A is for apple: the role of letter-word associations in the development of grapheme-colour synaesthesia

Jennifer Mankin, Julia Simner

Publication date
30-05-2017

Document Version
Accepted version

Citation for this work (American Psychological Association 7th edition)
Mankin, J., & Simner, J. (2017). *A is for apple: the role of letter-word associations in the development of grapheme-colour synaesthesia* (Version 1). University of Sussex.
https://hdl.handle.net/10779/uos.23444111.v1

Published in
Multisensory Research

Link to external publisher version
https://doi.org/10.1163/22134808-00002554

A is for apple: The role of letter-word associations in the development of grapheme-colour synaesthesia

Jennifer L. Mankin and Julia Simner

School of Psychology, University of Sussex

Correspondence should be addressed to Jennifer L. Mankin, School of Psychology, University of Sussex, Brighton. Email: J.Mankin@sussex.ac.uk

Keywords: grapheme-colour synaesthesia; literacy; alphabet acquisition; prototypical colour

Abstract

This study investigates the origins of specific letter-colour associations experienced by people with grapheme-colour synaesthesia. We present novel evidence that frequently observed trends in synaesthesia (e.g. *A* is typically red) can be tied to orthographic associations between letters and words (e.g., “A is for apple”), which are typically formed during literacy acquisition. In our experiments, we first tested members of the general population to show that certain words are consistently associated with letters of the alphabet (e.g. *A* is for *apple*), which we named *index words*. Sampling from the same population, we then elicited the typical colour associations of these index words (e.g. apples are red) and used the letter → index word → colour connections to predict which colours and letters would be paired together based on these orthographic-semantic influences. We then looked at direct letter-colour associations (e.g., *A* → red, *B* → blue…) from both synaesthetes and non-synaesthetes. In both populations, we show statistically that the colour predicted by index words matches significantly with the letter-colour mappings: that is, *A* → red because *A* is for *apple* and apples are prototypically red. We therefore conclude that letter-colour associations in both synaesthetes and non-synaesthetes are tied to early-learned letter-word associations.

Introduction

People with synaesthesia experience consistent and automatic quasi-perceptual experiences, such as experiencing taste or colour sensations when they hear words (Simner, 2012; Ward & Mattingley, 2006). The condition has enjoyed a recent surge of interest since its scientific “rediscovery” in the 1970s and 1980s (Cytowic, 1989; Cytowic & Wood, 1982; Marks, 1975). One idea that has gained traction is that experiences in synaesthesia often reflect intuitive, cross-modal associations common to synaesthetes and non-synaesthetes (Sagiv & Ward, 2006; Spector & Maurer, 2009; Ward, Huckstep, & Tsakanikos, 2006). Hence, for both synaesthetes experiencing synaesthesia, and non-synaesthetes making intuitive associations, brighter colours are associated with higher musical pitch (Ward et al., 2006), darker colours with rougher and harder surfaces (Simner & Ludwig, 2012; Ward, Banissy, & Jonas, 2008), and numbers with particular spatial locations (Jonas, Spiller, Jansari, & Ward, 2014). Studying synaesthesia can therefore elucidate universal cross-modal structures and cognitive processes. In the current study we look at similarities between synaesthetes and non-synaesthetes in the way they associate colours with graphemes (letters and numbers). We shall see that such associations are not random for either population, and can be predicted in part by linguistic influences (see also Mankin, in press; Simner, 2007) and in particular, by early-learned letter-to-word associations (e.g., *A* is for apple).

The current study focuses on grapheme-colour synaesthesia, a common variety of synaesthesia wherein graphemes (here, particularly letters) give rise to automatic associations with colours (e.g. *E* might be leaf green or *D* brown; Baron-Cohen, Burt, Smith-Laittan, Harrison, & Bolton, 1996; Simner, Glover, & Mowat, 2006; Ward, Simner, & Auyeung, 2005). This synaesthesia involves the cognitive processes involved in reading, which themselves involve a learned association between abstract symbols and sound or meaning. A common thread in large-scale investigations of synaesthesia is that synaesthetes and non-synaesthetes tend to agree on certain colour associations at above chance levels. For example, *A* tends to be red for both populations, *L* tends to be yellow, and so on. The largest studies showing these trends have been conducted in English (Jonas, 2010; Rich, Bradshaw, & Mattingley, 2005; Simner et al., 2005; Witthoft, Winawer, & Eagleman, 2015), although similar trends have been found in other languages such as Dutch and Hindi (Rouw, Case, Gosavi, & Ramachandran, 2014), Japanese (Nagai, Yokosawa, & Asano, 2015), German (Emrich, Schneider & Zedler, 2002; Simner et al., 2005) and Ukrainian (Lavrynenko, 2014). To explore how these colour-letter pairing trends are formed, we will first briefly review the previously identified sources of these trends. We will then investigate an as-yet-untested possibility: that the colours for letters may originate from early-acquired letter-to-word associations. We name this proposal, which has been raised previously but never tested, the ‘A is for apple’ hypothesis: simply put, *A* is red because *A* is for *apple* and apples are red.

We turn first to the colour-letter trends identified in synaesthetes. Three studies (Jonas, 2010; Rich et al., 2005; Simner et al., 2005) asked English-speaking synaesthetes for their letter-colour experiences and identified the colour that occurred at a higher-than-chance level for each letter. A fourth study (Witthoft et al., 2015) reported the most frequent (i.e. *modal*) colour choice for each letter from a large population of synaesthetes. Two of the four studies listed above (namely, Rich et al., 2005; Simner et al., 2005) also gave the same letter-colour association test to non-synaesthetes. Although a question such as "What colour is the letter A?" may seem nonsensical to non-synaesthetes, these subjects nonetheless showed agreement not only among themselves, but also with synaesthetes. The sources of some of these widespread associations were more obvious than others. There was a significant tendency for both synaesthetes and non-synaesthetes to associate a letter with the colour that begins with that letter: *R* with red, *Y* with yellow, *G* with green, *B* with blue, *V* with violet, and *P* with pink. Non-synaesthetes also strongly associated *W* with white and *O* with orange, while synaesthetes' associations here were not explicable by colour-name association: *O* with white, but *J* with orange. Furthermore, both groups showed strong shared associations across other letters as well: *A* with red, *D* with brown, *F* with green, *L* with yellow, *U* with grey, *X* and *Z* with black, and *I* with white and/or black. Disregarding for the moment the letter-colour pairs that are easily explicable by the initial letter of the colour name (e.g. red for *R*), how can we explain these trends across synaesthetic and non-synaesthetic populations?

First, there is some evidence that associations can be explicitly acquired from childhood toys or books featuring coloured letters. After a few synaesthetes reported letter-colour associations highly similar to coloured alphabet magnets (Witthoft & Winawer, 2006, 2013), Witthoft et al. (2015) found that in a large sample of 6,588 synaesthetes, 400 (about 6%) had 10 or more letter-colour associations that matched a well-known alphabet magnet set. Furthermore, just one in 150 synaesthetes showed similarities to childhood alphabet books in a study by Rich et al. (2005). A second possibility is that these common letter-colour pairings are indicative of more general associations, not specific to graphemes but to shapes and concepts. Pre-literate children consistently pair *X* with black and *O* with white, but show no inclination towards associating *A* with red and *G* with green, while literate children and adults do both (Spector & Maurer, 2008). Hence, some of grapheme-colour pairings may be based in literacy (e.g. *G* → green), while others may be naturally biased shape-colour pairings (e.g. *X* → black). This leads to the conclusion that some of these grapheme-colour pairings may be based in literacy (e.g. *G* → green), while those acquired earlier may be naturally biased shape-colour pairings (e.g. *X* → black). In a follow-up study, pre-literate children also further associated *I* and amoeboid shapes with white, and Z and jagged shapes with black, implying a more general natural bias for spiky or sharp shapes with black, and round or smooth shapes with white (Spector & Maurer, 2011). Finally, Brang, Rouw, Ramachandran, and Coulson (2011) showed that graphemes with similar visual features tended to have more similar colours, so the visual characteristics of graphemes do appear to have some influence on their associated synaesthetic colours.

Another explanation is that these shared associations might come from implicit linguistic, rather than explicit perceptual, characteristics of these colours and graphemes. Two studies found that the saturation and luminance of the colours associated with graphemes by synaesthetes are modulated by how frequently those graphemes appear in the synaesthetes' native language (in German, Beeli, Esslen, & Jäncke, 2007; in English, Smilek, Carriere, Dixon, & Merikle, 2007). Specifically, graphemes that are high in frequency (e.g., A, S) have colour associations that are more saturated (i.e., richer in colour; Beeli et al., 2007) and more luminant (i.e. brighter; Smilek et al., 2007). Simner et al. (2005) suggested that these effects may be better explained by also considering the linguistic frequency of the colour term (see also Simner & Ward, 2008): for example, that *A* is red for synaesthetes because A is a high frequency letter and *red* is a high frequency colour term, while low-frequency letters like *Q* tend to be paired with lower-frequency colour terms like *purple*. Although Simner et al. (2005) showed that high-frequency letters tend to be paired with high-frequency colour terms, they could not explain why those particular combinations arose: why is high-frequency *A* consistently red but not another high-frequency colour like blue? In other words, what is special about the connection between *A* and red in particular? The current study will attempt to answer this question by proposing that at least some of the letter-colour pairings of synaesthetes and non-synaesthetes are based on word associations acquired during early alphabet acquisition.

Our study will show that for the average person, each letter of the alphabet becomes associated with a particular word or words during alphabet acquisition, particularly through alphabet books (see Nodelman, 2001). These books commonly present a letter of the alphabet with a word beginning with that letter using the phrase "A is for…; B is for…" as a way to encourage children to make the connection between sound, spelling, and words. We will refer to these associated words, which are explicitly linked to the identity of the letter through repeated reading, as *index words*. Here we propose that the prototypical colour of the index word for each letter becomes associated with the letter itself. In short, this can be exemplified as, "*A* is red because *A* is for *apple,* and apples are red." This suggestion has been mentioned by studies investigating common grapheme-colour associations (e.g. Hancock, 2013; Spector & Maurer, 2011) but, to the best of our knowledge, it has never been empirically tested. Here we ask whether this sort of orthographic-semantic meditation (*A* → apple: orthographic mediation; apples → red: semantic mediation) has any basis in psychological reality when it comes to how letters are internally represented, both for synaesthetes and non-synaesthetes.

In order for this approach to be viable to explain letter-colour commonalities, three connections must be established. First, it must be the case that a letter (e.g. *A*) is consistently associated with a specific word beginning with that letter (e.g., *apple*) across a large proportion of the population. Second, it must also be the case that this index word (apple) has a consistently associated prototypical colour for most of the population (e.g., apples are predominantly conceptualised as red). Third, the prototypical colour of the index word must also be the preferred colour for the letter when people are asked to give direct letter-colour associations (e.g., *A* is red). If we find that all three are true, this will support an index-word explanation for grapheme colours. Therefore, the current study will elicit the index words for each letter, the colour of the index word's referent, and the colours directly associated with each letter for both synaesthetes and non-synaesthetes.

Experiment 1: *A* is for apple

Here we will ask whether certain words are consistently associated with each letter of the alphabet (e.g. "*A* is for apple"). We will refer to these highly-associated words as *index words*. Experiment 1 prepares for our subsequent investigation into whether index words have prototypical colours (Experiment 2) which influence letter-colour judgements (Experiment 3). As well as identifying index words, this first study will also explore linguistic characteristics of index words, and what determines their selection above other words in the language.

Method

Participants

Our participants comprised 315 non-synaesthete native English speakers from the USA. Participants were 43% female with a mean age of 33.2 years old (*SD* = 9.9); all were older than 20. We recruited our participants from Amazon's Mechanical Turk (see below). Since this platform is open to workers around the world, we necessarily selected our target sample (*N* = 315 non-synaesthete native speakers of American English) from a larger population by additionally testing the following subjects who were subsequently removed: 71 American English-speakers who self-declared synaesthesia (see *Procedure*), and participants who took the test but who were either non-native speakers (*N* = 55) or who were non-American English speakers, these being from India (22 total, 9 reporting synaesthesia), unspecified national origin, i.e. "white" (22 total, 5 reporting synaesthesia); and a further 18 from various national backgrounds. This left our final sample of 315 subjects.

Materials and procedure

Participants were recruiting using Amazon's Mechanical Turk (hereafter MTurk; www.mturk.com), a self-termed online "marketplace for work" for tasks requiring human intelligence. Workers can preview and complete experiments on the website, and are compensated with a small financial reward once their submissions are approved by the requester. MTurk has been validated as an effective research tool (Bankieris & Simner, 2015; Goodman, Cryder, & Cheema, 2013), and the reward we offered ($.20 per completed test) falls within the typical rate (Buhrmester, Kwang, & Gosling, 2011).

After giving demographic data (gender, age, nationality, and native and additional languages) participants began our test, which consisted of a series of phrases in the format "[Letter] is for…" (e.g. "A is for…"). Each letter was followed by a text box, and letters appeared in alphabetical order. This ordering was intentionally selected to evoke alphabet books and early literacy learning. Participants were given the following instruction: "In the box below each phrase, write the first English word beginning with that letter that you think of. Please answer as quickly and instinctively as possible." Participants then completed each sentence in alphabetical order and this continued to the end of the test ("Z is for…"), where a final question asked participants if they experienced synaesthesia (defined as "lifelong colours for letters or digits").

