remembering that 'F' is purple after 34 trials. Over the set of stimuli, we took the average colour distances—reflecting the average variability in chosen colours—and averaged these for a total score on each test. Over the population of participants who completed the tests, we computed a histogram of the average colour variability and fit a Gaussian distribution, working under the assumption that the low-variance scores represent true synaesthetes and the higher-variance scores represent non-synaesthetes. We took the mean of the Gaussian fit and set our threshold for inclusion at 1 standard deviation greater than that mean. For the Numbers-Colour, Letters-Colour, Weekdays-Colour, and Months-Colour tests, this yielded a threshold of approximately 70 (in arbitrary units of colour distance). For the Music Pitch-Colour, Chords-Colour, and Instrument-Colour tests, the distribution of scores yielded a slightly looser threshold of approximately 100.

Both of these data sets are imperfect in their inclusion criteria; the purpose of using both data sets is to see if the results differ between the most lenient and stringent criteria

for synaesthetic responses. We suggest that an unbiased population of synaesthetes might fall somewhere in between the two sets. As will be shown below, both data sets yielded similar and statistically significant conclusions, suggesting that our strategy of collecting data from thousands of self-reported synaesthetes was sufficient to detect patterns existing at the population level.

Exploratory and confirmatory factor analysis (EFA and CFA)

To determine a set of constructs underlying different synaesthesia types that account for their measured correlations, we conducted exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) on both data sets using the statistical software R (Steiger, 2009) in conjunction with the package SEM (Fox, Kramer, & Friendly, 2010). EFA is used to identify the number of underlying constructs, called factors, influencing the correlations among a set of observed variables. CFA tests whether a specified model of factors accounts for the observed variables in the predicted way using new data points (Decoster, 1998).

For each data set, we conducted EFA on a randomly selected half of the participants in our sample. In making our decisions, we consulted Osborne and Costello's recommendations for best practices in EFA (Osborne & Costello, 2005). For a constrained set of factors, loadings were extracted using an optimization process to minimize the maximum likelihood (ML) discrepency function (Steiger, 2009). EFA was run using both varimax (orthogonal) and promax (oblique) rotations of the loadings.

For a given number of factors used in EFA and the rotation type, a corresponding model was generated for the CFA, where each variable (synaesthesia type) was assigned to the factor with the corresponding strongest load. In the case of the varimax-rotated loadings from the EFA, underlying factors are not allowed to co-vary in our model, whereas they may in promax-rotated outputs (in which case we also analyze the resulting factor correlations). CFA was then used on the second half of the sample for each data set to determine the best-fitting model. The CFA model fit was evaluated on the basis of three goodness-to-fit indices: The Bayesian information criterion (BIC, minimum value), root mean squared error of approximation (RMSEA, values $\leq$0.08 though preferably values $\leq$ 0.06) (Browne & Cudeck, 1993; Hu & Bentler, 1999; Yu, 2002), and the standardized root mean residual (SRMR, values $\leq$0.08 though preferably values $\leq$0.05) (Hu & Bentler, 1999). In the main text, we present results using the BIC and RMSEA measure; we include SRMR measures in supplementary material for completeness.

Multidimensional scaling

Multidimensional scaling (MDS) is a set of techniques that displays the relative proximities between items in a visual configuration. This technique aims to find a comprehensive classification of the data, highlighting their differences as well as underlying dimensions. Each object is represented by a point, which is placed in multidimensional space so that the distances between two points best characterize the differences between the objects they represent. The greater the perceived differences between the objects, the greater the distance between them in the spatial configuration. We used Matlab's non-classical multidimensional scaling algorithm, which minimizes Kruskal's Stress-1 metric according to the normalized sum of the squares of inter-point distances (zero stress means a perfect fit, values $\leq$0.1 are considered a good fit, and values $\leq$0.15 are considered acceptable)

(Borgatti, 1997). We applied MDS to organize synaesthetic types into a space where the proximity between two types reflects their correlations.

Latent class analysis (LCA)

Latent class analysis (LCA) is a statistical method that organizes individuals with characteristics or symptoms into distinct classes (McCutcheon, 1987; Uebersax, 2009). Unlike factor loadings associated with each variable (synaesthesia type) in factor analysis, LCA provides a probability that an individual belongs to a class given that she has a particular pattern of synaesthesias.

LCA was performed using the random LCA package for R (Beath, 2011). Note that LCA assumes 'local independence', which requires that variables within a given class are independent of one another (which is generally not the case). Resultingly, if synaesthesia types have an underlying correlation structure (as we will demonstrate), LCA will yield ambiguous results because it is a suboptimal analysis tool in this condition. Indeed, that is what transpired here; nonetheless, we include the LCA results in Supporting information for completeness.

Results

Using the Stringent data set (n = 12,127), we first constructed a correlation matrix between the 22 different types of synaesthesia. We then excluded any elements with correlation <0.4 and permuted the matrix using the symmetric reverse Cuthill-McKee algorithm, which pushes larger elements of the matrix as close to the diagonal as possible. A thresholding of this matrix at 0.4 indicates the groupings of synaesthesia types that are moderately to strongly correlated with one another (Figure 1). It is apparent by eye that there are five distinct clusters: coloured sequences, coloured music, non-visual sequelae, coloured sensations, and spatial sequence synaesthesia (SSS) (which do not correlate with other types). Performing the same analysis on the Inclusive set yielded the identical result, but with slightly more noise (data not shown).

The separation into five clusters is further highlighted by MDS (Figure 2), which again reveals a clear clustering of synaesthesia types. The MDS on these data yielded a Stress-1 of 0.1102, which is considered to be an acceptable fit bordering a good fit (Borgatti, 1997).

Exploratory and confirmatory factor analysis

The above analyses indicate five groupings of synaesthesia types. We next turned to factor analysis to better understand these results. Factor analysis of both the Inclusive and Stringent sets using varimax-rotated factor loadings yielded the emergence of four factors with passable fits (Table 3).

