

Opinion piece



Cite this article: Mankin JL. 2019 Deepening understanding of language through synaesthesia: a call to reform and expand. *Phil. Trans. R. Soc. B* **374**: 20180350. <http://dx.doi.org/10.1098/rstb.2018.0350>

Accepted: 8 April 2019

One contribution of 16 to a discussion meeting issue ‘Bridging senses: novel insights from synaesthesia’.

Subject Areas:

cognition

Keywords:

assumptions, psycholinguistics, synaesthesia, colour measurement

Author for correspondence:

Jennifer L. Mankin

e-mail: j.mankin@sussex.ac.uk

Deepening understanding of language through synaesthesia: a call to reform and expand

Jennifer L. Mankin

School of Psychology, University of Sussex, Brighton BN1 9QH, UK

JLM, 0000-0002-2229-2791

In this paper, I present arguments and suggestions for the improvement of the scientific study of synaesthesia, and particularly grapheme-colour synaesthesia in relation to psycholinguistic research, although the principles I advocate can be easily adapted to any subfield of synaesthesia study. I postulate that the current state of research on synaesthesia in general, and on grapheme-colour synaesthesia in particular, suffers from a lack of exploratory evidence and essential groundwork upon which to build hypothesis-testing studies. In particular, I argue that synaesthesia research has been artificially bounded by assumptions about the nature of synaesthetic experiences, which constrain both the questions that researchers ask and the way in which they go about answering those questions. As a specific example, I detail how much of the current research on grapheme-colour synaesthesia is built to accommodate two major assumptions about the nature of colours for letters and for words—assumptions which I will contend are not universally true, and the exceptions to which point to a much richer and heterogeneous understanding of synaesthetic experience than current research practices capture. The top-down predetermination of what is important or meaningful to measure, and what is not, has subsequently impeded a full understanding of what synaesthesia is and how it works. I argue that these assumptions must be carefully addressed and evaluated, both for the particular case of grapheme-colour synaesthesia and for the field as a whole, to move towards a holistic and fruitful understanding of synaesthesia as a phenomenon and as a tool to study language, thought and perception. To that end, I propose specific recommendations for synaesthesia researchers to solidify and expand their understanding and to capture the actual experience of synaesthetes.

This article is part of a discussion meeting issue ‘Bridging senses: novel insights from synaesthesia’.

1. Introduction

In the last few decades, research interest has intensified in the phenomenon of synaesthesia, which is characterized by unusual and automatically evoked experiences (termed *concurrents*, such as a specific colour like blood red) in response to everyday stimuli (termed *inducers*, such as a letter like *A*; [1,2]. Grapheme-colour synaesthesia, in particular, experiencing automatic impressions of colour for letters and/or words, represents more than a quarter of all cases of synaesthesia, or about 1% of the general population [3–6]. Although still uncommon, this puts the number of grapheme-colour synaesthetes worldwide in the tens of millions. Grapheme-colour synaesthesia, therefore, presents a striking opportunity to study human perception and cognition using these extraordinary colour experiences.

Case studies conducted early on in the modern scientific investigation of synaesthesia focused primarily on describing the complex and varying experiences that a synaesthete might have. One widely cited early study [7] documented ‘coloured-hearing’ synaesthete EP’s experiences. Besides formalizing test-retest consistency over time as the main method of establishing the genuineness of

synaesthesia, this paper tested EP's associations with a range of stimuli, including words of various semantic categories, names and individual letters. Similarly, subsequent studies and reports sought to describe a variety of synaesthetic experiences across a large sample of the population. Among others, Day [8] describes data and patterns gleaned from hundreds of synaesthetes, even mentioning synaesthetes with 'two...[or] more than two' colours for a single letter (p. 13). Rich *et al.* [9] also collected synaesthete's descriptions of their experiences induced by letters, digits, days of the week, names and a variety of nouns. In other words, these studies often focused on capturing the range and inherent heterogeneity of synaesthetic experiences as far as possible.

In recent years, interest has grown in the value of grapheme-colour synaesthesia as a tool to investigate the underlying structure of language. In 2007, Simner suggested that language has a 'special' status in synaesthesia, based on the predominance of linguistic items as inducers, and called for more research into the essentially psycholinguistic basis of grapheme-colour synaesthesia. For subsequent research in this area, the goal is not only to better describe and understand synaesthesia but to use these systematic correspondences between language and colour to test theories regarding various linguistic and psycholinguistic processes. Thus far, several areas of psycholinguistic interest have been investigated using synaesthesia, including morphological processing [10–12], visual word recognition [13,14] and grapheme acquisition [15,16]. However, this focus on correspondences between linguistic features of words (such as frequency or morphological structure) and particular features of their synaesthetic colours (such as their hue, saturation and luminance) has resulted in, or at least coincided with, a narrowing of the variables collected to investigate these patterns. It may be no coincidence that this keen interest in the implications of synaesthetic colours for psycholinguistic theory was much invigorated after the development of the Synesthesia Battery [17]. The grapheme-colour test in the Battery made the quick online verification of synaesthetes possible for large-scale, detailed data collection, and simultaneously defined a real synaesthete specifically by their consistency in selecting one colour per capital letter on an HTML colour palette. Perhaps, as a side effect, it also concentrated researchers' interest specifically on synaesthetes with colours for written letters, rather than for spoken words (as with EP) or even necessarily for whole written words (as some synaesthetes have reported, e.g. [18]). In order to develop this measure—which, to be clear, has been invaluable for both researchers and synaesthetes—operationalization decisions were made that disregarded many of the qualitative details that earlier, less streamlined studies had documented. With this paradigm widely accepted as the predominant testing method for synaesthesia research, much of the variety of the synaesthetic experience is no longer measured by default in favour of a single, easily captured colour measure.