Results

Data validation

Our dataset comprised a series of words associated to letters, and we first minimally cleaned our data using the following criteria. First, responses that clearly referred to the same concept were combined (e.g. "apple" and "apples" both fell under "apple"). This was not done when the plural morpheme created two different words (e.g. "new" and "news" were not combined) nor when any other affixation gave rise to different concepts (e.g. "killer", "kill", and "killing" were not combined). We corrected spelling mistakes where the intended word was clear (e.g. "giraffee" was combined with "giraffe"). However, ambiguous responses were left as they were, and therefore counted as unique responses (e.g. "ca", which could have been intended as "cat", "car", "can", etc.).

Identifying index words

We next asked whether each letter had a particularly dominant *index word* from among the *response words* given by our participants. To begin, we calculated the agreement for each response word across our participants to measure whether different people gave the same word for each letter. Here, *agreement* indicates the percentage of subjects who agreed on a response for any given letter (e.g., over 80% of subjects agreed that *A* is for apple). Figure 1 shows the most commonly chosen word for each letter according to this metric.

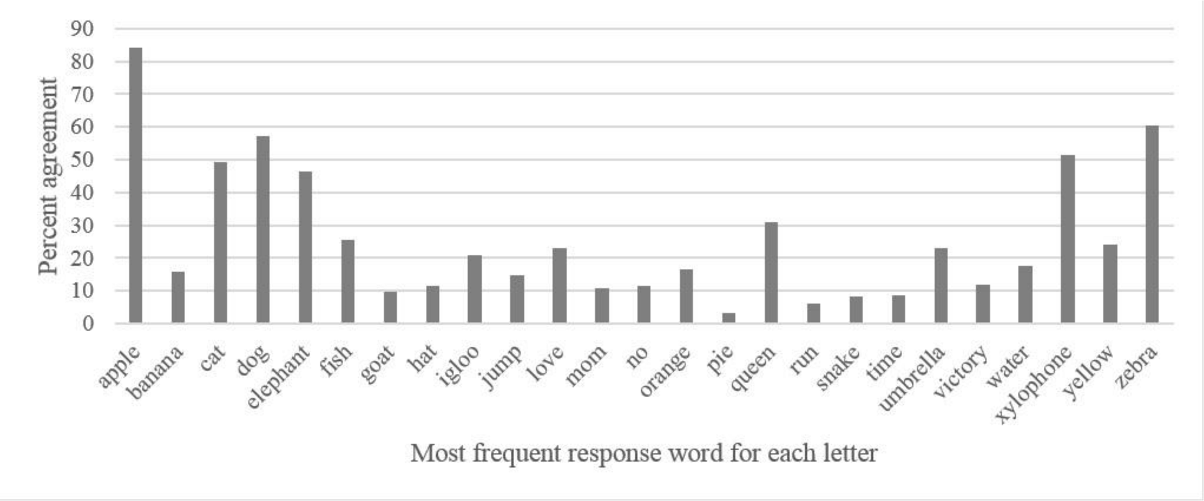


Figure 1. The highest-agreement index word for each letter, listed in alphabetical order with its agreement (as a percentage of all responses for that letter).

From Figure 1, it is immediately clear that while some letters have a clear index word out of all response words (i.e. *apple, dog, xylophone,* and *zebra* all have over 50% agreement), other letters do not (e.g. *pie* at 3.2%). We identified the three response words for each letter that had the highest agreement as potential *index words.* A list of these three index words for each letter and their percentage agreement can be found in Appendix 1. Our main focus is how these index words may help shape grapheme-colour associations (Experiments 2, 3), but in the following section we briefly explore what psycholinguistic factors underlie English speakers' choice of index words.

Characteristics of response and index words

We examined several possible predictors for how subjects chose response words for any given letter. In this section, we first analysed the entire set of response words given by our subjects (e.g., *A* is for *apple, animal, aardvark,* etc.) to increase our dataset. We considered several factors that might make these response words not only different from other words in the language (e.g., *animal* was a response word but *annex* was not) but which might also distinguish those that were chosen very often from those chosen less often (*apple* was chosen more often than *aardvark*). This analysis may help us understand why particular words might be more likely to contribute their prototypical colour to grapheme-colour trends. To this end, we examined several possible predictors for response word agreement, beginning with word frequency.

Higher-frequency words are reliably elicit quicker responses in behavioural tasks (e.g., Oldfield & Wingfield, 1965) so we first tested whether there was a tendency for response words to be high frequency. For this we entered lexical frequency as a predictor of agreement in response words across our subjects in a multiple regression using frequency measures from several different corpora (CELEX: Baayen, Piepenbrock, & van Rijn, 1993; Kučera-Francis: Kučera & Francis, 1967; HAL: Lund & Burgess, 1996; and SUBTLEX-US: Brysbaert & New, 2009). However, none of these measures predicted response word agreement (for all predictors, $t < 1.42$, $p > .155$), meaning that higher-frequency words were no more likely to have higher agreement among our participants than lower-frequency words. However, these frequency measures do not capture the task demands of our experiment. That is, we asked our participants to give a response *within each letter*, whereas the above frequency measures are all calculated from entire corpora across all letters. This means that while *xylophone*, for instance, may be one of the highest-frequency words that begins with *X*, it has a very low frequency in the language as a whole, because the frequency measures used above do not group frequency by spelling.

We therefore developed a new frequency measure: the frequency of each word per million within all occurrences of words *beginning with the same letter*, which we call *by-letter frequency*. To calculate by-letter frequency, we made use of the SUBTLEX-US database of American English film subtitles because this contains the same variety of (American) English used in our experiment, and it has been shown to predict response times to lexical decision and naming tasks better than older, more widely-used corpora (Brysbaert & New, 2009). For each of the *response words* generated in our study (e.g., *apple*, *animal*, *aardvark,* etc.) we divided its total count in the corpus by the sum of all counts for every word sharing that initial letter. For example, *apple* appears in the corpus 1,207 times, so we divided this raw frequency count by the total count of all words in the corpus beginning with *A*. We then multiplied by a million to produce a by-letter frequency per million (e.g., for *apple*, this was 303.9). As the resulting distribution was highly skewed, we also $\log_{10}$-transformed the values to reach a final log by-letter frequency for all of the words in the corpus.

Having established the *log_{10}-transformed by-letter frequency* for all of the words in the corpus, we then compared the response words generated by our participants to the rest of the words in the corpus to see whether their log by-letter frequency differed. We conducted a 2 (Word-type: Response/Non-response word) x 26 (Letter: A-Z) ANOVA predicting the log by-letter frequency. Our most important finding was a main effect of response word ($F(1, 74234) = 7842.73$, $p < .001$), indicating that response words have higher frequency (M = 6.03) than non-response words (M=1.74). That is, when people are asked to name the first word that they think of beginning with a particular letter, they tend to choose one of the most common words in English within that letter category (e.g., A is for *apple*, not *annex*). There was also a main effect of letter, ($F(25, 74234) = 407.82$, $p < .001$), reflecting the fact that the mean frequencies of the words in each letter group differed between letters. Finally, there was a significant interaction between response word and letter ($F(25,74234) = 5.73$, $p < .001$). Bonferonni-corrected post-hoc tests showed that mean log by-letter frequency was significantly higher for response vs non-response words for all letters (ps < .001, α corrected for 26 comparisons = .002) except the letter *X* ($t(29)= -1.39$, $p = .175$). This failed to reach significance because of the small number of words in English beginning with *X* combined with a very high frequency of *x* counted on its own as a word (raw by-letter frequency per million = 750714.3, log by-letter frequency = 5.88). The difference between response vs non-response word frequencies by letter is clearly illustrated in Figure 2.

We also calculated the percentile rank of each word within all words beginning with the same letter to evaluate this further. Out of the 2024 response words, only 56 (2.77%) fell below the median log by-letter frequency, and 257 (12.70%) fell outside the 75th percentile. This underscores that for the vast majority of response words, high frequency is a defining characteristic.

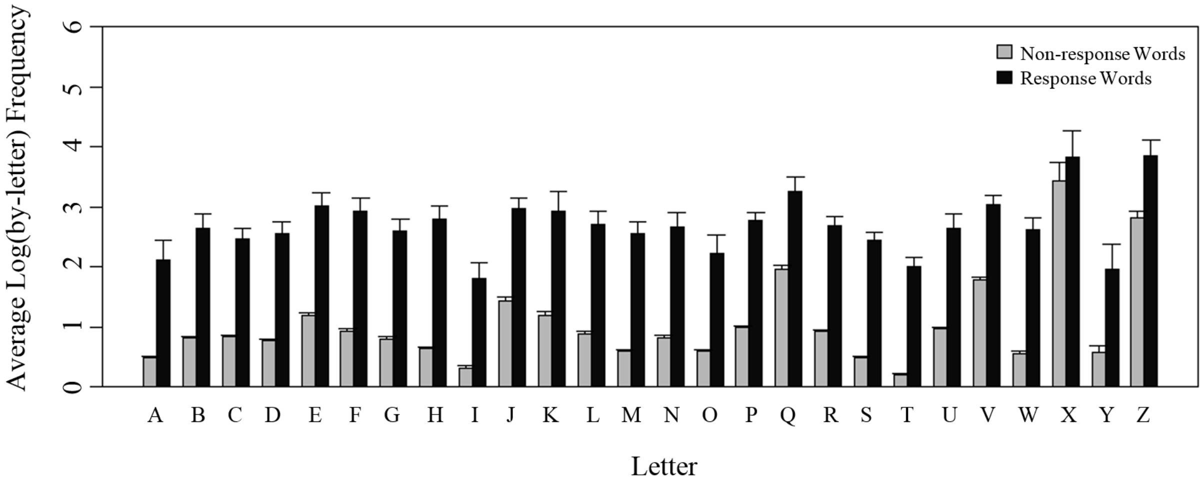


Figure 2. Comparison by letter between average log by-letter frequency of response words (dark bars) and non-response words (light bars). This difference was highly significant for every letter except *X*.

Having shown that response words tend to be high frequency, we now ask what determines the degree of agreement among respondents – that is, why some response words came up more often than others. The relationship between log-by letter frequency and agreement is illustrated in figure 3, showing that while the majority of response words were single, unique instances, the higher-agreement words tend to also have higher log by-letter frequency.

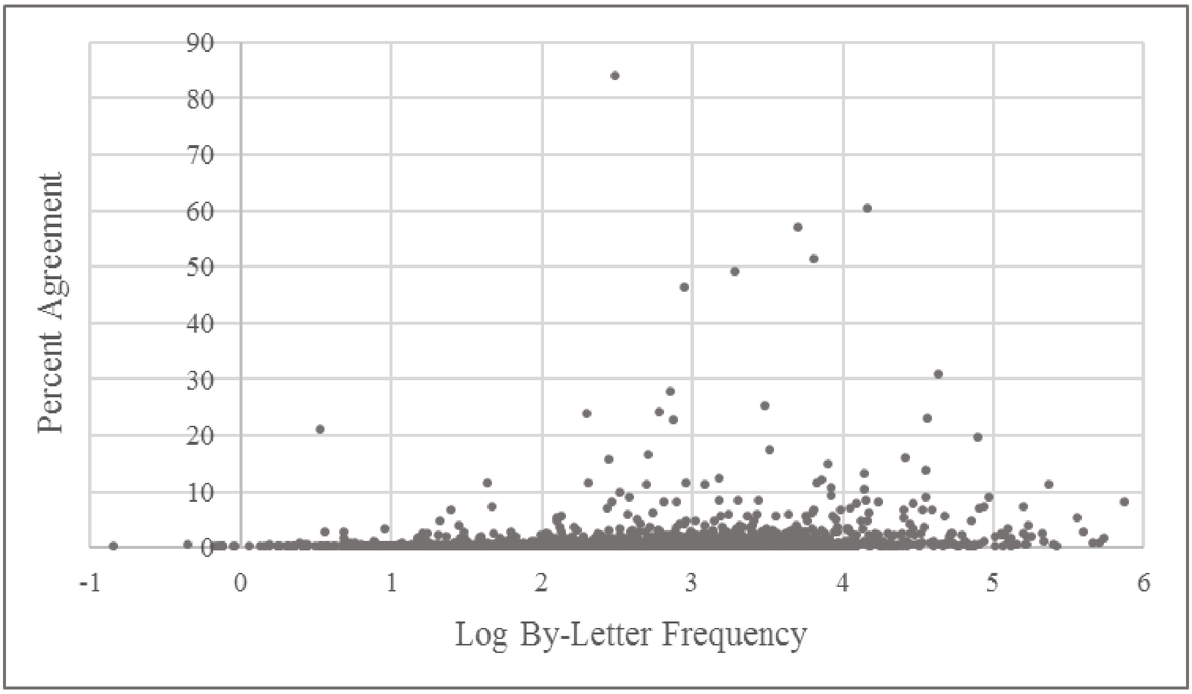


Figure 3. Scatterplot of log by-letter frequency versus percent agreement for all response words.