The corresponding loadings (Table 4) indicate relatively strong and clear assignments of variables to factors for both sets, and are consistent across sets.

Table 4 demonstrates that SSS did not load particularly strongly on any synaesthesia type, as often happens when a factor only contains a single variable (Osborne & Costello, 2005; Tabachnick & Fidell, 2001). In other words, SSS appears to be a fifth and independent type of synaesthesia, as indicated in Figures 1 and 2, and demonstrated further below.

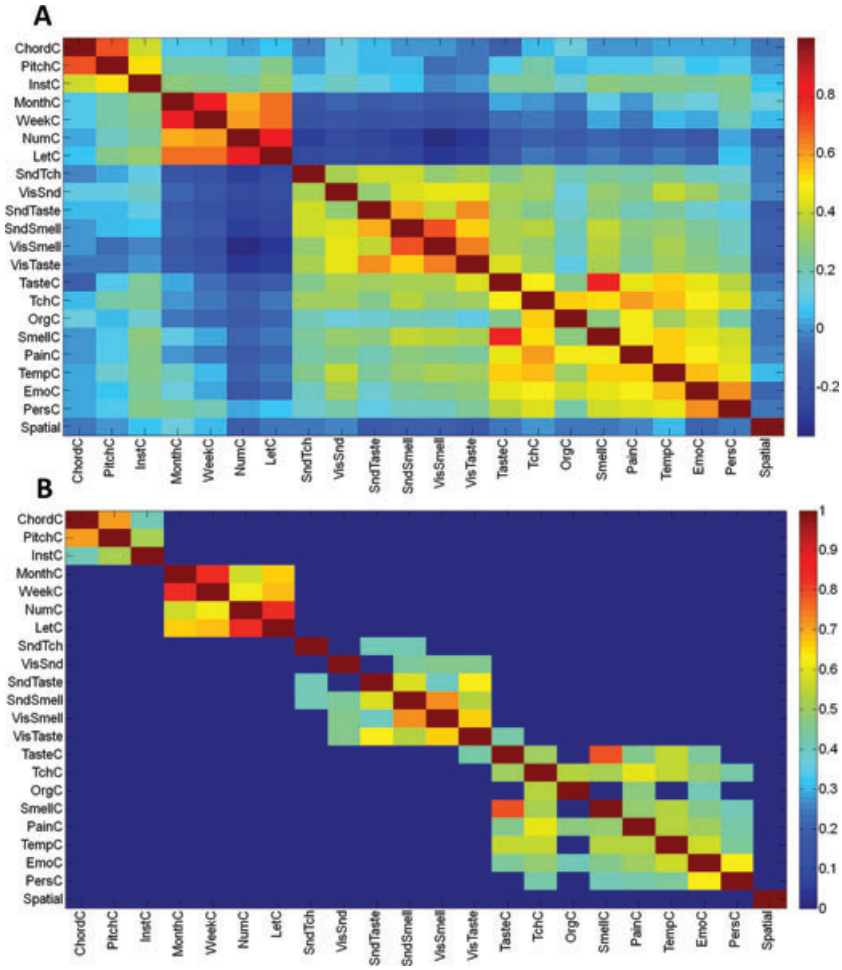

Figure 1. Correlations between types of synaesthesias reported by participants in the Stringent set (n = 12,127). (a) Sorted correlation matrix (b) Same data thresholded at 0.4 to clarify the structure. The character 'C' appended to types means 'Colour'; 'Num' means 'Numbers'; 'Let' means 'Letters'; 'Instr' means 'Instrument' (Timbre); 'Org' means 'Orgasm'; 'Emo' means 'Emotion'; 'Pers' means 'Personality'; 'Temp' means 'Temperature'; 'Vis' means 'Vision'; 'Snd' means 'Sound'; 'Tch' means 'Touch'.

Given the convergent results from the correlation analysis (Figure 1), MDS (Figure 2), and the factor analysis (Tables 3 and 4), we can now refer to these groupings in the following manner:

- Coloured Sequences: Colour experience triggered by overlearned, ordinal sequences (such as days of the week → colour or letters → colour).
- Musical Colours: Colour triggered by stimuli that are musical in nature, such as a pitch, a chord, or the timbre of an instrument.
- Coloured Sensation Synaesthesias: Colour experience triggered by physical touch or emotional sensation (such as pain → colour or emotion → colour).
- Non-Visual Sequelae: Synaesthesias in which a non-visual sense (such as smell, sound, touch, or taste) is triggered (such as sound → taste or vision → smell).

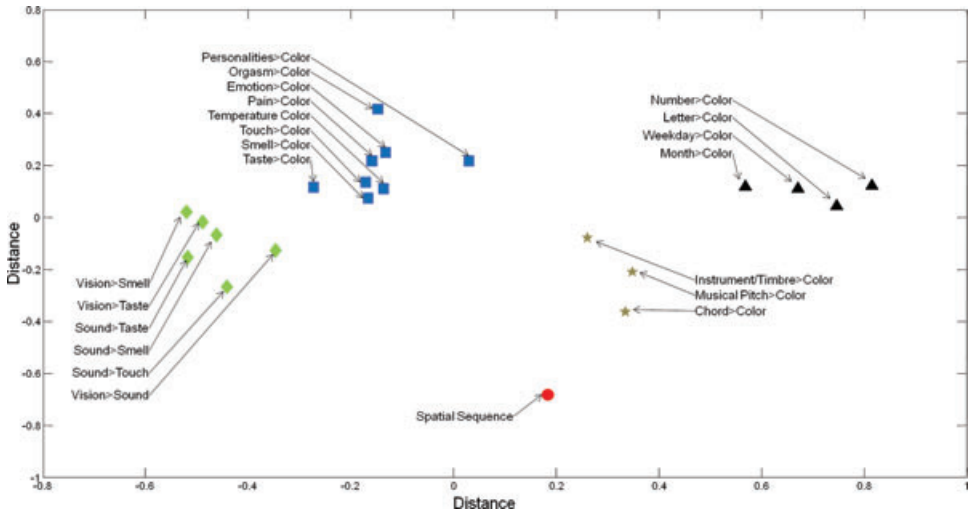

Figure 2. Results of two dimensional multidimensional scaling on the Stringent data set (n = 12,127). The distance matrix was defined by distance (i, j) = 1-correlation (i, j).