It is beyond the scope of this paper to argue whether these consistency measures are truly representative of what defines a synaesthete (cf. [1,2,19]), but I will argue that this narrowed operationalization of grapheme-colour synaesthesia must be expanded again to account for the inherent idiosyncrasy of synaesthetic experiences. I will suggest that relying on the predominant method of gathering synaesthetic concurrents specifically and only as single colours artificially restricts both the research questions that scientists in this field ask

and the type and breadth of data that they collect to investigate those questions. This paper will first discuss some potential issues with the assumptions that currently underlie research into grapheme-colour synaesthesia and suggest new directions to deepen our understanding.

2. Current assumptions

As research on synaesthesia diversifies, it is becoming clear that some of the assumptions frequently made about grapheme-colour synaesthesia may not be consistently true of all synaesthetes, and may, in fact, obscure the true nature of synaesthetes' experiences. Two of these assumptions, in particular, are challenged by recent results: one about the *number* of colours in a word, and the other about the *nature* of those colours. That is, synaesthesia researchers have frequently assumed that a single colour is sufficient to represent a synaesthete's colour association for a word [9,13,20], and that that single colour is straightforwardly derived from a single letter source within that word, typically the first letter or vowel [14,21]. These assumptions were indeed based on case studies and reports from synaesthetes themselves [7,9,18,22], and these reports strongly influenced the directions of subsequent work. However, recent studies have begun to offer synaesthetes more colour options, the results of which highlight major gaps in the scientific understanding of synaesthetic experiences.

3. Number of colours per word

The first assumption to address is that synaesthetes experience a single colour associated with each word. This widely accepted assumption is supported by circular design: synaesthesia researchers often only allow synaesthetes the option of a single colour per word. No further colours are asked for or accepted, leading to the impression that that one colour comprises the whole-word colour. Having only a single colour per word also simplifies the data processing and analysis of complex colour and linguistic information very substantially, which may also explain why this paradigm is useful to researchers. However, if this single whole-word colour does not, in fact, capture the actual experience of synaesthetes, then it does not matter how much is gained in elegance of analysis, when the underlying operationalization—and therefore the accuracy and generalizability of the results—is restricted from the outset.

Indeed, when given the chance to describe their word colours, case studies document multiple colours: for example, the city name *Catonsville* appealed to synaesthete MLS because of its 'browns and greens and this nice shiny N' [22], and synaesthete MD mentioned that *banana* was yellow although it 'should be dark blue and black' [9]. Simner *et al.* [14] also showed that letters downstream in a word could reinforce the synaesthete's colour experiences (e.g. the synaesthetic colour of *ether* named more quickly than that of *ethos*), implying that multiple colours in a word are pertinent to the whole word's colour. More recently, in a study (as yet unpublished) investigating the link between synaesthetic and canonical colour associations [23], the online test apparatus concluded with an open-answer question about where the colours of whole words came from, and an invitation to write any thoughts or feedback they wanted us to know. Between these two spaces, 16 out of the 20 synaesthetes who completed the test provided written descriptions of their synaesthetic

experiences. The voluntary descriptions of these colour experiences, many of them thoughtful and detailed, paint quite a different picture than a straightforward one-word-one-colour experience. Almost all of the comments describe experiencing multiple colours in a word. Several participants mentioned that the first letter tended to be influential, but are also subject to other influences downstream, as one synaesthete described: 'In the majority of cases, the whole word colour comes from one of the first few letters in the word...Often, others letters in the word affect the overall colour of a word, such as the word bleed—b words are often bright blue (for b), but bleed is very dark blue because of the black of the 'e's...' Several synaesthetes mentioned that the colour selection task, and particularly the option of choosing only one colour, did not reflect their actual experience. One synaesthete said that they experience 'several colours simultaneously—i.e. instead of mixing red and blue to get purple, imagine wearing those 3D glasses that have one blue lens and one red lens, so you're experiencing both colours at the same time. Only being able to choose one colour for a word is difficult and sometimes inaccurate [*sic*].' These detailed descriptions highlight the diversity and complexity of synaesthetes' experiences, as well as evident thoughtfulness about the systematic influences that form a synaesthete's colours.

Here, I note again that the grapheme-colour synaesthetes who are asked to provide single colours for whole words are often identified by the consistency of their colour associations with individual letters. It is not clear that synaesthetes with this subtype of synaesthesia can always reasonably be expected to have single or separate colours for whole words, beyond a composite of the colours of the individual letters—indeed, one would reasonably expect that for such synaesthetes, multiple colours in the word based on a combination of letter colours would be the norm, rather than the exception.