We repeated our multiple regression for frequency to predict the percentage agreement among participants, but used our new log by-letter frequency as a predictor along with 6 other predictors which we hypothesised may contribute to response word agreement: age of acquisition, imageability, familiarity (Bird, Franklin, & Howard, 2001; Gilhooly & Logie, 1980; Stadthagen-Gonzale & Davis, 2006), neighbourhood size (i.e., the number of words that differ from the target word by just one letter change, such as "car" and "mat" for *cat*), and behavioural reaction times in lexical decision and naming tasks (English Lexicon Project; Balota et al., 2007). Our final model (see Table 1) shows that only log by-letter frequency and imageability were significant predictors. In summary then, the most widely agreed-upon index words across subjects tend to be the most frequent words for that letter, and also the more highly imageable.

Table 1. Summary of the regression model predicting percent response word agreement. $R^2 = .059$, $F(2,1217) = 38.44$, $p < .001$.

	B	*SE B*	*t*	*p*
Intercept	-5.245	0.794	-6.609	0
Imageability	0.008	0.001	7.148	< .001
Log by-letter Frequency	0.968	0.146	6.617	< .001

Discussion

This experiment investigated whether letters of the alphabet have associations with words that are shared among language users. In our study, we asked participants to complete phrases of the type "A is for ____". We classified words that were generated by our participants as response words, and calculated how much agreement there was among participants for each response word. We found that some letters of the alphabet are indeed consistently associated with particular words and have high agreement, and we have termed these *index words* (e.g. *A* is for *apple*). We further demonstrated that the total set of response words from our subjects were higher in frequency than the remaining words in English, but only if frequency is considered within each initial letter. To do this, we created a new frequency measure: a log-transformed frequency per million within each letter based on the SUBTLEX-US corpus (Brysbaert & New, 2009) that more accurately related to the task that we had set our participants of choosing a word beginning with a particular letter (i.e., *xylophone* has a high log by-letter frequency because it is one of the more frequent words beginning with *X*, even though it is low frequency within the language overall). Using this by-letter frequency, we found that the most widely agreed upon index words are those that are the most frequent by this measure, and are also highly imageable. These types of words are the central feature of alphabet books commonly used in literacy pedagogy with demonstrable success (e.g. Nowak, 2015), and our findings suggest that at least some of these letter-word associations endure into adulthood. We will next ask whether the connection between index words and letters is strong enough to account for direct letter-colour associations in synaesthetes and non-synaesthetes. We do this by now establishing the colours of index words.

Experiment 2: Apples are red

In this experiment, we will focus on the top three highest-agreement response words that we identified in Experiment 1, which we have termed *index words*. We will seek to establish whether the index words for each letter refer to entities that have consistent, prototypical colours (e.g. what is the prototypical colour of an apple?). If these words do indeed have consistent colour associations, we will then be able to compare these colours with the colours associated with letters directly (see Experiment 3, below).

Method

Participants

Our participants comprised 146 English-speaking American non-synaesthetes, 49.3% female (*N* = 66) with a mean age of 37.7 years (*SD* = 12.8 years, range = 19 to 75 years). As in Experiment 1, all participants were recruited using MTurk. These participants had not taken part in Experiment 1. In order to match cultural and linguistic background with the index words gathered in Experiment 1, we excluded participants who were non-native speakers of English and/or were not Americans (*N* = 12). MTurk allows the requester to specify geographic location, so we required that our test would only be available to workers in the United States. (We had not specified this in our first experiment because we were unsure which language or cultural group would dominate our initial sample.) Therefore, nonspecific responses to nationality, e.g. “white”, “black”, were this time included in the analysis, so long as the geographic location was our target location. As before, we screened participants for self-reported grapheme-colour synaesthesia at the end of the test, using the question described in Experiment 1. We also tested but subsequently removed a further 90 participants because they had already taken part in Experiment 1 (N = 9) or potentially self-declared synaesthesia (by answering “Yes” N = 24, or “Don’t Know” N = 57 to our synaesthesia question; see *Procedure*). All respondents were compensated $0.40 for their participation. This left our final sample of 146 subjects.

Materials

The materials from this study were a subset of the words generated in Experiment 1, which had been elicited in that study using phrases such as “A is for ____; B is for ____...”. In the current study we selected only the top three highest-agreement response words as index words to be tested here. This resulted in a final list of 78 words (3 words x 26 letters); these items are listed in full with their percentage agreement from Experiment 1 in Appendix 1.

Procedure

We created our study using Qualtrics survey software (www.qualtrics.com) and posted its URL on MTurk. After a brief introduction and the collection of basic demographic information (age, gender, nationality, native and other languages spoken), the 78 target words were presented in a unique random order for each participant. For each word, participants were instructed to form a mental image and then provide the “strongest, most dominant colour.” Colours were selected from a drop-down list of basic colour terms (black, white, red, orange, yellow, green, blue, purple, pink, brown, gray). Participants were also asked to provide a confidence rating for how sure they were that the colour they had chosen was the best colour for each item, on a Likert scale from 1 (not sure at all) to 7 (very sure). The test required participants to give both a colour and a confidence score for every item before it would allow them to advance.

Results

Response validation

We conducted an initial check that participants had completed the test to a sufficient standard and in good faith. We identified and removed 12 subjects who were responding randomly or repeating the same colour-choice throughout on the following basis. First, we selected four items we believed should have a unique colour association*: banana*, *elephant*, *orange*, and *yellow* (coloured: yellow, grey, orange, and yellow, respectively). We then asked three independent raters to confirm our intuition, which they did in 100% in agreement, as did over 92% of our participants. We therefore removed any participants who differed from these independently-established responses for two (50%) or more of these standardised items. Next, we identified participants who had chosen the same colour repeatedly regardless of the item (e.g. *green* for most words). We calculated the mean number of times that each colour was chosen for each word across all participants, and established a first cutoff at 2.5 standard deviations above the mean, and a second cutoff at 3 standard deviations above the mean. For example, *orange* was selected an average of 4.5 times (*SD* = 3.0) by each participant, with a first cutoff of 12.0 and a second cutoff of 13.5. We removed any participants who selected more than two colours above the first cutoff (e.g. who selected *orange* 12 times or more), or one colour above the second cutoff (e.g. who selected *orange* 14 times or more). Using both these criteria, we identified 12 problematic participants and excluded them from the analysis. This left a final pool of 134 participants.

After removing inattentive subjects we next validated our dependent measure, which was the frequency with which any given colour was selected for our target items. For example, for *apple*, the most commonly selected colour was *red*, which was selected 122 times out of 134 responses, which gave a maximum colour frequency of 91.0%. As described above, participants also gave a confidence rating, which we used to validate their colour choice. To do this, we compared the mean confidence rating for each word to its max colour frequency, using a Spearman nonparametric correlation as our data were not normally distributed. This correlation showed that when participants were more consistent in their colour selection (i.e. items had higher max colour agreement), they were also more confident in their colour choice (Spearman's ρ [78] = .822, $p < .001$), thereby validating our dependent measure.

Do index words predict letter-colour trends?

We can now use our results from Experiments 1 and 2 (letter →index word → colour) to predict which colours would be most often associated with each letter if index words do indeed influence direct letter-colour pairings. To do this, we took into account how often each index word was chosen for its letter and how often a particular colour was chosen for each index word. We will illustrate our procedure using *A* and *apple*. We first calculated the percentage of each colour for each of the three index words, so for *apple* the responses were 91.04% red and 8.96% green, and so on for every colour. We then multiplied this result by the percentage of times that particular index word was selected for that letter. So, as *apple* accounted for 84.13% of all of the responses for *A*, this means that red accounts for 76.59% (.8413 x 91.04%) of the colour selections for *A* via *apple*. On the other hand, since *animal* was only given as a response for *A* 1.90% of the time and brown was selected for *animal* 58.21% of the time, the brown responses for *animal* only count as 1.11% (.190 x 58.21%) of the total proportional colour responses for *A*. We applied this procedure to all combinations of index words and colours, then we summed the resultant weighted colour responses for all three words for each letter – all the red responses, all the brown responses, etc. This gave a colour score for each colour within each letter. The highest colour score for each letter indicates the most *dominant colour* for that letter, and is therefore the colour we would predict if letter colour is mediated by index words. These predictions are detailed in full in Appendix 2.

This colour score predicts the most likely dominant colour for each letter, while still taking into account the amount of index word agreement. That is, the colour scores for *A* are 76.61% red, 7.60% green, etc., but the sum of the eleven colour scores for *A* is less than 100%. This is because each score reflects the percentage agreement for index words for each letter. In the case of *A*, for example, the three index words *apple, animal* and *aardvark* accounted for 87.08% of all responses for *A*, so the eleven different colour scores for *A* sum to 87.08%. The remaining response words for *A* (e.g. *automobile, ant,* etc.) account for the remaining 13.02% of all responses. Since we only collected colours for the top three index words, the colours for non-index response words like *automobile* or *ant* are not represented in the colour score, so the colour distribution of the remaining 13.02% for *A* is unknown. In contrast to *A*, the three index words for *S*, which are *snake*, *sister,* and *stop*, together only accounted for 17.8% percent of the response words for *S*, so the remaining 82.22% of the colour score variation for *S* is unknown and the highest colour score for *S* is very low (dominant colour: green, 3.93%). This distribution is illustrated in Figure 4 below for both *A* and *S*.

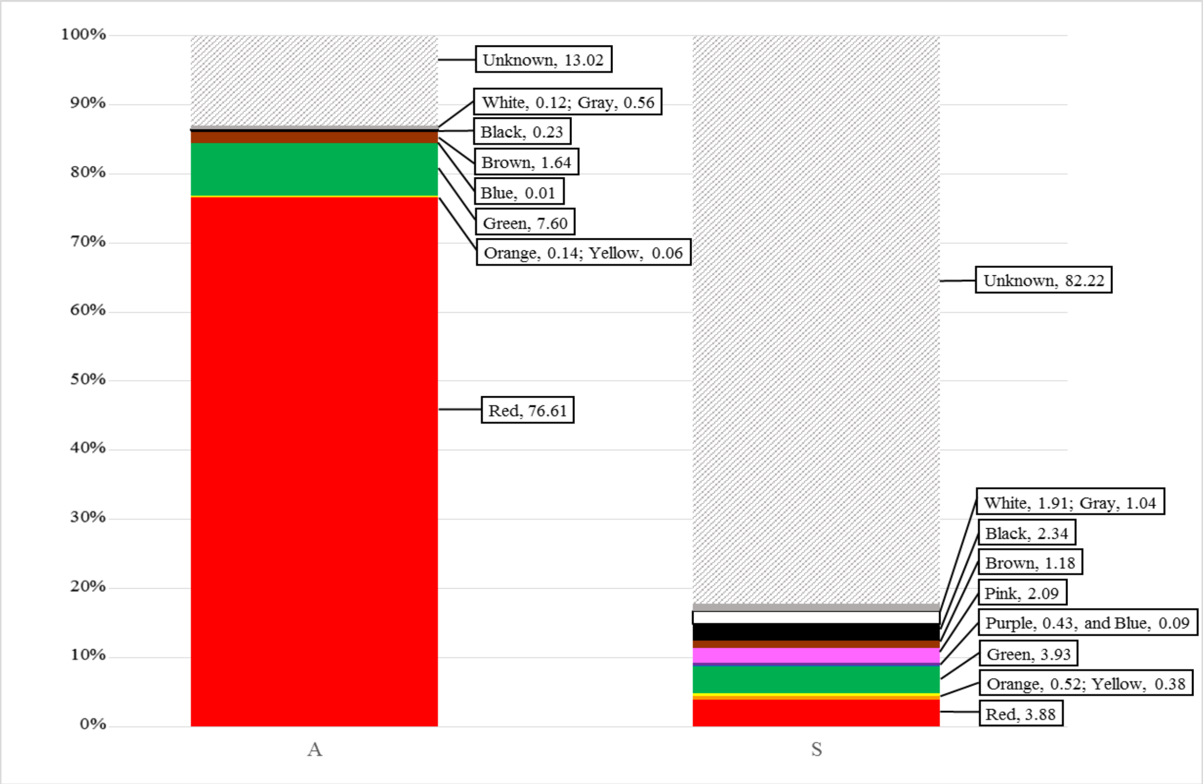


Figure 4. The weighted proportions of colour choices for *A* (left column) and *S*. In both cases, the diagonally barred "unknown" proportion shows how much of the overall colour distribution for each letter is accounted for by response words not included in our word → colour selection task (since we included only the top three index words for each letter). For *A* this is only 13.02%, but for *S* this is 82.22%. Each section is labelled with both the colour name and its colour score, in percent.

Comparison to previously reported letter-colour trends

Next, we evaluated whether the dominant colour that we predicted for each letter via index words, as calculated above, successfully matched with previously reported letter-colour associations in the literature. To do this, we return to the colour associations presented in the introduction, where we explored the grapheme colouring trends in both synaesthetes and non-synaesthetes. Table 2 compares the predictions from the current study with the previously published colour-letter associations across the studies reviewed in the literature.