- Spatial sequence synaesthesia: the perception of overlearned sequences as possessing objecthood (e.g., perceiving a relative spatial location for each weekday).

Independence between groups

The previous analyses suggest independence between groups. To further confirm this hypothesis, we present a closer look at two of the groups, SSS and CSS.

In SSS, overlearned ordinal sets such as numbers, letters, weekdays, or months are experienced as having distinct spatial configurations (Eagleman, 2009; Galton, 1880; Hubbard, Ranzini, Piazza, & Dehaene, 2009; Sagiv, Simner, Collins, Butterworth, & Ward, 2006). As seen in the factor analysis, SSS did not have loadings greater than 0.13 on any factor, indicating that the variable shares little in common with other measured variables in the domain (Osborne & Costello, 2005; Tabachnick & Fidell, 2001). As with Figures 1 and 2, this confirms that SSS is a type statistically independent of the others. To confirm

Table 3. CFA Results from varimax-rotated models: RMSEA and BIC scores as factors used in EFA increase. As our method for assigning factors to variables (synaesthesia types) based on EFA can sometimes yield factors without an assignment, 'Factors Used for CFA' indicates the actual number of factors used in the model for CFA. Values in italics indicate the best obtained scores.

Inclusive set (n = 19,133)								
Factors Used for CFA	1	2	3	*4*	5	6	7	6
RMSEA	0.113	0.085	0.067	*0.065*	0.077	0.083	0.076	0.072
BIC	1,468	283	−308	*−356*	−7	195	−21	−149
Stringent Set (n = 12,127)								
Factors Used for CFA	1	2	3	*4*	5	6	7	8
RMSEA	0.095	0.062	0.059	*0.051*	0.052	0.055	0.063	0.067
BIC	679	−444	−503	*−682*	−679	−597	−412	−301

Table 4. Corresponding varimax-rotated factor loadings from EFA for the best CFA model based on RMSEA and BIC measures for both Inclusive and Stringent data sets, not including spatial-sequence synaesthetes. The factor analysis was also performed using promax-rotated loadings, as well as using another measure (SRMR) instead of RMSEA and BIC. All these analyses yielded roughly consistent results, and are included in Supplementary Material.

	Inclusive set (n = 19,133)				Stringent set (n = 12,127)			
	F1	F2	F3	F4	F1	F2	F3	F4
Spatial sequence	0.128					−0.129		
Taste→Colour	*0.604*	0.459		0.139	*0.711*	0.266		
Smell→Colour	*0.603*	0.47		0.181	*0.692*	0.299		
Pain→Colour	*0.688*	0.243		0.162	*0.697*	0.105		
Personalities→Colour	*0.62*	0.202	0.205	0.159	*0.608*		0.186	
Touch→Colour	*0.654*	0.38		0.194	*0.684*	0.186		0.16
Temperature→Colour	*0.685*	0.329		0.146	*0.699*	0.243	0.101	
Orgasm→Colour	*0.528*	0.179		0.137	*0.491*			0.162
Emotion→Colour	*0.72*	0.236		0.195	*0.709*	0.148		
Sound→Smell	0.274	*0.803*		0.111	0.27	*0.788*		
Vision→Smell	0.273	*0.805*			0.248	*0.798*	−0.107	
Vision→Sound	0.35	*0.552*		0.204	0.339	*0.503*	−0.105	0.11
Sound→Touch	0.287	*0.521*		0.107	0.231	*0.436*	−0.175	
Sound→Taste	0.237	*0.696*		0.108	0.205	*0.613*	−0.159	
Vision→Taste	0.288	*0.764*			0.311	*0.72*		
Weekdays→Colour	0.184		*0.755*			−0.132	*0.881*	
Months→Colour	0.236		*0.7*		0.121		*0.882*	
Numbers→Colour		−0.155	*0.898*			−0.313	*0.665*	0.124
Letters→Colour			*0.943*			−0.279	*0.735*	0.181
Instrument→Colour	0.165	0.108		*0.93*	0.37		0.281	*0.514*
Chords→Colour	0.252	0.125		*0.773*		0.117		*0.806*
MusicPitch→Colour	0.144	0.151		*0.871*	0.152		0.177	*0.893*

this finding, we next examined the overall probability of each synaesthesia type within our population (Figure 3, dark bars), and compared this to the probability of having each type if a participant also has SSS (Figure 3, light bars). As seen in the figure, there was no significant difference between these probability distributions ($p = 0.7974$, Kolmogorov–Smirnov goodness-of-fit, spatial bar excluded), meaning that possessing SSS makes a participant no more or less likely to have any other type. These findings are consistent with the isolation of SSS in Figures 1 and 2 and in our factor analysis. Collectively, these results reveal that SSS is an independent type of synaesthesia: it does not preferentially cluster with any other type.