Based on these reports, it is clearly the case that at least some synaesthetes experience multiple colours per word as a matter of course, and that the single whole-word colour does not accurately represent their synaesthetic experience. The next logical questions, such as how common it is for a given synaesthete to experience multiple colours for words, what proportion of synaesthetes have multi-coloured words, and how important these multiple colours are to a synaesthete's word processing, have never been systematically investigated. That is, researchers investigating the cognitive, perceptual or linguistic connections between synaesthetic colours and words have no evidence that it is safe to ignore the multiple colours that a synaesthete may experience for a single word. One could point out that we also do not have sufficient evidence that it is *not* safe to ignore them, but I am not proposing hypotheses, but rather advocating research design based on evidence rather than assumptions. Only two studies to date have explicitly obtained multiple colours per word in an experimental setting: our compound words study [12], and a case study by Blazej & Cohen-Goldberg [10]. In both these studies, the one- versus two-colour difference was of experimental interest, but to date, no study has investigated how many colours synaesthetes experience or report per word as a baseline or as a function of length, morphological structure, frequency, etc. This is a fundamental gap in the scientific portrayal of grapheme-colour synaesthesia, but also represents a golden opportunity both to expand our understanding of synaesthesia itself and gain a clearer reference point from which to explore exactly how synaesthesia maps onto language, as I will elaborate on further below.

4. Sources of whole-word colours

The second, related, assumption is that the colour of a word derives straightforwardly from the colour of a particular grapheme—that is, if *R* is purple, so will be *rain* and *rhythm* and any other word beginning with *R*. Experimental reports often state that synaesthetes report word colours that correspond systematically to the colour of the first letter [9,18,21,22] or the first vowel [14]. However, these studies are limited in that they, again, typically only obtain a single colour per word. It could then be the case that when asked for the colour of a whole word, synaesthetes report the most salient letter-colour, starting from the beginning of the word. This is unsurprising, as the importance of the initial letter is well-established in the visual word recognition literature (e.g. [24–27]). However, because synaesthetes only have the option to provide one colour, only the first-letter colour is recorded,¹ and becomes conflated with the whole-word colour, rather than the first colour of potentially several within the word.

Additionally, studies linking whole-word colour to a particular letter were often investigated with colour-term descriptions (e.g. synaesthetes asked to describe in words the colour they experienced for *R* and *rain*), where using the same colour category when describing the colour of the word and the colour of its initial letter counted as a match (e.g. [9,21]). Even when more precise colours were collected using colour palettes, they were often interpreted in terms of colour categories (e.g. [13]). The use of colour categories rather than particular colours selected by synaesthetes may obscure subtle but important distinctions within a colour category, which may map meaningfully onto correspondingly subtle influences of spelling, meaning and canonical colour.

However, simply obtaining more nuanced colours is not enough, when synaesthetes also report that what they are asked by researchers to provide does not adequately reflect their experiences. Again, from the same study described above [23], one synaesthete characterized their colours as 'metallic' in nature sometimes or like the northern lights... my colours have grades and variants and shades. A photoshop like pallet [*sic*] is sometimes hard to choose a colour of a word with because my colours can be detailed and varying within a space.' Tellingly, one synaesthete described experiencing colours in a way that cannot be captured by a colour palette, and mentioned doing their best to make a decision based on other factors: '[S]ome of the words I saw with two colours in a stripe effect based on dominant letter and meaning. I did not know how to indicate this and so chose the strongest colour if it was dominant.' These reports not only provide further evidence that the simple single-colour test may not be adequate, but also introduce the idea of letter dominance, an as-yet-unstudied quality of the synaesthetic experience.

5. Letter dominance

The most striking recurrent theme across the responses described above [23] was the repeated reference to 'strong' or 'dominant' letters, which have a greater influence on the colour of the word than 'weaker' letters. This is sometimes the first letter, but not always. One synaesthete explains:

Some letters are stronger than others, and so in those cases when they are the first letter of a word, they color the entire word. Other letters are less strong, and so those tend to blend together to make a new color. If it's a blend of strong letters, then it's like a

mural of very distinct, individual colors. And then every once in a while there are letters that have a pattern with multiple colors.

This characteristic of letter dominance has, to date, never been experimentally investigated by synaesthesia researchers,² but in this sample, over half ($N = 13$, 65%) of the synaesthetes independently and spontaneously mentioned letter dominance as a primary influence in determining word colour.³ Many of them do note that the first letter of the word often dominates the whole-word colour, but explicitly point out that other dominant letters further on in the word have a disproportionate influence on word colour, for example: 'For words, it can sometimes be the colour of the first letter, or a mixture of several dominant colours in the word.' Another participant explained that letter dominance can be the primary source of whole-word colour: 'The colour of the word is formed from the dominant letters. For example 'f' is a dominant letter and will give the word a green tinge. 'a' is dominant letter and will give the word a yellow colour [*sic*].' It is particularly interesting to note that only three of the synaesthetes mentioned a strong or dominant colour; this dominance was usually described as a property of the letter itself, not the colour *per se*—although the colour of the letter is strengthened within the word because of its association with the dominant letter. Therefore, it seems that dominant letters may have a disproportionate influence on the synaesthetic colours of whole words, regardless of the nature of the colour associated with that letter and possibly independent of, or despite, position in the word. If letter dominance is indeed a quality of synaesthesia that is common among grapheme-colour synaesthetes, this would imply that attempting to capture letter-to-word colouring patterns without taking letter dominance into account presents an incomplete picture of the true nature of whole-word colouring.