Table 2. An abbreviated summary of the previously reported trends in grapheme-colour associations (first three columns) for comparison with the current study (far right column) predicting associations

via index words. The first column reports the letters that show associations across all three “significance studies”, i.e. Simner et al. (2005), Rich et al. (2005), and Jonas (2010), who reported the pairings that were statistically significant. The second column reports only the most commonly chosen letter-colour pair (Witthoft et al., 2015). Colour-letter pairings from any of the first three columns that match with the current study are highlighted in bold.

Colour	Synaesthetes, significance studies	Synaesthetes, Witthoft et al.	Non-synaesthetes, all studies	Current study
Black	I**XZ**	**XZ**	**XZ**	CHPU**XZ**
White	**I**O	**I**O	**I**W	G**I**
Red	**A**R	**A**R	**A**R	**A**JL
Orange	J	HJKN	**O**	**O**T
Yellow	**Y**	CLS**Y**	L**Y**	B**Y**
Green	-	EFG	FG	S
Blue	B	BDT**W**	BI	**W**
Purple	V	P**Q**V	PVM	**Q**
Pink	P	-	P	-
Brown	**D**	-	**D**HT	**D**KMNV
Grey	X	-	UX	EFR

The colour-score-predicted dominant colour matched with previously observed trends for nine letters, or 35% of the alphabet: *X, Z, I, A, O, Y, W, Q*, and *D*. However, two of these letters, *Q* and *W*, were associated by synaesthetes and not by non-synaesthetes, so we will focus here on the seven letters (*X, Z, I, A, O, Y,* and *D*, 27% of the alphabet) that matched between previous associations reported for non-synaesthetes and predicted dominant colours from our own non-synaesthete participants. The cumulative probability of obtaining seven or more matches out of 26 by chance, given an equal 1/11 = .09 probability of each of the eleven colour terms being selected for any given letter, is approximately one in 147.8, or $p = .007$. This is a promising result, but in order to quantify it further, we will gather our own direct letter-colour associations *from the same population* in Experiment 3, below, to match our results by linguistic and cultural background.

Discussion

In the current experiment, we asked participants to tell us the prototypical colours that they thought of for index words identified in Experiment 1 (e.g. “apples are red”). We then calculated a colour score, combining the agreement within each letter from Experiment 1 with these colour associations, that allowed us to make predictions about which colours would be most dominantly paired with which letters, and how strong that association would be. By comparing our predictions with previously published results from studies of synaesthetes and non-synaesthetes, we showed that our predictions were able to account for nine letter-colour associations (*X, Z, A, I, O, Y, W, Q*, and *D*) that were not previously directly explicable. The current experiment provides a new explanation for previously reported trends in grapheme-colour pairs. We suggest that during the course of literacy acquisition, alphabet books and other classroom materials pair index words with letters, such as *apple* and *A*, and this orthographic association is internalised as an association between *A* and red via the prototypical semantic colour for *apple*. We further suggest that the same effect can explain *Q* with purple (via *queen*), *D* with brown (via *dog*), *W* with blue (via *water*), and *X* and *Z* with black (via *x-ray* and *zebra*).

We will now test the predictions of our dominant colours directly using American English-speaking participants. Thus far, all of the grapheme-colour trend studies have used populations and methods different from the current study. In order to match as closely as possible for the influences of cultural and sociolinguistic background, we will collect our own direct letter-colour pairings for further analysis from the same population.

Experiment 3: *A* is red

This third experiment will directly gather letter-colour associations from a similar population that provided index words (Experiment 1) and the colours of those index words (Experiment 2). As described in the introduction, some studies have already sought to establish colour trends in non-synaesthetes but used populations with cultural and linguistic backgrounds different from the current study. As the previous trends come from British (Simner et al., 2005; Jonas, 2010), Australian (Rich et al., 2005), and mixed nationality (Witthoft et al., 2015) participants, we will obtain our own letter-colour associations from the same population as the previous experiments (i.e. English-speaking American MTurk workers), which allows us to control for location and language. We still expect to find some of the same general patterns of letter-colour associations as have been previously reported. More importantly, we will be able to directly and numerically compare these letter-colour choices with the dominant-colour predictions made by our colour score from Experiment 2.

Method

Participants

We tested a final sample of 175 American English-speaking non-synaesthetes, 56% female (*N* = 98) with an average age of 33.30 (*SD* = 10.10 years, range = 20 - 82 years). Following the same procedure described in Experiments 1 and 2, additional participants were removed from the analysis if they declared a nationality or native language other than American and English. We also only used responses from participants who answered “No” to our screening question for synaesthesia, which asked, “Do you experience synesthesia? In other words, were the colors you gave in this task associations that you've known about all your life?” This question was different from Experiments 1 and 2 because of the difference between the tasks – those experiments asked for the prototypical colour of real-world objects, rather than synaesthetic colours. On the basis of nationality, language background, and/or self-reported synaesthesia, 110 participants were removed. Finally, five participants were excluded because they had already participated in a previous experiment. This gave us a final pool of 175 participants.

Materials and procedure

The recruitment, instructions, and survey apparatus were similar to that reported in Experiment 2 with the following adjustments. First, the list of words used in Experiment 2 were switched out for the 26 letters of the English alphabet. Participants were instructed to choose the colour for each letter that “seems to fit the letter best.” We also removed the confidence rating task. All participants provided colours freely for all letters (i.e. they could choose the same colour as many times as they liked) but were required to provide a colour for every letter. The order of the letters was randomised for each participant, and the order of colour options was also randomised for each letter. This was especially important as Simner et al. (2005) showed that non-synaesthetes tend to associate colours with letters in the order that the colours are easiest to generate.

Results

Data validation

We conducted a basic data validation procedure similar to that outlined in Experiment 2. For each colour, we calculated the distribution of colours for each participant, and the overall mean number of selections for each colour. Thirteen participants were excluded because they had more than one colour selected above 2.5 standard deviations from the mean, and/or they had chosen the same colour for more than 25% (in this case, 7 or more) of the letters. This resulted in a final pool of 162 participants.

Letter-colour associations

In this analysis, we found the most frequently selected (i.e. highest-agreement) colour for each letter by calculating the proportions of colour selections for each letter. Table 3 below compares this *modal colour* to the letter-colour data from previous studies of letter-colour associations.

Table 3. An abbreviated summary of the previously reported trends in grapheme-colour associations (first three columns) for comparison with the current study (far right column) directly associating letters and colours. Colour-letter pairings from any of the first three columns that match with the current study are highlighted in bold; letters that appear more than once were tied for modal colour.

Colour	Synaesthetes, significance studies	Synaesthetes, Witthoft et al.	Non-synaesthetes, all studies	Current Study
Black	I**XZ**	**XZ**	**XZ**	**XZ**
White	IO	IO	I**W**	H**W**
Red	**AR**	**AR**	**AR**	**A**FK**R**
Orange	J	HJKN	**O**	C**O**
Yellow	**Y**	C**L**S**Y**	**LY**	H**LY**
Green	-	EF**G**	F**G**	**G**S
Blue	**B**	**B**D**T**W	**BI**	**BI**J**T**U
Purple	**V**	**P**Q**V**	**PV**M	J**PV**
Pink	P	-	P	Q
Brown	**D**	-	**D**HT	**D**MN
Grey	X	-	UX	E

To further establish the connections between letters and colours, we next conducted an analysis of *statistically significant* letter-colour pairings, following Simner et al. (2005). We did this to distinguish between colours that are simply selected often overall as opposed to associations that occur at a frequency significantly beyond chance. In this analysis, we first counted the total number of times each colour was selected for all participants and all words, from which we calculated the baseline probability of each colour being selected. We then used a binomial distribution to calculate the probability of each letter-colour pairing occurring at above chance levels. These *significant colours* are described in Table 4, below.

Table 4. For each colour term, the letters are listed which were significantly associated with that colour at least at $p < .05$ level in our non-synaesthete participants' responses. Letters appear more than once if they were significantly associated with more than one colour. For each associated colour, letters are listed in separate columns if they were significant at $p < .01$ (middle column) or $p < .001$ (right column).

Colour	Significant at $p < .05$	Significant at $p < .01$	Significant at $p < .001$
Black	B D		X Z
White	E F T	H I	W Z
Red	F S		A R
Orange	J N		C O
Yellow	H		L Y
Green	C D E F	S T	G
Blue	A I J	U	B T
Purple	I Q	J	P V
Pink	I L	K Q	P
Brown	H T	B K M	D N U
Grey	D U	G Q X	E

The results of the binomial analysis, first, confirm the modal (i.e. most selected) colour as statistically significant for every letter. However, a comparison between modal and significant colours reveals how the degree of agreement can vary for each letter. For example, the modal colour for *S* was green, but red was also significantly associated with *S*, with only 1.23% of agreement separating them. Eleven letters differed between their modal colour and second most frequently selected colour by less than 5%, and three (*H*, *J*, and *K*) had exact ties for their modal colour. For this reason, we will include both first and second most selected colours for each letter for the purposes of evaluating the predictions of the index word colour score, and refer to these as *first* and *second modal* colours. In this way, we can distinguish the colours that are selected most often (i.e. *modal* colours) from those that are associated at a statistically significant level (i.e. *significant* colours).

Index word predictions vs letter-colour associations

We can now directly compare the letter-colour combinations predicted by the colour score from Experiment 2 with the actual letter-colour associations collected in the current experiment. We will address this in two ways: first, by examining the relationship between levels of agreement; and second, by comparing the colours themselves.

To begin, we can evaluate the relationship between the colour score that we calculated in Experiment 2 and the max colour agreement described above with a correlation. First, we removed the letters from the analysis for which the colour association could be explained another way – namely, the initial letters of colour terms (i.e. *B, W, R, O, Y, G, V, P*). As the introduction describes, the letter-colour agreement for these letters is very high because of their connection to colour terms (e.g. *R* for red); indeed, the binomial analysis above showed that all initial letters of colour terms were significantly associated with their colour at $p < .001$, which may mask an effect for non-colour-term letters. After removing these colour-term letters, the remaining 18 letters showed a highly significant correlation between colour score (from Experiment 2) and colour agreement (from the current experiment; Spearman's $\rho = 0.70$, $p = .001$). In other words, the letters that had higher agreement in the colour obtained via their index words also had higher agreement for their directly chosen colour. The plot in figure 5 illustrates this correlation.