Similarly, we turn to the cluster we have labelled 'coloured sequence synaesthesia' (CSS), a group comprising Weekdays-Colour, Months-Colour, Numbers-Colour, and Letters-Colour. We separated out all participants who possessed any one of these types, and computed the probabilities that they possessed any other type (Figure 4). The grouping of these types is readily apparent as supported by our previous analyses (Figures 1 and 2, Tables 3 and 4). For example, in Figure 4A, a participant with coloured weekdays is 72.40% likely to also have coloured months, 66.06% likely to have coloured numbers, and 56.50% likely to have coloured letters. However, such a synaesthete is no more likely than chance to possess other types (e.g., Smell-Colour). A comparison of the probability

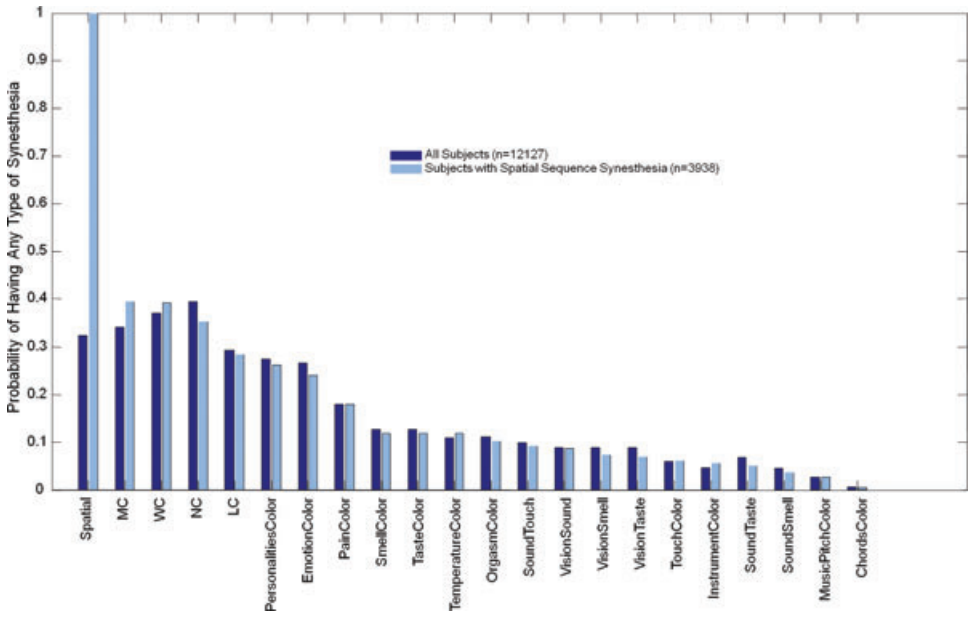

Figure 3. Probability of having a certain type in two separate populations within the Stringent data set. The dark bars show the overall probabilities (n = 12,127) of having any type. The light bars show the probabilities of spatial synaesthetes (n = 3,938) having other types. Comparison of both distributions showed no significant difference ($p = 0.7974$, Kolmogorov–Smirnov goodness-of-fit, with the Spatial category excluded from the analysis), indicating that possession of SSS is orthogonal to the possession of other types.

distributions represented by the dark and light bars (Figure 4A–D) shows a significant difference ($p < 0.002$, Kolmogorov–Smirnov test); however, when we exclude the four CSS types, there is no longer a significant difference between the distributions ($p = 0.4$). These results produce two conclusions: possessing any one of the four CSS types (1) affects the probability of having another CSS type, and (2) does not alter the probability of having any other type of synaesthesia. This finding is consistent with expectations from the groupings revealed by the factor analysis (Table 4), the MDS (Figure 2) and correlation matrix analysis (Figure 1).

Substructure within groups

We performed further factor analysis with different choices of rotations (promax instead of varimax), different model measures (SRMR instead of RMSEA/BIC), and analyses of the Inclusive data set as well as the Stringent data set (Supporting information). These all yielded roughly similar results, but with the interesting result that the different approaches drew out what appear to be subtle substructures in the groups (Tables S4 and S6). Specifically, for the Inclusive set in the BIC-case (Table S4) one can see a split in the colour sequence synaesthesias between weekdays and months, on the one hand, and letters and numbers on the other. (The two subgroups still have a strong correlation (0.683) between them, meaning they are best viewed as subgroups rather than separate grouping). In the Stringent case (Table S6), coloured sensation synaesthesia splits coloured pain, touch, and orgasm away from coloured emotion, personality, and temperature. We tentatively label the first group 'coloured sensation-physical', and the

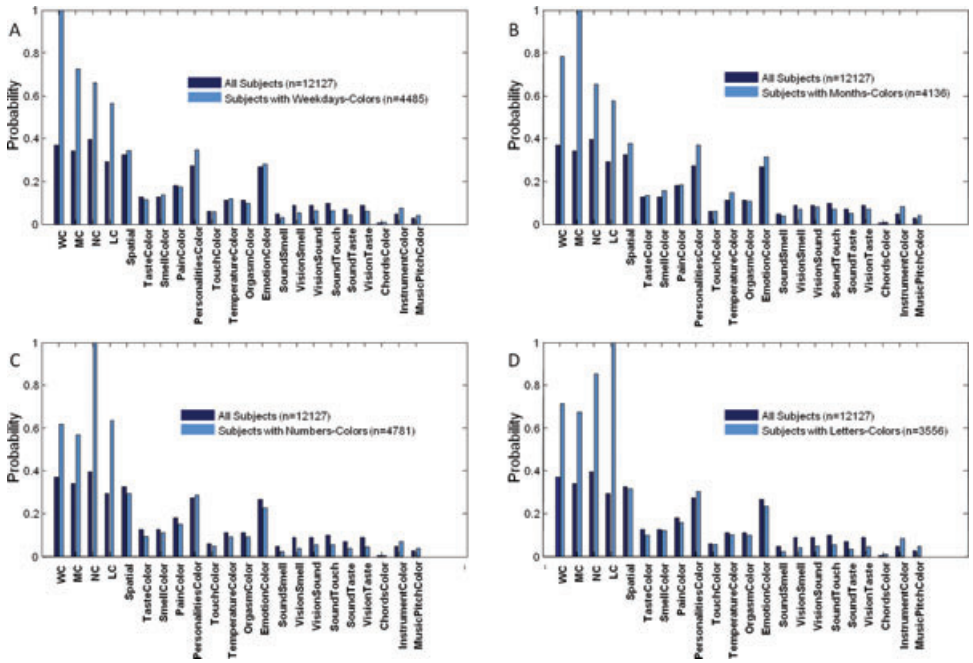

Figure 4. Possessing any type of coloured sequence (letter, number, weekday, month) gives a high probability of possessing other types of coloured sequences. (A–D) Probabilities of having any type given a particular CSS (light bars), and the overall probabilities of having any type (dark bars). Comparison of both distributions excluding the CSS types were found to be insignificant ($p = 0.4$, Kolmogorov–Smirnov test), meaning that the distribution of other types is independent of CSS types.