As these comments have shown, the quality of letter dominance may be a widespread and influential characteristic of synaesthetic colour associations. Subsequent studies may investigate whether this dominant quality is a type of personality associated with the letters, as in ordinal linguistic personification (i.e. a different subtype of synaesthesia which associates personality traits with graphemes, e.g. C is a timid male or 8 an perfectionistic female; [33,34]). Alternatively, letter dominance may not be a full-fledged personality trait, but rather a conceptual property of each letter, or further related to properties of the associated colours (e.g. lightness and saturation), particular hues, letter frequency, position in the alphabet sequence, shape and so on. It may be that there are trends in the dominance of certain letters as there are trends in their associated colours (e.g. [9,16,35–37]), which could be connected to the processes involved in recognizing graphemes and words and therefore uncover a new tool for studying these psycholinguistics mechanisms. On the other hand, a letter's dominance may also be linked to its perceptual form, such that more visually complex, repeated, bold-faced, or capital letters may be more dominant.⁴ Naturally, all of the above is speculation at this point; aside from the reports of synaesthetes, as of yet I have no experimental results to support any claims about the characteristics of dominant letters, nor can I produce evidence that such a quality is widespread among grapheme-colour synaesthetes aside from that described above. The lack of existing evidence, however, is by no means proof that letter dominance is *not* an interesting or influential synaesthetic quality, when these qualitative reports by synaesthetes have not yet been systematically investigated.

The essential point is that the importance, impact and influence of a variety of qualities on synaesthetic associations—of which letter dominance is one example—is not only under-evidenced but often completely unknown. It is precisely this lack of research, combined with qualitative reports from synaesthetes consistently identifying these qualities as important, that should motivate and inspire researchers to expand the documented, scientific understanding of synaesthesia to more fully capture synaesthetes' experiences.

6. Implications for the psycholinguistics of synaesthesia

In presenting these arguments, I have no desire to dispute the existence of trends or 'rules' providing the underlying structure that scaffold synaesthetic experience (cf. [38]). Indeed, the entire logic of studying language using synaesthetic associations depends on the fundamental assumption that synaesthetic concurrents for linguistic inducers are in some way systematically related to the features of those inducers [39,40]. However, the limits imposed by the current methods of obtaining the colour associations for those linguistic items, and the assumptions that guide the design of those methods, similarly restrict the generalizability of those studies' conclusions.

As an example of this problem, Mankin *et al.* [12] reported that synaesthetes were more likely to report two colours for a compound word, rather than one, when the compound was lower in frequency. The authors pointed to this as evidence of dual- or multiple-route models of compound processing, on the assumption that a single colour for a compound was an index of its lexicalization as a single word. However, no study thus far has established, using a large bank of randomly selected high- and low-frequency words with a variety of morphological structures, how many colours synaesthetes experience for any given word. Without that context, it is impossible to say whether Mankin *et al.*'s [12] results were specific to compound words, or if any sort of multi-syllabic word would show the same frequency-based pattern regardless of morphological structure. Additionally, no study thus far has established whether this correspondence between the number of colours and the frequency of a word is, in fact, a pattern unique to synaesthetes. It could be the case that non-synaesthetes, if asked to assign any number of colours to a list of compound words, would also be more likely to select two colours for lower-frequency compounds. If this were the case, it would not only give us some very interesting insight into the mapping between colours and words as a more general cognitive correspondence, but could save researchers having to recruit synaesthetes specifically for studies of this type.

As mentioned above, interest in studying psycholinguistic questions with synaesthesia is a subfield generating increasing interest. For researchers interested in fine-grained linguistic phenomena, using synaesthetic experiences requires an equally precise and sensitive measure of colour concurrents that adequately captures the salient elements of the synaesthetic experience; the critical question is what are, and are not, salient elements. For researchers interested in studying, as another example, reading and visual word recognition, a single colour per word is simply not sufficient. In order to contribute meaningfully to the long-standing debate over how different letters inform the recognition of a word [41], researchers must be able to model how each letter contributes to the overall

synaesthetic colour experience of that word, and how those contributions may change in different contexts of interest (e.g. with different syllable structures, with letters in different positions within the word, etc.). If indeed qualities such as letter dominance are influential in determining word colouring, such models must account for them as well. For instance, when trying to understand the representation of the word *rain* based on its letters, a researcher might expect that the word *rain* is purple based on a purple *R*. However, if the yellow letter *N* is particularly dominant, the colour of *rain* may be yellow rather than purple. This would be inexplicable without including all of the letters in the representation of the word, and would appear to the researcher to be a word that did not match its expected letter colour, for an unknown reason. Such a model including both the colour of each letter and its strength might allow researchers to actually visualize the influence of different letters on word recognition using those colour patterns. However, this research must be founded on a much more substantial groundwork that puts specific contrasts in the context of the full spectrum of synaesthetic experiences.