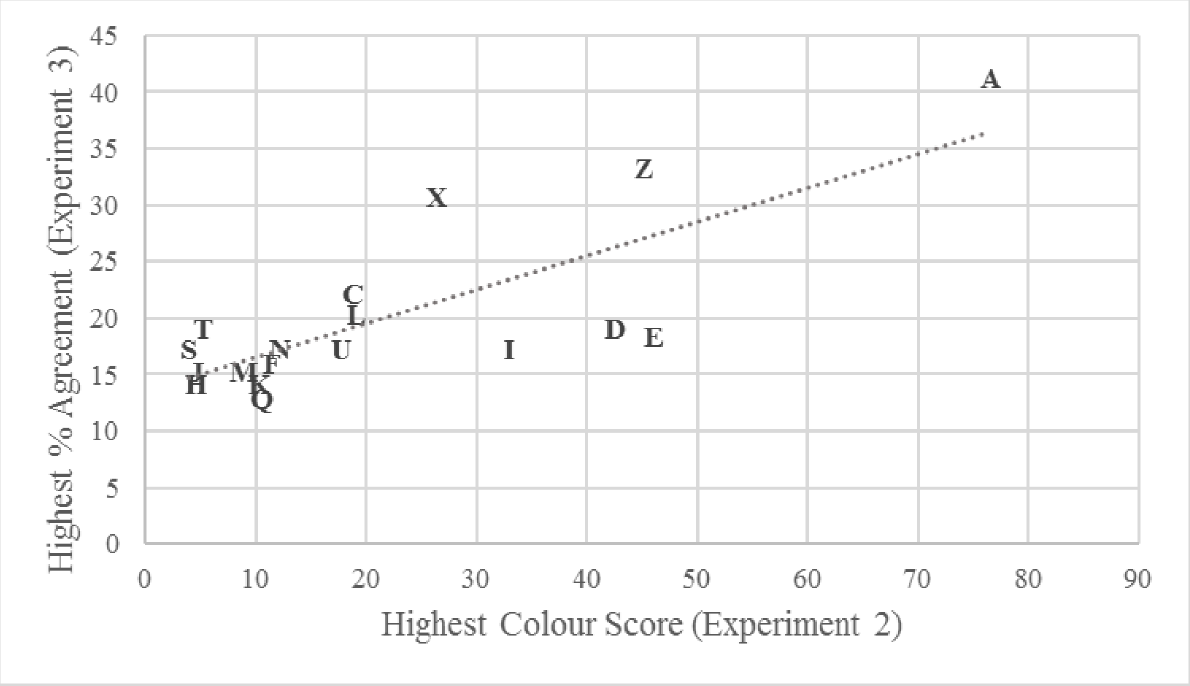
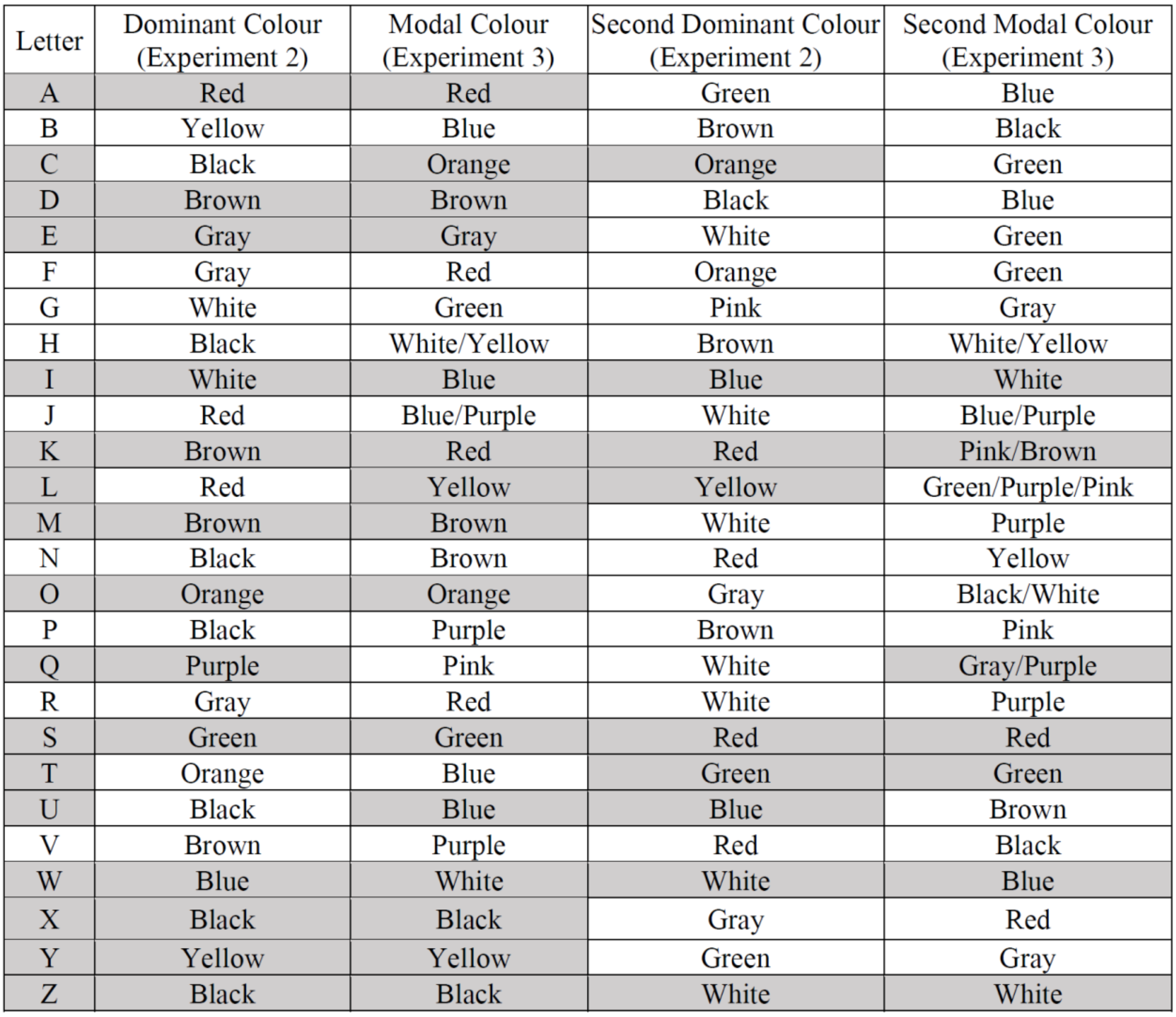


Figure 5. Scatterplot showing the correlation between the highest colour score obtained via index words in Experiment 2 and the highest agreement reported for direct colour association by non-synaesthete participants. Each of the 18 letters included in this correlation represents its respective point.

We will next explore *how* index words predict colours for letters. Figure 6 shows the first and second dominant colours predicted by the colour score from Experiment 2 and the first and second modal colour associations from Experiment 3. Altogether, there was a match between at least one of the dominant colours and one of the modal colours for 17 out of 26 letters (65%), and 22 out of a possible 52 combinations of first/second dominant colour and first/second modal colour, a hit rate of 38.5%. In order to understand how likely this pattern was to emerge by chance, we used a Monte Carlo simulation. For each iteration of the simulation, we randomly generated two pairs of colours, sample 1 and sample 2, 26 times, representing the top two colours that our participants selected for each letter via index words (dominant colours from Experiment 2) and directly (modal colours from the current experiment). Both sample 1 and sample 2 were composed of mutually exclusive colours – that is, each randomly generated pair of colours had to consist of two different colours, not the same colour twice. We then counted the number of times there was a match between the colours in sample 1 and in sample 2, reflecting the same colour-matching process between the dominant colours from Experiment 2 and the modal colours from the current experiment that we conducted on our data from our participants (see figure 6 for an illustration). We then repeated this process of generating two random pairs of colours for each of 26 letters and counting the number of matches for one million (10000000) iterations, which gave us a simulation of colour matches at chance level.

We then used this simulation to estimate of the probability of obtaining the patterns we observed. As noted above, we found at least one match for 17 out of 20 letters in our data. Our simulation found that 1,352 iterations out of 10000000 resulted in 17 or more letters with at least one colour match. Dividing this result by 10000000 gives a decimal probability of having matches for 17 or more letters by chance, comparable to a p-value, which in this case was p = .001. This by-letter analysis counted both a single match (e.g. sample 1: red, green; sample 2: blue, red) and a double match (e.g. sample 1: red, *blue*; sample 2: *blue*, red) as a single hit for a given letter. However, we also wanted to know the probability of obtaining the total number of matches we observed, so that a single match and a double match would count as two hits. Counting double matches as two hits rather than one means that there were 52 total possible hits for each iteration, since there were two hits possible for each of 26 letters.

The Monte Carlo simulation, again iterated 1000000 times, indicated that the probability of observing 22 out of 52 possible hits, as we did in our data, was .000017, or $p < .001$. Given this extreme unlikelihood that choosing colours by chance would lead to the pattern of matches we describe, we rather suggest that the colours for letters of the alphabet are systematically derived from index words.



Letter	Dominant Colour (Experiment 2)	Modal Colour (Experiment 3)	Second Dominant Colour (Experiment 2)	Second Modal Colour (Experiment 3)
A	Red	Red	Green	Blue
B	Yellow	Blue	Brown	Black
C	Black	Orange	Orange	Green
D	Brown	Brown	Black	Blue
E	Gray	Gray	White	Green
F	Gray	Red	Orange	Green
G	White	Green	Pink	Gray
H	Black	White/Yellow	Brown	White/Yellow
I	White	Blue	Blue	White
J	Red	Blue/Purple	White	Blue/Purple
K	Brown	Red	Red	Pink/Brown
L	Red	Yellow	Yellow	Green/Purple/Pink
M	Brown	Brown	White	Purple
N	Black	Brown	Red	Yellow
O	Orange	Orange	Gray	Black/White
P	Black	Purple	Brown	Pink
Q	Purple	Pink	White	Gray/Purple
R	Gray	Red	White	Purple
S	Green	Green	Red	Red
T	Orange	Blue	Green	Green
U	Black	Blue	Blue	Brown
V	Brown	Purple	Red	Black
W	Blue	White	White	Blue
X	Black	Black	Gray	Red
Y	Yellow	Yellow	Green	Gray
Z	Black	Black	White	White

Figure 6. Comparison between dominant colour predictions via colour score from Experiment 2 and modal letter-colour agreement from Experiment 3 for non-synaesthete participants. Shaded cells indicate a match between dominant and modal colour(s). Cells with more than one colour indicate a tie in agreement.

Discussion

This experiment collected colour associations for letters from a large group of English-speaking American non-synaesthetes; previous studies have examined British (Simner et al., 2005), Australian (Rich et al., 2005), and mixed nationality (Witthoft et al., 2015) English speakers. We first calculated the top two *modal*, or most frequently selected, colours for each letter, and used a binomial analysis to show which colours were also *significantly* associated with each letter. We also compared our modal colours to other studies and found that our participants chose colours for letters similar to those previously reported. We then compared the *dominant* colours for each letter that we had predicted using index words (from Experiments 1 and 2) with the *modal* colours, and showed that index-word-based dominant colour matched with directly associated modal colours significantly beyond what chance would predict.

First, our data provide a strong initial indication that the index words associated with letters (*A* is for apple) may indeed influence letter-colour pairings through orthographic-semantic associations. We also saw that despite sharing a common language with the English speakers in previous studies, for our participants some letters differed in their modal colour associations from previously reported trends (e.g. *M* with brown, *E* with grey) that nonetheless matched the dominant colour predicted by index words. We will explore the implications of this in depth in the general discussion, but first we will examine whether the index word route described above is also a meaningful predictor of letter-colour associations for self-reported synaesthetes as for non-synaesthetes.

Experiment 4: Synaesthetes

To further evaluate the influence of the index word route, we turn now to the trends reported for grapheme-colour synaesthetes. It may be that while non-synaesthetes rely on overlearned index words to form associations in a task that is not particularly meaningful for them (“What colour is the letter *K*?”), synaesthetes might rather rely on other implicit or systematic processes to determine colour associations. Support for this idea comes from Simner et al. (2005), who found that while synaesthetes and non-synaesthetes shared some implicit ‘rules’ in associating letters with colours, they also showed certain differences as well. Therefore, index words may predict different letter-colour pairs for synaesthetes. To test this, we will repeat the index word → colour → letter analysis above, this time using the responses from participants who self-reported experiencing grapheme-colour synaesthesia.

Method

Participants

A total of 88 self-reported synaesthete participants took part across the two colour-gathering experiments. In Experiment 2 (index words → colours) we identified 24 self-reported synaesthetes as described below, 14 female (58%) with a mean age of 35.58 (*SD* = 12.42). In Experiment 3 (letters → colours) there were 64 self-reported synaesthetes, 37 female (60.94%), with a mean age of 35.56 (*SD* = 11.06 years). All participants were American English speakers.

In both experiments, we used a self-report question to evaluate synaesthesia. For Experiments 1 and 2, this question was, “Do you experience synesthesia (lifelong colors for words, letters, or digits)?” and in Experiment 3, “Do you experience synesthesia? In other words, were the colors you gave in this task associations that you've known about all your life?” As explained in Experiment 3, above, this question was different between the experiments because Experiments 1 and 2 did not specifically ask for synaesthesia-like associations, whereas Experiment 3 did. For this present analysis, we included participants who answered “Yes” to these questions for either question. We acknowledge that this self-report measure is far less stringent than the widely accepted objective validation method of testing synaesthetes’ associations repeatedly over time (Baron-Cohen, Harrison, Goldstein, & Wyke, 1993; Eagleman, Kagan, Nelson, Sagaram, & Sarma, 2007). For this reason, the results reported below should be considered an initial investigation pending further research.

Materials

All materials were the same as in Experiments 2 and 3. We used the same list of index words collected from non-synaesthetes in Experiment 1, and the testing apparatus was identical to Experiments 2 and 3[1].

---

[1] We decided to use the same index words collected from non-synaesthetes in Experiment 1 because we wanted to compare to the results for non-synaesthetes directly, as well as avoid any effects of bidirectionality (cf Weiss, Kalckert, & Fink, 2009). That is, if synaesthetes experience red for *A*, that automatic experience may make them more likely to choose a red item as an index word for *A*. Although this is also a possibility in assigning colours to index words, we believe this is less likely to pose a problem. First, the instructions explicitly asked participants to form a mental image and report the most dominant colour in that image (see Experiment 2, above, for details). This requires participants to focus their attention on the concept, not the word itself, and would likely lessen any impact of word-based synaesthetic colour on their index word-colour choice (Mattingley, 2009). As synaesthetes have been shown to have enhanced mental imagery (e.g. Barnett & Newell, 2008; Price, 2009), we believe they would be particularly *good* at this task, rather than influenced unduly by

Procedure

All participants were recruited and tested using MTurk and Qualtrics as reported in Experiment 2 and 3, above. There was no explicit indication in either task that synaesthesia was of interest until the last question, the self-report of synaesthesia. Therefore, self-reported synaesthetes and non-synaesthetes were tested together, and only separated into groups using their response to the synaesthesia question after they had completed the experiments.

Results

Self-reported prevalence of synaesthesia

The self-report of synaesthesia questions allowed us, first, to take an informal measure of the prevalence of self-reported grapheme-colour synaesthesia in a random sample of American English speakers. For both experiments, we compared the proportion of self-reported synaesthetes (i.e. "Yes" answers) to non-synaesthetes (i.e. "No" answers) within English-speaking Americans; participants responding "Don't Know" (Experiment 2 only) were excluded. For Experiment 2, there was a synaesthesia prevalence of 16.44% (24/170 total), and for Experiment 3 a prevalence of 27.20% (65/239 total). Combining both tasks, the overall prevalence of self-reported synaesthesia was 21.76% (89/409 total). This is an unexpectedly high proportion, more than four times the 4.9% prevalence of self-reported grapheme-colour synaesthesia found by Carmichael, Down, Shillcock, Eagleman, and Simner (2015), and well beyond their estimate of 1-2% of grapheme-colour synaesthesia in the general population. However, we note an important caveat with this measure beyond self-report. MTurk requires a description of the task before workers decide to accept. For Experiment 3, this was "You will provide color associations for a list of letters," which might have attracted people with synaesthetic experiences to take part. More informatively, we can compare the gender ratio of those who report synaesthesia to those who did not. Combining both experiments, the synaesthetes were 57.3% female, whereas the non-synaesthetes were 51.1% female. A chi-square test showed that there was no significant difference in gender between synaesthetes and non-synaesthetes (χ^2 (1) = 1.31, *p*

---

synaesthetic colours. Finally, we excluded any participants who consistently reported unexpected or unusual colours in this task (see Data Validation), so we are confident that our self-reported synaesthete participants completed this task just as well as non-synaesthetes.