second 'coloured sensation-conceptual'. Again, a strong correlation between the factors (0.8) means they are subgroups rather than truly independent clusters. Further, both sets described above indicate a split of Taste→Colour and Smell→Colour synaesthesias away from the rest of the coloured sensation group, with a moderately strong correlation of 0.575 between these cases. We label these 'coloured flavour'.

The substructures pulled out of the promax-rotated loadings can also be seen upon closer inspection. In the case of CSS splitting, there are much tighter correlations between numbers with letters and weekdays with months, and this can also be seen by proximity in the MDS plot. For the split that occurs between taste→colour with smell→colour from the other coloured sensation synaesthesias, the correlation matrix clearly indicates that these are tightly coupled (more so than either of these with any other synaesthesia type). This is not as readily discernable in the MDS plot, however, despite being in the same general region. This is most likely due to the fact that MDS does not provide a perfect mapping of higher dimensional data.

This set of results is consistent with the correlation visualizations above in Figures 1 and 2, revealing five distinct groupings of synaesthesia types, with some amount of substructure within the groups. Our final model for the structure of the data is summarized visually in Figure 5, along with the probability in the Stringent data set of possessing one or more synaesthesias in that group. Our overall findings are generally consistent with the hypothesis that each of these synaesthesia groups can be 'turned on' with independent probability in any individual subject.

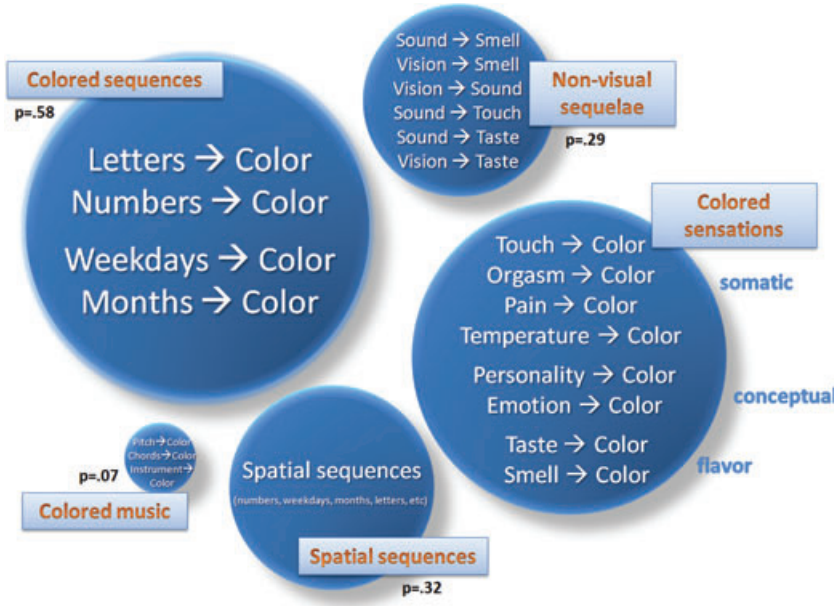

Figure 5. Schematic of the model: five distinct groups of synaesthesia. The radius of each type is proportional to the probability of independently expressing that type (computed from the Stringent data set, n = 12,127).

Discussion

Using several converging analyses on data from thousands of participants in the Synaesthesia Battery, we have found that synaesthesia types cluster into distinct groups. Our findings reveal that a person possessing more than one type of synaesthesia is more likely to have types that belong in the same cluster, rather than different clusters. The different clusters appear to be roughly statistically independent of one another, like independent 'modules' that can be on or off with independent probabilities. We here summarize the main findings.

First, coloured letters, numbers, weekdays, and months are highly likely to co-occur (Figures 1,2, and 4, Tables 1 and 2), but with other types of synaesthesia they co-occur no more than predicted by chance (Figure 4). Given the overlearned and ordinal nature of these stimuli, we group these types under the label 'coloured sequence synaesthesia (CSS)'. This cluster appears to contain a deeper substructure in which coloured letters and numbers are slightly more likely to co-occur as a pair, as are coloured weekdays and months (Table S4).

Second, the triggering of colour by musical pitches, chords, and timbres also cluster; we collectively label these 'coloured music synaesthesias'.

A third cluster is characterized by the triggering of non-visual experiences such as smell, taste, sound, and touch (Figures 1 and 2). We have labelled these 'non-visual sequelae'.

Fourth, several types of colour synaesthesias are triggered by stimuli that involve sensation: these include triggers such as personality and emotion (conceptual), as well as touch, pain, and the experience of orgasm (physical). Furthermore, colour experiences triggered by smell and taste fall within this same cluster (Figures 1 and 2), but form a

recognizable subgroup. Given the close relationship between the two chemical senses in composing the perception of flavour (Cytowic & Eagleman, 2009, Chapter 6), we subcluster these latter two types under the label 'coloured flavour synaesthesia'.

Finally, the experience of spatial localization for overlearned sequences appears to be independent of all other types; possessing it does not influence the probabilities of having any other type (Figure3). As we have done previously, we label this type 'spatial sequence synaesthesia (SSS)' (Eagleman, 2009).

Our factor analysis (Tables 1 and 2), MDS (Figure 2), and analysis of joint probabilities (Figures 3 and 4) jointly corroborate the categories we have proposed above.