7. Recommendations for future work

The preceding sections have argued that current research work on grapheme-colour synaesthesia has in large part potentially neglected to measure or account for critical aspects of the nature of synaesthesia because of *a priori* assumptions about what is important to measure and what is not. This is by no means to say that the preceding decades of research work are useless, or even wrong; scientific progression is by necessity a slow, careful, and incremental process building on existing ideas and filling in outstanding gaps. What I propose here is to expand our current understanding by embracing the richness and complexity of synaesthesia as the individuals who experience it report it to be. Of course, researchers studying synaesthesia cannot account for all the individual nuances in synaesthetic experiences if they want to be able to identify underlying patterns and systematic relationships. At some point, the map becomes the territory, and the idiosyncratic quirks of letter and colour that, to the synaesthete, are a valued and cherished part of their perception of the world become indecipherably complex to the scientist attempting to understand them. However, imposing limitations on the complexity of the experience in order to obtain data that researchers can meaningfully analyse assumes from the start that finding a simple, parsimonious explanation is the ideal result, instead of understanding the experience itself. Here I advocate taking a step back to understand and describe what synaesthetes do in fact typically experience first, and to attempt to identify patterns and causal relationships only when the overall picture is more fully developed.

8. Call for exploratory studies

At this moment there is so much still unknown about synaesthesia that it is difficult to know where to begin. An essential first step must be to establish the baseline range of experiences of a large number of synaesthetes, and only then to explore the full variety of those experiences for potential underlying trends or correspondences. Several well-known papers have done this for specific aspects of synaesthetic experience—for instance, Rich *et al.* [9] and Simner *et al.* [37] on the

population-level trends in letter-colour pairings. By sampling from a large number of synaesthetes and non-synaesthetes, these studies established that particular letters tend to have a particular colour concurrent—for example, *A* is famously red (at least in English; [16,35]), *Q* purple, *D* brown, etc. That is, these studies set the groundwork for future research investigating *why* these trends exist by first acquiring colours and letters from a large sample. Once that broader picture was established, the trends within could be identified.

A similar process should be undertaken to identify what, if any, aspects of letters and words besides hue may be consistent or salient for synaesthetes. The first critical consideration is to embrace the inherent heterogeneity of synaesthetic experience and compile a picture of synaesthesia in all of its manifestations. This requires large-scale surveys or tests of grapheme-colour synaesthetes, allowing them to elaborate on the qualities their synaesthetic associations may have, without disregarding anything on an *a priori* basis. Once a large and thorough inventory of the full spectrum of synaesthetic experiences is available, this database may then provide a wealth of information on potential underlying systematicity between graphemes and colours. The value of such exploratory and data-gathering work is clearly evident in large database studies such as Witthoft *et al.*'s [42] study on learned patterns in grapheme-colour associations, and Novich *et al.*'s [4] work on subgroups and comorbidity of synaesthesia subtypes. These large databases are invaluable, and I argue here that they need expansion, not replacement. At the moment, the Synesthesia Database, compiled primarily via the Synesthesia Battery (www.synesthete.org), does not provide synaesthetes the option to give this sort of data, so new tests and measures must be developed that capture the full range of synaesthetic experiences for letters and words. Such research could establish whether some letters are more likely than others to be, for example, sparkly, see-through, reflective or shiny, or which letters may be more or less likely to have a strong colour impression, or to have more than one colour. Ideally, these new tests and studies would first be piloted with small groups of synaesthetes to identify potentially informative qualities of letter-colours that may be as yet unknown to researchers. It may, of course, be the case that few or none of these qualities—which synaesthetes are frequently at pains to describe to researchers, when given the chance—turn out to be consistently or systematically mapped onto particular graphemes. They may well not be scientifically interesting or informative at all, except as curiosities or embellishments. But until research of sufficient rigour and scope has reliably established this, there is no reason to dismiss them as such—and good reason not to, should these understudied qualities turn out to be scientifically relevant.

At the word level, this groundwork is even more critically lacking. Researchers interested in the synaesthetic colours of words are in sore need of norming studies to establish baselines, composed of associations from synaesthetes on large sets of words with a variety of properties. These norming databases would not be primarily concerned with the specific colours of particular words, but rather the systematic mapping between linguistic features of words and corresponding features of the synaesthetic associations for those words. Without a substantive database of words with a variety of synaesthetic associations available to compare to, it is extremely difficult to tell what is actually of psycholinguistic (or psychological, or cognitive, or perceptual) interest and what

may simply be the prevailing way that synaesthetes experience colours for words of a particular length, or syllable structure, or prosodic pattern, or morphological construction, or shape, etc. For researchers interested in grapheme-colour synaesthesia as a tool for investigating psycholinguistic phenomena, I believe that large-scale normed databases of words and synaesthetic associations are crucial not only for identifying population-level trends worth following up with carefully controlled tests, but also for providing essential context for any work on the psycholinguistic features of synaesthetic associations. Beyond grapheme-colour synaesthesia, researchers interested in personalities for letters, or colours for music, or any other synaesthetic pairing, could make use of similar norming databases of a variety of inducers to provide context for their tests.

9. Sufficient sample size

A critical feature of norming studies like the ones proposed above will be input from a large number of synaesthetes with a variety of experiences. In the current research climate emphasizing replicability and sufficient statistical power, synaesthesia researchers cannot continue depending on case studies and sample sizes that barely scrape into double digits to make sweeping statements about the nature of synaesthesia. As previously mentioned, many of the highly cited influential early papers on synaesthesia—which contributed significantly to shaping the assumptions that this paper has sought to challenge—are indeed case studies, such as EP [18], MLS [22], JW [14] and JIW [43]. Naturally, before synaesthesia was well known or even recognized as a legitimate psychological phenomenon, and before a quick standardized test for verifying synaesthesia had been developed, case studies played a critical role in exploring the characteristics of synaesthetic experiences. Now that there are online communities of synaesthetes thousands strong and easily shareable online tests, statistically unreliable sample sizes should similarly be a thing of the past when researchers wish to generalize to the synaesthete experience, or more broadly, to the underlying processes of language in everyone. Otherwise, researchers risk establishing research precedents that are meant to reflect the experiences of millions of people based on the particular characteristics of a few dozen.