= .253). This result agrees with recent work indicating that there is no gender bias in synaesthesia (Simner & Carmichael, 2015). We will return to our recruitment method in the general discussion.

Data validation

Using the same data validation procedure as outlined in Experiment 2 and 3, we excluded participants who had not completed the task as instructed. This led to the exclusion of five participants from Experiment 2, for a final sample of 19; and five participants in Experiment 3, for a final sample of 59 and a final grand total of 78.

Do index words predict letter colours for self-reported synaesthetes?

Using the same index word agreement percentages from Experiment 1 and the colours reported by the first group of 19 self-reported synaesthetes, we calculated a colour score to predict the colour for each letter using the index word route. As described in detail in Experiment 2, above, this score represents the colour distribution for each letter that is accounted for by the colours of its top three index words, with the highest score indicating the most dominant colour for each letter. Appendix 2 details the highest colour score and the associated dominant colour for each word.

The results demonstrate that self-reported synaesthetes and non-synaesthetes performed this task very similarly. This is likely due to the fact that index-word → colour associations are based in real prototypical colour (e.g. *dogs* are brown). However, predicted colour did differ for six letters: *G, H, J, K, S,* and *T.* Next, we calculated the highest agreement modal colour as well as significantly associated colours for each letter (see Experiment 2, above, for details). These results are compared to those in previous studies for synaesthetes in Table 5, below.

Table 5. Comparison of the results of letter-colour associations between the current study (far right columns) and previous studies of trends in letter-colour pairings. For comparison, we also include modal colours for non-synaesthetes from Experiment 3 of the current study. Matches between our self-reported synaesthetes and previous studies are highlighted in **bold**.

Colour	Synaesthetes, significance studies	Synaesthetes, Witthoft et al.	Synaesthetes, Current Study	Non-synaesthetes, Current Study
Black	I**XZ**	**XZ**	NSU**XZ**	**XZ**
White	IO	IO	**I**W	HW
Red	**AR**	**AR**	**AR**	**A**FK**R**

Orange	J	HJK**N**	**N**O	CO
Yellow	**Y**	C**L**S**Y**	E**LY**	H**LY**
Green	-	**EFG**	**EFG**JN	**G**S
Blue	**B**	**B**D**T**W	**B**C**T**	**B**IJ**T**U
Purple	**V**	**PQV**	**QV**	J**PV**
Pink	P	-	K**P**	Q
Brown	**D**	-	**D**HM	**D**MN
Grey	X	-	-	E

We also calculated the letter-colour pairs that had a significant binomial distribution. The results are summarised as before in Table 6, below.

Table 6. For each colour term, the letters are listed which were significantly associated with that colour at least at $p < .05$ level in our self-declared synaesthete participants' responses. Letters appear more than once if they were significantly associated with more than one colour. For each associated colour, letters are listed in separate columns if they were significant at $p < .01$ (middle column) or $p < .001$ (right column).

Colour	Significant at $p < .05$	Significant at $p < .01$	Significant at $p < .001$
Black	D S U X	Z	
White	H	I	W
Red			A R
Orange			O
Yellow	E	L	Y
Green	E		G
Blue	D I	T	B
Purple	Q Z		P V
Pink		I K	P
Brown	B T	D H M	
Grey	J	G	

As with non-synaesthetes, the correlation between colour score and colour agreement was significant for the 18 letters that are not the first letters of colour terms (Spearman's $\rho = .533$, $p = .023$). This correlation is illustrated in figure 7, below.

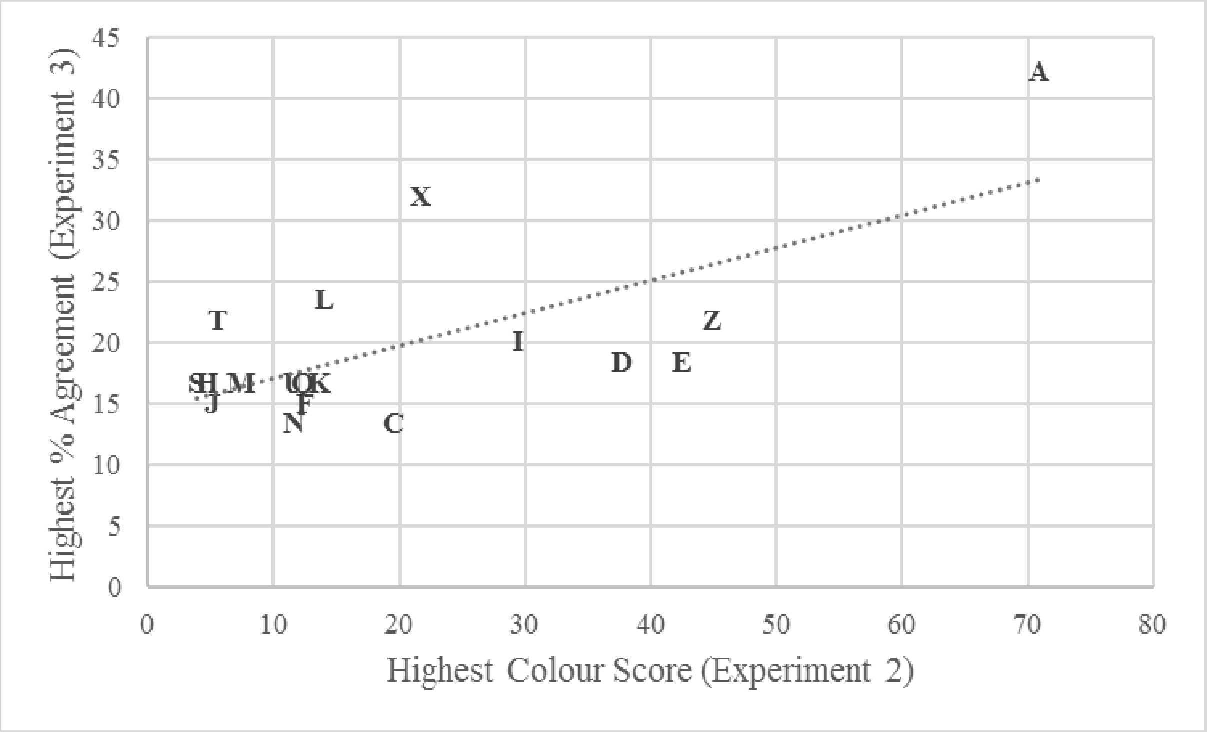
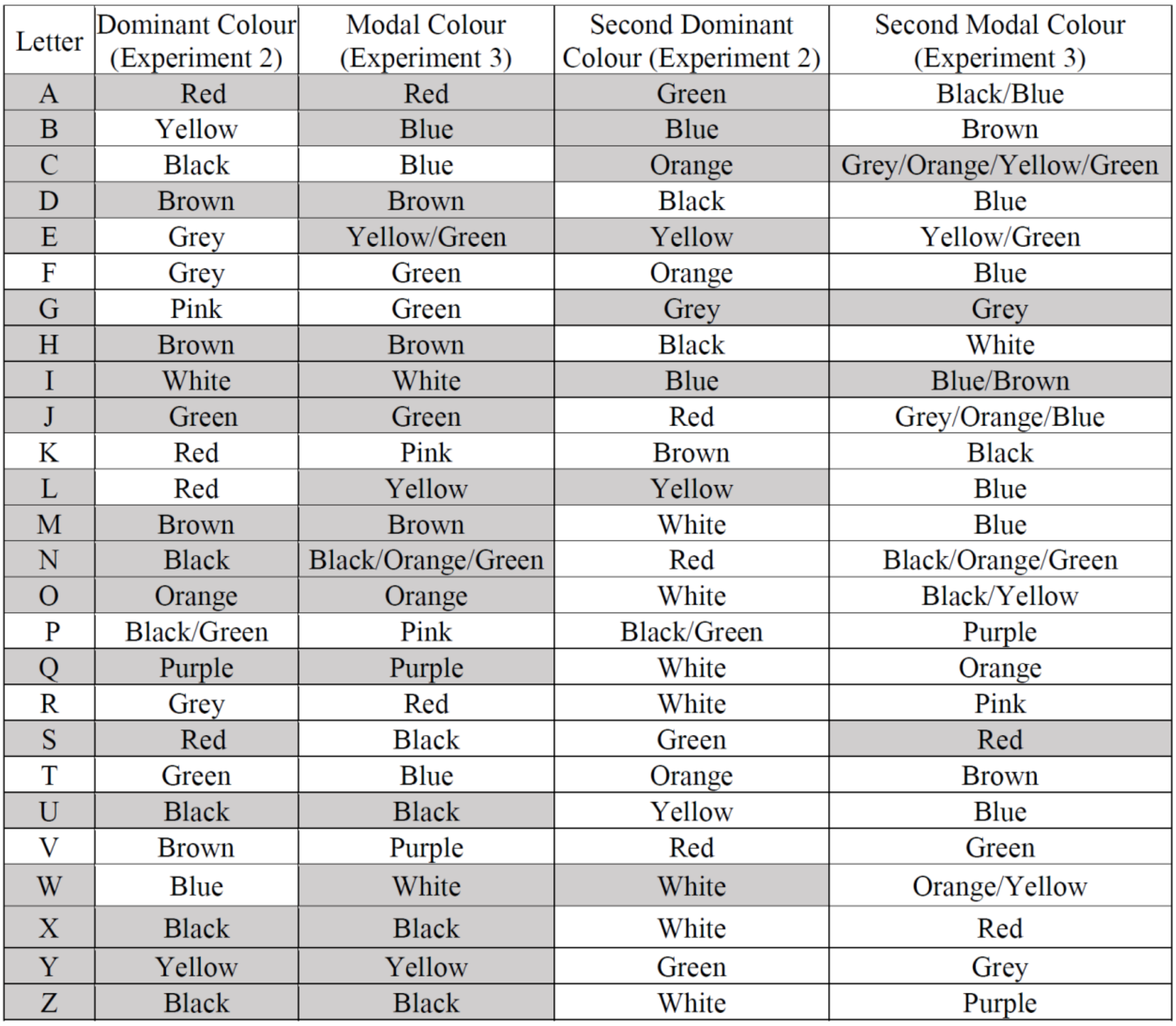


Figure 7. Scatterplot showing the correlation between the highest colour score obtained via index words and the highest agreement reported for direct colour association for self-reported synaesthete participants. Each of the 18 letters included in this correlation represents its respective point.

As summarised in figure 8, below, index words successfully predicted letter colour associations for 20 letters (76.9% of all letters), with 21 matches overall between colour score and agreement out of a possible 52 (maximum two matches per letter); this is a hit rate of 40.38%. We evaluated the probability of obtaining this pattern of results as in Experiment 3, using a simulation of matching two pairs of randomly generated colours for 26 letters, iterated 10000000 times. We again conducted two of these Monte Carlo simulations. The first counted the number of iterations in which there was at least one match for 20 out of 26 letters as 8/10000000, or $p < .001$. The second counted the number of iterations in which there were overall 21 matches out of 52 possible matches across the entire alphabet and found 62/10000000, or $p < .001$. This indicates, as with synaesthetes, that the matches between index-word-predicted dominant colours and directly-associated modal colours are very unlikely to be coincidental. Rather, these consistent patterns suggest that letter-colour associations are influenced by semantic colours transferred to the letters via the index words associated with those letters during literacy acquisition.



Letter	Dominant Colour (Experiment 2)	Modal Colour (Experiment 3)	Second Dominant Colour (Experiment 2)	Second Modal Colour (Experiment 3)
A	Red	Red	Green	Black/Blue
B	Yellow	Blue	Blue	Brown
C	Black	Blue	Orange	Grey/Orange/Yellow/Green
D	Brown	Brown	Black	Blue
E	Grey	Yellow/Green	Yellow	Yellow/Green
F	Grey	Green	Orange	Blue
G	Pink	Green	Grey	Grey
H	Brown	Brown	Black	White
I	White	White	Blue	Blue/Brown
J	Green	Green	Red	Grey/Orange/Blue
K	Red	Pink	Brown	Black
L	Red	Yellow	Yellow	Blue
M	Brown	Brown	White	Blue
N	Black	Black/Orange/Green	Red	Black/Orange/Green
O	Orange	Orange	White	Black/Yellow
P	Black/Green	Pink	Black/Green	Purple
Q	Purple	Purple	White	Orange
R	Grey	Red	White	Pink
S	Red	Black	Green	Red
T	Green	Blue	Orange	Brown
U	Black	Black	Yellow	Blue
V	Brown	Purple	Red	Green
W	Blue	White	White	Orange/Yellow
X	Black	Black	White	Red
Y	Yellow	Yellow	Green	Grey
Z	Black	Black	White	Purple

Figure 8. Comparison between dominant colour predictions via colour score from Experiment 2 and modal letter-colour agreement from Experiment 3 for self-reported synaesthete participants. Shaded cells indicate a match between dominant and modal colour(s). Cells with more than one colour indicate a tie in agreement.