Coloured sequence synaesthesia

With regard to the cluster of coloured sequence synaesthesia in particular, Rich *et al.* (2005) has previously noted that letters, numbers, weekdays, and months tended to co-occur, which they labelled 'linguistic-colour' synaesthesia. We prefer the term coloured sequence to emphasize that the synaesthesia is not triggered by language in general, but rather by overlearned sequences in particular (Cytowic & Eagleman, 2009; Eagleman, 2009). This feature suggests the possibility of a separate neural basis for *ordinal stimuli* triggering colour perception. A candidate network of areas involved in overlearned ordinal stimuli has recently been identified with neuroimaging, consisting of the middle temporal gyrus and temporoparietal junction in the right hemisphere, and the inferior frontal gyrus in the left hemisphere (Pariyadath,Churchill, & Eagleman, 2009). Studies are underway in our laboratory to determine whether activity in this network directly correlates with increased activity in visual colour processing areas. In this framework, the proposed neural basis of CSS is increased crosstalk between brain regions involved in overlearned sequences and colour regions, while other clusters would presumably be explained by quite different neural pathways.

Limitations of the analysis

Current technology only allows us to test and verify some of the synaesthesia forms claimed by each participant, but not all of them. This necessarily introduces noise into the data. Therefore, this study has the potential limitation that registrants on the Synaesthesia Battery may not have all been synaesthetic, but instead were testing themselves for the possibility, misunderstanding synaesthesia, or malingering. Because not all synaesthesia types can currently be verified by consistency testing, these results should be interpreted with appropriate caution.

However, there are several reasons to believe that the results nonetheless contain a sufficiently high signal-to-noise ratio to detect clear patterns in the data. First, the sample size of thousands of participants lends itself to washing out noise. Second, the testing process is lengthy (average time to complete is approximately 15 minutes), and this time investment generally diminishes the enthusiasm of malingerers; we only include batteries that were fully completed. Third, our clustering results are statistically significant—a feature that would not be expected to emerge from noise.

An additional concern is the possibility of response bias, as regards the untested types of synaesthesia (e.g. coloured sensation). It is possible that some people reported several together because they 'seemed' like they should go together. This possibly seems unlikely to explain the extremely high correlation found among thousands of

respondents; nonetheless, it can put directly to the test in the future as we develop offline tests for these other types.

Despite our efforts, there remain limitations inherent in data collected online, but we hope that the population reported here is sufficiently large to allow us to detect underlying structure in a previously undifferentiated population.

Independent neural bases?

The most plausible hypothesis for why synaesthetic brains differ from non-synaesthetic brains is that synaesthesia reflects an increased degree of crosstalk between normally separated brain areas, such that activity in one area kindles activity in another (Baron-Cohen, Harrison, Goldstein, & Wyke, 1993; Cytowic & Eagleman, 2009; Grossenbacher, 2001; Hubbard & Ramachandran, 2005; Hubbard, Arman, Ramachandran, & Boyman, 2005; Nunn *et al.*, 2002; Ramachandran and Hubbard, 2001; Rouw & Scholte, 2007; Weiss, Zilles, & Fink, 2005).

However, our findings here suggest that synaesthesia is not a single phenomenon. Instead, several independent neural mechanisms may share the common property of increased crosstalk.

Many studies have suggested that synaesthesia runs in families (Tomson *et al.*, 2011). Our findings may support a model of independent gene expression in which each synaesthesia grouping is caused by an independent genetic module. Alternatively, our findings could also represent mosaic expression of a single gene product rather than different genetic bases (i.e., expressed with independent probabilities in different brain regions in different individuals). Disambiguating these hypotheses remains open for future study. It should be noted that these possibilities are non-exclusive; a full accounting of the groupings may include examples of each of these. Studies of the patterns of inheritance in family trees (Barnett *et al.*, 2008) find that family trees appear to have representations from multiple clusters—for example, a set of twins in one family reports colored sequence synaesthesia, while the third sibling appears to also have taste-colour synaesthesia. However, these studies do not yet disambiguate the two hypotheses.

In conclusion, we have found with strong statistical significance that there is an underlying clustering of synaesthesia types: the distribution is highly non-random. Further, we find that individuals can span multiple clusters and that these clusters appear to be statistically independent of one another. This leads us to believe that these clusters function as independent modules. Having at least one type synaesthesia in a module implies a high likelihood of having another synaesthesia type from the same module. Conversely, our data indicate that possessing a synaesthesia type in one module has no bearing on the probability of having a synaesthesia type from another module. Further research is required to identify the underlying neural and genetic bases of these clusters. Current studies in our laboratory are examining this using both neuroimaging and family linkage analyses (Tomson *et al.*, 2011). Collectively, these studies will give a more rigorous understanding of the variety of phenomena that have traditionally fallen under the single umbrella term of 'synaesthesia'.

Acknowledgements

We thank the Guggenheim Foundation and Mind Science Foundation for grants to support this research. Thanks to Stephanie Tomson for feedback on the manuscript.

References

Amin, M., Olu-Lafe, O., Claessen, L. E., Sobczak-Edmans, M., Ward, J., Williams, A. L., & Sagiv, N. (2011). Understanding grapheme personification: A social synaesthesia? *Journal of Neuropsychology, 5*, 255–282. doi:10.1111/j.1748-6653.2011.02016.x

Andreasen, N. C. (2000). Schizophrenia: The fundamental questions. *Brain Research Brain Research Reviews, 31*, 106–112. doi:10.1016/S0165-0173(99)00027-2

Asher, J. E., Aitken, M. R., Farooqi, N., Kurmani, S., & Baron-Cohen, S. (2006). Diagnosing and phenotyping visual synaesthesia: A preliminary evaluation of the revised test of genuineness (TOG-R). *Cortex, 42*, 137–146. doi:10.1016/S0010-9452(08)70337-X

Baratoux, J. (1887). De l'audition colorée. *Le Progrès Médical, 6*, 495–496; 515–517; 538–539.