10. Collaboration

The two recommendations above—establish a baseline, and test large samples—will be difficult for any single researcher, or even a well-funded laboratory group, to achieve on their own. Luckily, working alone is not necessary. Recent papers combining data from different laboratories in different languages have showcased the value of collaboration across institutions and nations [35,36], providing insight across populations and languages that would have been prohibitively expensive or impractical in isolation. Establishing a baseline successfully, with participation from synaesthetes in sufficient numbers to develop a clearer picture of synaesthesia as a phenomenon, should be a priority of the synaesthesia research community as a whole, in the interest of creating resources and challenging assumptions that enhance the entire community's ability to ask and answer good scientific questions.

The other side to collaboration is also critical to understanding synaesthesia better: namely, listening to synaesthetes themselves. Synaesthetic experiences can and have been objectively verified to be real (e.g. [17,44,45], among many others), but the nature of those experiences—that is, what a given synaesthete actually sees, or hears, or feels, when presented with an inducing stimulus—cannot be objectively ascertained by technological or behavioural methods. In short, researchers desiring to learn what the experience of synaesthesia is actually like must *ask synaesthetes*; otherwise, they will be investigating synaesthesia as they assume it must be, rather than as it actually is experienced. Indeed, Smilek & Dixon [46] similarly advocated for first-person reports from synaesthetes to be integrated into the research design and interpretation process, citing many early successes in small case studies that synergistically combined experimental and subjective data.

Researchers desiring to do this would be well advised to invite the input of synaesthetes themselves throughout the process of developing and implementing new tests and studies. As an example, I am currently undertaking the development of a new multi-colour test for words in order to address the concerns I have articulated here. Once a beta version of the test has been finalized, I will make it available to online synaesthete communities for play testing and invite feedback and suggestions. By iteratively incorporating synaesthetes' and researchers' feedback, I hope to develop a mutually satisfactory and beneficial research tool that will eventually be available to the synaesthesia community as a whole, both scientists and synaesthetes. Simply put, collaborations like this between researchers and synaesthetes will be essential to fully understanding synaesthesia.

11. Recommendations for individual researchers

For an individual researcher designing a new study involving synaesthetes, I believe the best practice will be to make choices, not assumptions. That is to say, it is not the case that obtaining a single colour for each word, or any of the other assumptions I have addressed above, is always inappropriate. Rather, each researcher should consider their options carefully and choose how many colours to obtain, or what other qualities or ratings to ask for from synaesthetes, without dismissing any of these options out of hand. On the other side of this coin, it can be very difficult to know what data may, in fact, be useful to collect, and what may be superfluous, since not enough is currently known about the overall nature of synaesthetic experiences to make this determination from the outset based on the literature. Until the exploratory studies described above are able to shed light on this process, I suggest that the individual researcher consider the following when designing their study.

1: Describe rather than compare. Although there are many external pressures on researchers to test hypotheses and report significant results, outside of pure scientific curiosity (e.g. [47]), research on synaesthesia is in need of descriptive and exploratory studies to lay the groundwork for subsequent hypothesis-testing studies. As described above, consider developing a study to better understand and capture synaesthetic experiences, rather than making a specific *a priori* hypothesis. Only proceed with a specific hypothesis if there is enough extant contextualizing research to be confident that the manipulation you are intending to use, and the way you will measure it, is truly reflective of synaesthetes' experiences.

2: Build on synaesthetes' experiences. Whether designing a study from scratch or building on or replicating an existing one, incorporate the ideas and experiences of synaesthetes into the design process wherever possible. This can be achieved by posing the research question informally to synaesthetes as part of piloting the study. Communities such as the Synaesthesia Research Facebook group, the UK and American Synaesthesia Associations (uksynaesthesia.com and synesthesia.info), the International Association of Synaesthetes, Artists and Scientists (IASAS; theiasas.com), and the subreddit r/Synesthesia are examples of large associations of synaesthetes who frequently take an active interest in researchers' work. Inviting ideas from the synaesthete community may not only elicit new ideas or hypotheses but may also generate interest in the study for later participation.

3: Make choices, not assumptions. Approach your study without deciding *a priori* which elements of the synaesthetic experience may be informative and which can be disregarded. Use the recurring ideas from synaesthetes via (2), above, to expand your ideas about which variables may be relevant for a study about, for instance, the colours of words. Once you have a list of qualities that may be of interest—for example, the colours of individual letters within the word, each letter's position in the word, the dominance of those letters, how many colours are associated with each word, and the morphological structure and number of letters in the words of interest—you can then review this list in practical terms. As a first step, you may find that some qualities are simple to measure, even if a formal test does not yet exist: for instance, measuring letter dominance could be achieved by simply listing all the letters of the alphabet and asking synaesthetes to rate the dominance of each letter on a Likert scale. Such a test is both simple to create and to complete, so it would be worth including this synaesthetic quality in your research design. The remaining qualities, such as measuring multiple colours per word, may be more challenging, and you may not have the time or resources to create an appropriate apparatus. In this case, if you do have to use a suboptimal test (e.g. one that only collects one colour per word, or collects multiple colours but without an explicit mapping between the colours of the letters and the colour of the whole word), you can decide to do so with a clear understanding of what the potential confounds and limitations of your study may be. Finally, if you are intending to perform a confirmatory statistical test (rather than an exploratory analysis), chose to include variables that have a clear purpose and expected effect, either known or hypothesized. Qualities that are both difficult to measure and unclear in their expected influence—such as, at this point, the shininess of a particular letter—may necessarily be shelved until a better test apparatus is available, or exploratory studies show such a quality to be influential in word colouring.