Discussion

In this experiment, we 1) calculated a prevalence and gender ratio of self-reported synaesthesia in a sample of American English speakers, 2) calculated the colour scores for each letter and 3) gathered letter-colour associations for these self-reported grapheme-colour synaesthetes. This allows us to evaluate the index word route as an influence on trends in letter-colour pairings.

Although the total number of matches was lower for self-reported synaesthetes, the number of letters for which there was at least one match was higher. We note that while seven letters had a match between first dominant colour and first modal colour for non-synaesthetes, this first-order match accounted for over half of the matches for synaesthetes. Self-reported synaesthetes also had only one double match (for *I*), while non-synaesthetes had five (for *I, K, S, W,* and *Z*). This, as well as the much smaller number of significant letter-colour matches for synaesthetes – 34 for synaesthetes versus 59 for non-synaesthetes – may be due to the much lower numbers of synaesthete participants for both experiments, but particularly Experiment 2 (index words to colours). Even given the small number of synaesthetes, it is striking that index words were able to predict colour associations for the majority of the letters of the alphabet. The nuances and implications of these findings will be explored in more depth below.

General Discussion

This study has quantified for the first time the influence of *index words* on the development of letter-colour associations in English (i.e. words starting with a particular letter that are strongly associated with that letter; e.g., *A* is for apple). We examined letter-colour associations that have been previously reported in the literature and also elicited our own responses from American self-reported synaesthetes and non-synaesthetes. We first gathered the set of words that were most commonly associated with each letter of the alphabet and identified the top three for each letter as *index words* (e.g., *A* is for *apple, animal,* and *aardvark*). We then asked what colour was prototypically associated with each index word (e.g., apples are red). Next we calculated a *colour score* predicting the dominant colours for each letter. This colour score combined the percentage agreement from letters to index words, and the percentage agreement from index words to colours. We then compared the colours predicted by index words to the colours reported in direct letter-colour associations. For both synaesthetes and non-synaesthetes, we found that the dominant and/or second-dominant colour (i.e. the colours with the highest and second-highest colour score for each letter; for *A*, red and green respectively) matched the most common direct letter-colour associations for 17 out of 26 letters for non-synaesthetes, and 20 out of 26 letters for self-reported synaesthetes, a rate much higher than chance. This is the first indication that orthographic-semantic associations (beyond red for *R*; e.g., Rich et al., 2005; Simner et al., 2005) have some measurable influence on the development of these population-wide letter-colour pairings. We will address the implications of these findings in order, beginning with a discussion of the nature of index words themselves and their implications for literacy acquisition. We follow this with a discussion of the impact of index words on letter-colour associations in synaesthetes and non-synaesthetes.

Index words in literacy learning

Despite the large body of work on children's acquisition of the alphabet, there is little data on the index words that accompany alphabet instruction. Worden and Boettcher (1990) found that children's success at naming a word beginning with a particular letter (i.e. naming what we have termed *index words*) followed the same pattern as other measures of alphabet literacy, increasing steadily with age, but it was the most difficult of their alphabet knowledge tasks. In that study it was not clear whether children did indeed acquire particular index words, despite training in these associations via alphabet books, because Worden and Boettcher (1990) only recorded the number of letters for which the children could name an index word (i.e. their rate of successful naming), but not which words they produced. The impact of index words in literacy acquisition is also not clear, since children often fail to connect the spelling, sounds, and meanings of words, even when explicitly coached by their parents using alphabet books (Davis, Evans, & Reynolds, 2010). However, there is abundant evidence that alphabet books do promote alphabet learning (e.g. Both-de Vries & Bus, 2014; Brabham, Murray, & Bowden, 2006; Evans, Saint-Aubin, & Landry, 2009; Murray, Stahl, & Ivey, 1996; Nowak, 2015). Despite this apparent conflict between the efficacy of alphabet-book-based literacy training and the actual ability of children to produce index words, little research has addressed whether index words as such have any direct influence on the acquisition of alphabet or general literacy, or indeed whether they have any enduring connection with individual letters. This study attempted to provide a first indication of this connection into adulthood by showing that some particular letters do have index words, with very high levels of agreement.

Index words in determining letter-colour associations

Our results suggest that index words are one influence (among many; see Introduction) on letter-colour pairings in synaesthetes and non-synaesthetes. Our results show that not only can index words predict the matching of letters with colours, but they can also explain *why* particular colour-letter pairs consistently recur across large groups of both synaesthetes and non-synaesthetes (e.g., why *A* tends to be red but not blue). We will here discuss the influence of index words on letter-colour associations for different groups of letters that appear to show similar patterns: the initial letters of colour terms; the letters for which index words provide a new explanation for consistent colour association trends; and the letters for which we have not yet been able to identify an index-word-based source.

First, we will look at the initial letters of colour terms, namely *B, W, R, O, Y, G, V,* and *P*. On the one hand, it is tempting to disregard the consistent colour association between *R* and red, *B* and blue, etc., since they are obviously the first letters of basic colour terms. However, the relationship between initial letter and colour term is not quite so straightforward – for instance, *B* is consistently associated with blue for both groups, even though *black* is a more frequent colour term (although we note that *black* was the second highest directly associated colour term for non-synaesthetes). The same conflict applies to *P* for *pink/purple* and *G* for *green/gray*. We suggest that *P* may be associated with *pink* because of the typical association of *V* with purple (via *violet*), and *G* with *green* due to higher frequency. More fundamentally, however, colour terms are a clear example of the first letter of a particular word becoming associated with that word's colour due to an orthographic-semantic connection. While this is not exactly the same type of index word influence as we have explored above, it is still a linguistic connotation that is fossilised into an automatic, explicit colour association for synaesthetes, and an intuitive colour association for non-synaesthetes. In other words, for synaesthetes, the colours denoted by colour terms have become indelibly associated with their initial letters past the point of conscious association and into automatic perception (e.g. Dixon, Smilek, & Merikle, 2004; Gray et al., 2006; Mattingley, Rich, Yelland, & Bradshaw, 2001; Mills, Boteler, & Oliver, 1999). However, synaesthetes who have these colour-term associations must form them *after* reaching the realisation that graphemes represent a series of independent sounds which, taken together, can create a word. As Nodelman (2001) summarises, understanding and appreciating the meaning of a phrase like *A is for apple* is a complex process that children struggle with, even when they are explicitly instructed in it. A synaesthete child must realise that the symbol *B* makes the phonetic sound /b/, that the sequence B + L + U + E represents the word *blue*, and connect this abstract representation of the word *blue* with semantic knowledge of the colour blue, before they can form a connection between *B* and blue synaesthetically. Therefore, even colour-term-based synaesthetic colours are necessarily fundamentally rooted in the literacy acquisition process.

The second group of letters to consider are those that have consistent colour associations not explicable by colour terms, but which could nonetheless be predicted from index words such as *apple*. For non-synaesthetes, index-word-predicted dominant colour matched directly-associated modal colour for seven letters, *A, D, E, M, S, X,* and *Z,* and for self-reported synaesthetes for eleven letters, *A, D, H, I, J, M, N, Q, U, X,* and *Z.* The number of matches in each subject-group show that index words have a particularly strong influence on synaesthetes, especially since synaesthetes had fewer significant letter-colour pairs overall: 34 for synaesthetes versus 59 for non-synaesthetes. This strong influence might stem from the nature of the development of synaesthetic associations. A child with a genetic predisposition to synaesthesia may find the explicit link between letters and words with strong colours, as alphabet books often present them, to be a compelling formative influence during literacy, as they are predisposed to develop such connections (see Brang & Ramachandran, 2011). Meanwhile, the non-synaesthete child may learn these letter-word-colour connections, but as the associations are less salient, the influence of index words on colour associations may be somewhat diminished as there is no explicit perceptual experience of colour with letters.

The final group of letters are those for which we could not explain the systematic colour trends: *F, K,* and *T* for synaesthetes, and *F, H, J,* and *N* for non-synaesthetes. For example, if the green of *F* is from, for example, *frog,* we have no evidence of it. The lack of matches could be due to several factors. First, some letters (e.g. *N*) had a lack of agreement in index words, indicating that there was no particular word favoured above all others. Another reason we could not explain certain trends is because there was low agreement in direct letter-to-colour mapping; if a letter has no strongly associated colour, any attempt to predict a colour for that letter will fail. Furthermore, there could also have been a lack of agreement in the colour of index words. Although we found that index words were highly imageable, this does not mean they necessarily have a prototypical colour. To use *F* as an example, although *fish* had high agreement in Experiment 1 (25.4%), it could not predict the direct colour associated with *F* because there was no strong prototypical colour to transfer, as blue, silver/grey, or yellow/gold could all be possibilities. On the other hand, a word like *frog* does have a strong prototypical colour (i.e. green), but it had low agreement among our response words for *F* and was therefore not included as an index word. Finally, we have suggested that index words given by adults likely reflect their early learning from alphabet books, but this may not be the case. Certain low-frequency, low-age-of-acquisition words, like *frog, dragon,* or *jungle*, may appear frequently in alphabet- or storybooks for children but are seldom encountered by adults. Therefore, it may be that the high-imageability, low-age-of-acquisition, strongly prototypically coloured words that are frequently encountered by children may be strong influences on the development of letter-colour associations during childhood, but are too infrequent for adults to have been elicited by our study. The next step, currently underway by our lab, is to collect index words from children during literacy acquisition for a clearer representation of the words that are influential during the period of synaesthetic colour association formation.

We must also briefly address the use of data from self-reported and unverified synaesthetes in our analyses. Carmichael, Down, Shillcock, Eagleman, & Simner (2015) randomly tested a large (N = 2847) population for synaesthesia and found that 4.9% self-reported having synaesthesia, while only 1.2% scored below the threshold for genuineness of <1 using an objective measure (see Carmichael et al., 2015; Eagleman et al., 2007). We can therefore estimate that approximately 25% of our self-referred synaesthetes would be confirmed as objectively genuine. We point out that our group of self-reported synaesthetes almost certainly did contain some genuine synaesthetes; while we were unable to test them to ascertain exactly how many, the verification of genuine synaesthesia is an ongoing question in the field, and the number of “true” synaesthetes in our sample would vary depending on which definition we used (see Carmichael et al., 2015; Rothen, Seth, Witzel, & Ward, 2013). Our results nevertheless show that the two groups exhibited commonalities in their letter-colour pairings in line with the trends in these associations that we set out to investigate, which we attribute in some part to the shared influence of index word associations in childhood literacy acquisition. We plan to expand these promising preliminary results with verified synaesthetes in the future.

Our study has provided evidence to support one influence on letter-colour associations that has often been suggested previously. Index words are particularly useful as they allow us to explain *why* particular letter-colour pairs recur. Besides the oft-cited *A is for apple*, we can also now suggest that, for example, *D* is brown because of *dog* and *Q* is purple because of *queen*. A clear way to test these semantic influences on grapheme-colour trends would be to conduct a similar study as that described above in non-English languages. Our favoured example in this study, *apple*, happens to sit at the intersection of index word, red colour, and first letter, so it is difficult to establish which way the direction of influence runs: is *apple* for A because A is already red, or does the influence of *apple* help lend *A* its redness? In Spanish, for example, *A* may be for *agua* "water" or *árbol* "tree" (*apple* is ineligible, as it is *manzana* in Spanish), in which case we might find that Spanish speakers may be more likely to attribute blue or green to *A* rather than red. Such an investigation could also ask whether the ease-of-colour-generation effects are stronger than these semantic influences.

In conclusion, this study has suggested that both self-reported synaesthetes and non-synaesthetes are influenced in the formation of letter-colour associations by the prototypical colour of index words for each letter. We propose that this semantic influence, rooted in the process of literacy acquisition, works in conjunction with other salient linguistic factors to form lifelong associations between letters and colours. The agreement across a population in index words can explain both why particular letters and colours are paired, and why these associations occur in both synaesthetes and non-synaesthetes. Ultimately, this indicates that grapheme-colour synaesthesia is not a random pairing across modalities, but is rooted in literacy and language-based systematicity.

Acknowledgements

The research leading to these results has received funding for author JS from the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement n. [617678]. The authors would like to thank Prof. Jamie Ward for his help with our research.

References

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). The CELEX lexical database {CD-ROM}.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., … Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*(3), 445–459.

Bankieris, K., & Simner, J. (2015). What is the link between synaesthesia and sound symbolism? *Cognition*, *136*, 186–95.

Barnett, K. J., & Newell, F. N. (2008). Synaesthesia is associated with enhanced, self-rated visual imagery. *Consciousness and Cognition*, *17*(3), 1032–9.