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 synaesthesia. *Cognition, 106*(2), 871–893. doi:10.1016/j.cognition.2007.05.003

Baron-Cohen, S., Burt, L., Smith-Laittan, F., Harrison, J., & Bolton, P. (1996). Synaesthesia: Prevalence and familiality. *Perception, 25*, 1073–1079. doi:10.1068/p251073

Baron-Cohen, S., Harrison, J., Goldstein, L. H., & Wyke, M. (1993). Coloured speech perception: Is synaesthesia what happens when modularity breaks down? *Perception, 22*, 419–426. doi:10.1068/p220419

Baron-Cohen, S., Wyke, M. A., & Binnie, C. (1987). Hearing words and seeing colours: An experimental investigation of a case of synaesthesia. *Perception, 16*, 761–767. doi:10.1068/p160761

Beath, K. (2011). *randomLCA: Random effects latent class analysis*. Retrieved from http://cran.r-project.org/web/packages/randomLCA/index.html

Borgatti, S. P. (1997). *Multidimensional Scaling*. Retrieved from http://www.analytictech.com/borgatti/mds.htm

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Calkins, M. W. (1895). Synaesthesia. *American Journal of Psychology, 7*, 90–107.

Clavière, J. (1898). L'audition colorée. *L'année Psychologique, 5*, 161–178.

Cytowic, R. E. (2002). *Synesthesia: A Union of the Senses*. Cambridge, MA: MIT Press.

Cytowic, R. E., & Eagleman, D. M. (2009). *Wednesday is indigo blue: Discovering the brain of synesthesia*. Cambridge, MA: MIT Press.

Day, S. (2005). Some demographic and socio-cultural aspects of synaesthesia. In Robertson LCS, N. (Ed) *Synesthesia: Perspectives from cognitive neuroscience.* Oxford: Oxford University Press.

DeCoster, J. (1998). *Overview of factor analysis*. Retrieved from http://www.stat-help.com/notes.html

DeRochas, A. (1885). L'audition colorée. *La Nature*, *620*, 306–307.

Downey, J. E. (1911). A case of colored gustation. *American Journal of Psychology, 22*, 528–539.

Eagleman, D. M. (2009). The objectification of overlearned sequences: A large-scale analysis of spatial sequence synesthesia. *Cortex*, *45*(*10*): 1266–1277. doi:10.1016/j.cortex.2009.06.012

Eagleman, D. M., & Goodale, M. A. (2009). Why color synesthesia involves more than color. *Trends in Cognitive Sciences, 13*(7), 288–292. doi:10.1016/j.tics.2009.03.009

Eagleman, D. M., Kagan, A. D., Nelson, S. S., Sagaram, D., & Sarma, A. K. (2007) A standardized test battery for the study of synesthesia. *Jounal of Neuroscience Methods, 159*, 139–145. doi:10.1016/j.jneumeth.2006.07.012

Fox, J., Kramer, A., & Friendly, M. (2010). *Structural equation models*. Retrieved from http://cran.r-project.org/web/packages/sem/index.html

Galton, F. (1880). Visualized numerals. *Nature, 21*, 252–256.

Giradeau, C. (1885) De l'audition colorée. *L'Encéphale, 5*, 589–597.

Grossenbacher, P. G., & Lovelace, C.T. (2001). Mechanisms of synesthesia: Cognitive and physiological constraints. *Trends in Cognitive Sciences, 5*(1), 36–41. doi:10.1016/S1364-6613(00)01571-0

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55. doi:10.1080/10705519909540118

Hubbard, E. M., Arman, A. C., Ramachandran, V. S., & Boynton, G. M. (2005). Individual differences among grapheme-color synesthetes: Brain-behavior correlations. *Neuron, 45*, 975–985. doi:10.1016/j.neuron.2005.02.008

Hubbard, E. M., & Ramachandran, V. S. (2005). Neurocognitive mechanisms of synaesthesia. *Neuron, 48*, 509–520. doi:10.1016/j.neuron.2005.10.012

Hubbard, E. M., Ranzini, M., Piazza, M., & Dehaene, S. (2009). What information is critical to elicit interference in number-form synesthesia? *Cortex*, *45*, 1200–1216. doi:10.1016/j.cortex.2009.06.011

Hunter, R. S. (1948). Photoelectric color-difference meter. *Journal of the Optical Society of America, 38*, 661. doi:10.1364/JOSA.48.000985

Jones, C. L., Gray, M. A., Minati, L., Simner, J., Critchley, H. D., & Ward, J. (2011). The neural basis of illusory gustatory sensations: Two rare cases of lexical–gustatory synaesthesia. *Journal of Neuropsychology*, *5*, 243–254. doi:10.1111/j.1748-6653.2011.02013.x

Krohn, W. D. (1892). Pseudo-chromesthesia, or the association of colors with words, letters, and sounds. *American Journal of Psychology, 5*, 20–41.

Kruskal, J. B., & Wish, M. (1977). *Multidimensional Scaling*. Beverly Hills, CA: Sage Publications.

McCutcheon, A. C. (1987). *Latent class analysis*. Beverly Hills: Sage Publications.

Nunn, J. A., Gregory L. J., Brammer, M., Williams, S. C., Parslow, D. M., Morgan, M. J. . . . Gray, J. A. (2002). Functional magnetic resonance imaging of synesthesia: Activation of V4/V8 by spoken words. *Nature Neuroscience, 5*, 371–375. doi:10.1038/nn818

Osborne, J., & Costello, A. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation, 10*(7). Retrieved from http://pareonline.net/pdf/v10n7.pdf

Palmeri, T. J., Blake, R., Marois, R., Flanery, M. A., & Whetsell, W., Jr. (2002). The perceptual reality of synesthetic colors. *Proceedings of the National Academy of Sciences USA, 99*, 4127–4131. doi:10.1073/pnas.022049399

Pariyadath, V., Churchill, S. J., & Eagleman, D. M. (2008). Why overlearned sequences are special: Distinct neural networks in the right hemisphere for ordinal sequences. *Nature Precedings*. Retrieved from http://hdl.handle.net/10101/npre.2008.2452.1

Philippe, J. (1893). Résumé d'une observation d'audition colorée. *Revue Philosophique, 36*, 330–334.