4: Design a test. As the synaesthesia research community grows, more synaesthesia tests, such as the Synesthesia Battery (synesthete.org) and the Synaesthesia Toolkit (syn-toolkit.org) will be available to researchers interested in synaesthesia. However, avoid operationalizing your research question in a particular way simply because the test already exists, e.g. collecting only one colour per word because the extant tests only offer that option. Let reports from synaesthetes and your research question drive test design, rather than the reverse, whenever possible.

5: Gather feedback from synaesthete participants. As standard, include the opportunity in your test or questionnaire for synaesthetes to give feedback about their experience taking your test and any comments or ideas about the synaesthetic experience that you have elicited. These reflections can provide context for puzzling results, or may suggest an explanation you had not previously considered.

12. Summary

As a matter of urgency, all future research into synaesthesia, and especially research interested in complex and nuanced influences of language and meaning, must consider carefully the assumptions that underlie current research design. It is past time to discard research practices that are retained not because they accurately capture synaesthetic experiences, but because they simplify the collection and analysis of data for researchers. This paper has already identified some specific ways that current research could be improved or expanded. For instance, I have shown in-depth that the current assumptions that influence research into synaesthetes' colour experiences may consistently disregard crucial details about synaesthetic colours for words, which in turn directly impacts the applicability of that research to the larger questions about natural language processing that researchers are attempting to address.

This paper has argued that synaesthesia research on the whole, and particularly the emergent and growing field of research on the psycholinguistics of synaesthesia, must take careful stock of the methods and assumptions currently employed. In particular, I have argued that the currently established methods of obtaining and analysing colours for words are not sufficient to reflect the level of variation and detail that synaesthetes experience as a matter of course—details which may be fundamental to understanding and using synaesthesia as a tool for investigating language. The as-yet-unstudied quality of letter dominance is among many other basic elements of the synaesthetic experience for words that lacks a coherent baseline picture. Therefore, I have also called for large-scale exploratory studies examining the way synaesthetes subjectively understand and interpret their experience of colours for words, and further research establishing how various aspects of synaesthetic colours map systematically onto linguistic features for words. I believe it is crucial to examine carefully both the assumptions researchers make and the conclusions we can reasonably draw in order to contribute meaningful insight into the psycholinguistic investigation of synaesthesia.

Data accessibility. This article has no additional data.

Competing interests. We declare we have no competing interests.

Funding. We received no funding for this study.

Endnotes

¹A minority of synaesthetes report word colours based on vowels rather than first letters/consonants [21,28]. It is not yet clear why this distinction exists, although recent evidence has suggested that vowels and consonants contribute differently to word recognition [29–31]. Further discussion is forthcoming in Mankin and Simner [32].

²Researchers have used the term 'dominant' to describe the strongest colour in a compound word [12] and the most influential letter in determining a whole word's colour ([9,10,14,21]; etc.). However, the quality of 'letter dominance' described at present is a different characteristic, apparently referring to an intrinsic quality of the letter itself; i.e. colours are dominant in the word because they derive from letters

with this quality. Despite the potential for confusion, I have again used the term ‘dominant letter’ here because it is the term most frequently used by synaesthetes themselves to describe this quality.

³I emphasize here that this study was conducted online, without any contact between synaesthetes. However, as they were a part of a

database specifically for synaesthesia research, some of them may have participated in studies before, and possibly had contact with each other. Nevertheless, it is unlikely that these participants obtained the idea from other sources or colluded to produce these descriptions.

⁴My thanks to an anonymous reviewer for this suggestion.