Baron-Cohen, S., Burt, L., Smith-Laittan, F., Harrison, J., & Bolton, P. (1996). Synaesthesia: prevalence and familiality. *Perception*, *25*(9), 1073–9.

Baron-Cohen, S., Harrison, J., Goldstein, L. H., & Wyke, M. A. (1993). Coloured speech perception: is synaesthesia what happens when modularity breaks down? *Perception*, *22*(4), 419–26.

Beeli, G., Esslen, M., & Jäncke, L. (2007). Frequency correlates in grapheme-color synaesthesia. *Psychological Science*, *18*(9), 788–92.

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers : A Journal of the Psychonomic Society, Inc*, *33*(1), 73–79.

Both-de Vries, A. C., & Bus, A. G. (2014). Visual processing of pictures and letters in alphabet books and the implications for letter learning. *Contemporary Educational Psychology*, *39*(2), 156–163.

Brabham, E. G., Murray, B. a., & Bowden, S. H. (2006). Reading Alphabet Books in Kindergarten: Effects of Instructional Emphasis and Media Practice. *Journal of Research in Childhood Education*, *20*(February 2015), 219–234.

Brang, D., & Ramachandran, V. S. (2011). Survival of the synesthesia gene: why do people hear colors and taste words? *PLoS Biology*, *9*(11), e1001205.

Brang, D., Rouw, R., Ramachandran, V. S., & Coulson, S. (2011). Similarly shaped letters evoke similar colors in grapheme-color synesthesia. *Neuropsychologia*, *49*(5), 1355–8.

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–90.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon’s Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspectives on Psychological Science*, *6*(1), 3–5.

Carmichael, D. A., Down, M. P., Shillcock, R. C., Eagleman, D. M., & Simner, J. (2015). Validating a standardised test battery for synesthesia: Does the Synesthesia Battery reliably detect synesthesia? *Consciousness and Cognition*, *33*, 375–385.

Cytowic, R. E. (1989). *Synaesthesia: A Union of the Senses*. New York: Springer Verlag.

Cytowic, R. E., & Wood, F. B. (1982). Synesthesia. II. Psychophysical relations in the synesthesia of geometrically shaped taste and colored hearing. *Brain and Cognition*, *1*(1), 36–49.

Davis, B. J., Evans, M. A., & Reynolds, K. P. (2010). Child miscues and parental feedback during shared alphabet book reading and relations with child literacy skills. *Scientific Studies of Reading*, *14*(4), 341–364.

Dixon, M. J., Smilek, D., & Merikle, P. M. (2004). Not all synaesthetes are created equal: projector versus associator synaesthetes. *Cognitive, Affective & Behavioral Neuroscience*, *4*(3), 335–43.

Eagleman, D. M., Kagan, A. D., Nelson, S. S., Sagaram, D., & Sarma, A. K. (2007). A standardized test battery for the study of synesthesia. *Journal of Neuroscience Methods*, *159*(1), 139–45.

Evans, M. A., Saint-Aubin, J., & Landry, N. (2009). Letter names and alphabet book reading by senior kindergarteners: An eye movement study. *Child Development*, *80*(6), 1824–1841.

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation*, *12*(4), 395–427.

Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data Collection in a Flat World: The Strengths and Weaknesses of Mechanical Turk Samples. *Journal of Behavioral Decision Making*, *26*(3), 213–224.

Gray, J. A., Parslow, D., Brammer, M., Chopping, S., Vythelingum, G., & Fytche, D. (2006). Evidence Against Functionalism from Neuroimaging of the Alien Colour Effect in Synaesthesia. *Cortex*, *42*(2), 309–318.

Hancock, P. (2013). Synesthesia, alphabet books, and fridge magnets. In J. Simner & E. M. Hubbard (Eds.), *The Oxford Handbook of Synesthesia* (pp. 83–102). Oxford University Press.

Jonas, C. N. (2010). *Effects of synaesthetic colour and space on cognition*.

Jonas, C. N., Spiller, M. J., Jansari, A., & Ward, J. (2014). Comparing implicit and synaesthetic number-space associations: Visuospatial and verbal spatial-numerical associations of response codes. *Quarterly Journal of Experimental Psychology (2006)*, *67*(7), 1262–73.

Kučera, H., & Francis, W. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Lavrynenko, O. (2014). A study of the cognitive synesthetic mechanism as a supplement to the sound imitation theory of lingual genesis.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers*, *28*, 203–208.

Mankin, J. L. (n.d.). The psycholinguistics of synaesthesia: What can we learn from studying synaesthetes? *PsyPAG Quarterly*.

Marks, L. E. (1975). On colored-hearing synesthesia: Cross-modal translations of sensory dimensions. *Psychological Bulletin*, *82*(3), 303–331.

Mattingley, J. B. (2009). Attention, automaticity, and awareness in synesthesia. *Annals of the New York Academy of Sciences*, *1156*, 141–67.

Mattingley, J. B., Rich, a N., Yelland, G., & Bradshaw, J. L. (2001). Unconscious priming eliminates automatic binding of colour and alphanumeric form in synaesthesia. *Nature*, *410*(6828), 580–2.

Mills, C. B., Boteler, E. H., & Oliver, G. K. (1999). DIGIT SYNAESTHESIA : A CASE STUDY USING A, *16*(November 1995), 181–191.

Murray, B. A., Stahl, S. A., & Ivey, M. G. (1996). Developing phoneme awareness through alphabet books. *Reading and Writing*, *8*(4), 307–322.

Nagai, J., Yokosawa, K., & Asano, M. (2015). Biases and regularities of grapheme–colour associations in Japanese nonsynaesthetic population. *The Quarterly Journal of Experimental Psychology*, *0*(0), 1–13.

Nodelman, P. (2001). A is for... What? The Function of Alphabet Books. *Journal of Early Childhood Literacy*, *1*(3), 235–253.

Nowak, S. N. (2015). *ABCs by Design: The Role of Alphabet Book Design and Children's Alphabetic*

Behaviours in Emergent Literacy Skill Acquisition. Unversity of Guelph.

Oldfield, R. C., & Wingfield, a. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, *17*(4), 273–281.

Price, M. C. (2009). Spatial forms and mental imagery. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *45*(10), 1229–45.

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*(1), 53–84.

Rothen, N., Seth, A. K., Witzel, C., & Ward, J. (2013). Diagnosing synaesthesia with online colour pickers: maximising sensitivity and specificity. *Journal of Neuroscience Methods*, *215*(1), 156–60.

Rouw, R., Case, L., Gosavi, R., & Ramachandran, V. (2014). Color associations for days and letters across different languages. *Frontiers in Psychology*, *5*(MAY).

Sagiv, N., & Ward, J. (2006). Crossmodal interactions: lessons from synesthesia. *Progress in Brain Research*, *155*(2001), 259–71.

Simner, J. (2007). Beyond perception: synaesthesia as a psycholinguistic phenomenon. *Trends in Cognitive Sciences*, *11*(1), 23–9.

Simner, J. (2012). Defining synaesthesia. *British Journal of Psychology (London, England : 1953)*, *103*(1), 1–15.

Simner, J., & Carmichael, D. A. (2015). Is synaesthesia a dominantly female trait? *Cognitive Neuroscience*, 1–9.

Simner, J., Glover, L., & Mowat, A. (2006). Linguistic determinants of word colouring in grapheme-colour synaesthesia. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *42*(2), 281–9.

Simner, J., & Ward, J. (2008). Synaesthesia, color terms, and color space: Color claims came from color names in Beeli, Esslen, and Jäncke (2007). *Psychological Science*, *19*(4), 412–414.

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. *Cognitive Neuropsychology*, *22*(8), 1069–85.

Smilek, D., Carriere, J. S. a, Dixon, M. J., & Merikle, P. M. (2007). Grapheme frequency and color luminance in grapheme-color synaesthesia. *Psychological Science*, *18*(9), 793–5.

Spector, F., & Maurer, D. (2008). The colour of Os: Naturally biased associations between shape and colour. *Perception*, *37*(6), 841–847.

Spector, F., & Maurer, D. (2009). Synesthesia: a new approach to understanding the development of perception. *Developmental Psychology*, *45*(1), 175–89.

Spector, F., & Maurer, D. (2011). The colors of the alphabet: naturally-biased associations between shape and color. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(2), 484–95.

Stadthagen-Gonzale, H., & Davis, C. J. (2006). The Bristol norms for Age of Acquisition, Imageability, and Familiarity. *Behavior Research Methods*, *38*(4), 598–605.

Ward, J., Banissy, M. J., & Jonas, C. N. (2008). Haptic perception and synaesthesia. In M. Grunwald (Ed.), *Human Haptic Perception: Basics and Applications* (pp. 259–265). Basel: Birkhäuser Basel.

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

Ward, J., & Mattingley, J. B. (2006). Synaesthesia: an overview of contemporary findings and controversies. *Cortex*, *42*, 129–136.

Ward, J., Simner, J., & Auyeung, V. (2005). A comparison of lexical-gustatory and grapheme-colour synaesthesia. *Cognitive Neuropsychology*, *22*(1), 28–41.

Weiss, P. H., Kalckert, A., & Fink, G. R. (2009). Priming letters by colors: evidence for the bidirectionality of grapheme-color synesthesia. *Journal of Cognitive Neuroscience*, *21*(10), 2019–26.

Witthoft, N., Winawer, J., & Eagleman, D. M. (2015). Prevalence of Learned Grapheme-Color Pairings in a Large Online Sample of Synesthetes. *PLOS ONE*, *10*(3), e0118996.

Worden, P., & Boettcher, W. (1990). Young children’s acquisition of alphabet knowledge. *Journal of Literacy Research*, *22*(3), 277–295.

Appendix 1. For each letter, the number of unique items given; the top three most common index words; and the percentage agreement for those index words. Index words with higher than 20% agreement are *italicised.*

Letter	**Index Word**	**Agreement (%)**	**Letter**	**Index Word**	**Agreement (%)**
A	*apple*	*84.13*	N	no	11.43
	animal	1.9		night	6.67
	aardvark	0.95		nose	6.03
B	banana	15.87	O	orange	16.51
	boy	13.33		open	12.06
	bear	12.38		octopus	11.75
C	*cat*	*49.21*	P	pie	3.17
	car	10.48		pear	2.86
	cookie	2.86		penguin	2.86
D	*dog*	*57.14*	Q	*queen*	*31.11*
	dad	4.76		quiet	8.89
	door	2.54		question	7.3
E	*elephant*	*46.35*	R	run	6.35
	egg	8.57		rest	5.08
	eagle	4.76		rat	4.76
F	*fish*	*25.4*	S	snake	8.25
	fox	8.25		sister	5.71
	food	5.71		stop	3.81
G	goat	9.84	T	time	8.57
	girl	8.25		tiger	5.71
	giraffe	7.3		turtle	5.4
H	hat	11.43	U	*umbrella*	*22.86*
	house	6.67		under	16.19
	happy	6.35		up	5.4
I	*igloo*	*20.95*	V	victory	11.75
	ice	11.75		vendetta	8.89
	ice cream	5.71		violin	8.57
J	jump	14.92	W	water	17.46
	joke	9.21		wax	3.81
	jack	7.94		whale	3.81
K	*kite*	*23.81*	X	*xylophone*	*51.43*
	kangaroo	11.75		xray	27.94
	king	6.98		xerox	3.81
L	*love*	*23.17*	Y	*yellow*	*24.13*
	lion	11.43		yes	13.65
	laugh	3.49		you	8.25
M	mom	10.79	Z	*zebra*	*60.63*
	man	8.89		zoo	19.68
	monkey	8.25		zero	3.49

Appendix 2. Summary of the dominant colours predicted for each letter using the letter → index word → colour route. For each letter, the dominant colour (defined as the colour within each letter with the highest colour score) is reported along with its colour score for non-synaesthetes (Experiment 3) and self-reported synaesthetes (Experiment 4). Where the dominant colour differs between the two groups, this is highlighted in the synaesthete column in *italics*.

	Non-synaesthetes		*Synaesthetes*	
	Dominant colour	Colour score	Dominant colour	Colour score
A	Red	76.61	Red	70.89
B	Yellow	15.93	Yellow	15.74
C	Black	18.82	Black	19.55
D	Brown	42.52	Brown	37.69
E	Grey	46.06	Grey	42.47
F	Grey	11.47	Grey	12.46
G	White	5.82	*Pink*	6.47
H	Black	4.49	*Brown*	4.66
I	White	33.03	White	29.47
J	Red	4.59	*Green*	5.06
K	Brown	10.26	*Red*	13.63
L	Red	19.12	Red	13.97
M	Brown	8.85	Brown	7.42
N	Black	12.21	Black	11.63
O	Orange	17.83	Orange	17.78
P	Black	2.56	Black	2.11
Q	Purple	10.57	Purple	12.31
R	Grey	4.33	Grey	5.28
S	Green	3.93	*Red*	3.91
T	Orange	5.2	*Green*	5.63
U	Black	17.78	Black	11.63
V	Brown	7.74	Brown	8.29
W	Blue	16.11	Blue	13.15
X	Black	26.46	Black	21.72
Y	Yellow	25.99	Yellow	26.43
Z	Black	45.13	Black	44.85