Ramachandran, V. S., & Hubbard, E. M. (2001). Psychophysical investigations into the neural basis of synaesthesia. *Proceedings of the Royal Society B: Biological Sciences, 268*, 979–983. doi:10.1098/rspb.2001.1576

Rich, A. N., Bradshaw, J. L., & Mattingley, J. B. (2005). A systematic, large-scale study of synaesthesia: Implications for the role of early experience in lexical-colour associations. *Cognition, 98*, 53–84. doi:10.1016/j.cognition.2004.11.003

Robertson, L. C., Sagiv, N., eds. (2004). *Synesthesia: Perspectives from cognitive neuroscience*. New York: Oxford University Press.

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

Sagiv, N., Simner, J., Collins, J., Butterworth, B., & Ward, J. (2006). What is the relationship between synaesthesia and visuo-spatial number forms? *Cognition, 101*, 114–128. doi:10.1016/j.cognition.2005.09.004

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

Simner, J., & Ward, J. (2006). Synaesthesia: the taste of words on the tip of the tongue. *Nature, 444*, 438. doi:10.1038/444438a

Smilek, D., Dixon, M. J., Cudahy, C., & Merikle, P. M. (2002). Synesthetic color experiences influence memory. *Psychological Science, 13*, 548–552. doi:10.1111/1467-9280.00496

Smith, H. L. (1905). Synaesthesia. *Bulletin of the Johns Hopkins Hospital, 16*, 258–263.

Steiger, J. H. (2009). *Confirmatory factor analysis with R*. Retrieved from http://statpower.net

Steiger, J. H. (2009). *Exploratory factor analysis with R*. Retrieved from http://statpower.net

Suarez de Mendoza, F. (1890). *L'audition colorée*. Paris: Octave Doin.
Tabachnick, B. G., & Fidell, L. S. (2001) *Using Multivariate Statistics*. Boston: Allyn and Bacon.
Tang, J., Ward, J., & Butterworth, B. (2008). Number forms in the brain. *Journal of Cognitive Neuroscience*, *20*(9), 1547–1556. doi:10.1162/jocn.2008.20120
Tomson, S., Avidan, N., Lee, K., Sarma, A. K., Tushe, R., Milewicz, D, . . . , Eagleman, D. M. (2011). The genetics of colored sequence synesthesia: Suggestive evidence of linkage to 16q and genetic heterogeneity for the condition. *Behavioral Brain Research*, *223*, 48–52. doi:10.1016/j.bbr.2011.03.071
Uebersax, J. (2009). *LCA Frequently Asked Questions*. Retrieved from http://john-uebersax.com/stat/faq.html
Volkmar, F. R., Lord, C., Bailey, A., Schultz, R. T., & Klin, A. (2004). Autism and pervasive developmental disorders. *Journal of Child Psychology and Psychiatry, 45*, 135–170.
Walsh, V. (1996). Perception: The seeing ear. *Current Biology, 6*, 389–391. doi:10.1016/S0960-9822(02)00503-1
Ward, J., & Simner, J. (2003). Lexical-gustatory synaesthesia: Linguistic and conceptual factors. *Cognition, 89*, 237–261. doi:10.1016/S0010-0277(03)00122-7
Ward J., & Simner, J. (2005). Is synaesthesia an X-linked dominant trait with lethality in males? *Perception, 34*, 611–623. doi:10.1068/p5250
Ward, J., Tsakanikos, E., & Bray, A. (2006). Synaesthesia for reading and playing musical notes. *Neurocase, 12*, 27–34. doi:10.1080/13554790500473672
Weiss, P. H., Zilles, K., & Fink, G. R. (2005). When visual perception causes feeling: Enhanced cross-modal processing in grapheme-color synesthesia. *Neuroimage, 28*, 859–868. doi:10.1016/j.neuroimage.2005.06.052
Willems, P. J. (2004). *Genetic hearing loss*. New York-Basel: Marcel Dekker, Inc.
Witthoft, N., & Winawer, J. (2006). Synesthetic colors determined by having colored refrigerator magnets in childhood. *Cortex, 42*, 175–183. doi:10.1016/S0010-9452(08)70342-3
Yu, C. Y. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes* [Dissertation]. Los Angeles, UCLA.

Received 13 September 2010; revised version received 12 April 2011

Supplementary Data

The following supporting information is available for this article:

Table S1. All varimax-rotated factor Analysis results for both Inclusive and Stringent data sets.
Table S2. Varimax-rotated factor loadings for Inclusive and Stringent sets based on the best fit using SRMR.
Table S3. Promax-Rotated Factor loadings and corresponding factor correlations for the Inclusive set (n=19,133) based on best RMSEA (0.05) and SRMR (0.044) scores (first model to cross 0.05 for RMSEA and SRMR).
Table S4. Promax-Rotated Factor loadings and corresponding factor correlations for the Inclusive set (n=19,133) based on best BIC score (minimum score found).
Table S5. Promax-Rotated Factor loadings and corresponding factor correlations for the Stringent set (n=12,127) based on best RMSEA (0.045) and SRMR (0.043) scores (first model to cross 0.05 for RMSEA and SRMR).
Table S6. Promax-Rotated Factor loadings and corresponding factor correlations for the Stringent set (n=12,127) based on best BIC score (minimum score found).
Figure S1. Latent Class Analysis (LCA) on the Inclusive set (n=19,133).

Figure S2. Conditional Probability Matrices for (a) the Inclusive set (n=19,133) and (b) the Stringent set (n=12,127).

Supporting Information may be found in the online version of this article.