References

1. Simner J. 2012 Defining synaesthesia: a response to two excellent commentaries. *Br. J. Psychol.* **103**, 24–27. (doi:10.1111/j.2044-8295.2011.02059.x)
2. Simner J. 2012 Defining synaesthesia. *Br. J. Psychol.* **103**, 1–15. (doi:10.1348/000712610X528305)
3. Carmichael DA, Down MP, Shillcock RC, Eagleman DM, Simner J. 2015 Validating a standardised test battery for synesthesia: does the synesthesia battery reliably detect synesthesia? *Conscious Cogn.* **33**, 375–385. (doi:10.1016/j.concog.2015.02.001)
4. Novich S, Cheng S, Eagleman DM. 2011 Is synaesthesia one condition or many? A large-scale analysis reveals subgroups. *J. Neuropsychol.* **5**, 353–371. (doi:10.1111/j.1748-6653.2011.02015.x)
5. Simner J, Mulvenna C, Sagiv N, Tsakanikos E, Witherby SA, Fraser C, Scott K, Ward J. 2006 Synaesthesia: the prevalence of atypical cross-modal experiences. *Perception* **35**, 1024–1033. (doi:10.1068/p5469)
6. Simner J, Carmichael DA. 2015 Is synaesthesia a dominantly female trait? *Cogn. Neurosci.* **6**, 68–76. (doi:10.1080/17588928.2015.1019441)
7. Baron-Cohen S, Wyke MA, Binnie C. 1987 Hearing words and seeing colours: an experimental investigation of a case of synaesthesia. *Perception* **16**, 761–767. (doi:10.1068/p160761)
8. Day S. 2005 Some demographic and socio-cultural aspects of synesthesia. In *Synesthesia: perspectives from cognitive neuroscience* (eds LC Robertson, N Sagiv), pp. 11–33, New York, NY: Oxford University Press.
9. Rich AN, Bradshaw JL, Mattingley JB. 2005 A systematic, large-scale study of synaesthesia: implications for the role of early experience in lexical-colour associations. *Cognition* **98**, 53–84. (doi:10.1016/j.cognition.2004.11.003)
10. Blazej LJ, Cohen-Goldberg AM. 2015 Multicolored words: uncovering the relationship between reading mechanisms and synesthesia. *Cortex* **75**, 160–179. (doi:10.1016/j.cortex.2015.11.017)
11. Hung W-Y, Simner J, Shillcock R, Eagleman DM. 2014 Synaesthesia in Chinese characters: the role of radical function and position. *Conscious Cogn.* **24**, 38–48. (doi:10.1016/j.concog.2013.12.009)
12. Mankin JL, Thompson C, Branigan HP, Simner J. 2016 Processing compound words: evidence from synaesthesia. *Cognition* **150**, 1–9. (doi:10.1016/j.cognition.2016.01.007)
13. Asano M, Yokosawa K. 2012 Synesthetic colors for Japanese late acquired graphemes. *Conscious Cogn.* **21**, 983–993. (doi:10.1016/j.concog.2012.02.005)
14. Simner J, Glover L, Mowat A. 2006 Linguistic determinants of word colouring in grapheme-colour synaesthesia. *Cortex* **42**, 281–289. (doi:10.1016/S0010-9452(08)70353-8)
15. Asano M, Yokosawa K. 2013 Grapheme learning and grapheme-color synesthesia: toward a comprehensive model of grapheme-color association. *Front. Hum. Neurosci.* **7**, 757. (doi:10.3389/fnhum.2013.00757)
16. Mankin JL, Simner J. 2017 A is for apple: the role of letter-word associations in the development of grapheme-colour synaesthesia. *Multisensory Res.* **30**, 409–446. (doi:10.1163/22134808-00002554)
17. Eagleman DM, Kagan AD, Nelson SS, Sagaram D, Sarma AK. 2007 A standardized test battery for the study of synesthesia. *J. Neurosci. Methods* **159**, 139–145. (doi:10.1016/j.jneumeth.2006.07.012)
18. Baron-Cohen S, Harrison J, Goldstein LH, Wyke M. 1993 Coloured speech perception: is synaesthesia what happens when modularity breaks down? *Perception* **22**, 419–426. (doi:10.1068/p220419)
19. Ward J, Mattingley JB. 2006 Synaesthesia: an overview of contemporary findings and controversies. *Cortex* **42**, 129–136. (doi:10.1016/S0010-9452(08)70336-8)
20. Barnett KJ, Feeny J, Gormley M, Newell FN. 2009 An exploratory study of linguistic-colour associations across languages in multilingual synaesthetes. *Q. J. Exp. Psychol.* **62**, 1343–1354; quiz 1354–1355. (doi:10.1080/17470210802483461)
21. Ward J, Simner J, Auyeung V. 2005 A comparison of lexical-gustatory and grapheme-colour synaesthesia. *Cogn. Neuropsychol.* **22**, 28–41. (doi:10.1080/02643290442000022)
22. Mills CB, Viguers ML, Edelson SK, Thomas AT, Simon-Dack SL, Innis JA. 2002 The color of two alphabets for a multilingual synesthete. *Perception* **31**, 1371–1394. (doi:10.1068/p3429)
23. Mankin JL. 2017 The psycholinguistics of synaesthesia. PhD thesis, University of Sussex, UK. See <http://sro.sussex.ac.uk/id/eprint/76640/>
24. Aschenbrenner AJ, Balota DA, Weigand AJ, Scaltritti M, Besner D. 2017 The first letter position effect in visual word recognition: the role of spatial attention. *J. Exp. Psychol. Hum. Percept. Perform.* **43**, 700–718. (doi:10.1037/xhp0000342)
25. Johnson RL, Eisler ME. 2012 The importance of the first and last letter in words during sentence reading. *Acta Psychol.* **141**, 336–351. (doi:10.1016/j.actpsy.2012.09.013)
26. Schoonbaert S, Grainger J. 2004 Letter position coding in printed word perception: effects of repeated and transposed letters. *Lang. Cogn. Processes* **19**, 333–367. (doi:10.1080/01690960344000198)
27. White SJ, Johnson RL, Liversedge SP, Rayner K. 2008 Eye movements when reading transposed text: the importance of word-beginning letters. *J. Exp. Psychol. Hum. Percept. Perform.* **34**, 1261–1276. (doi:10.1037/0096-1523.34.5.1261)
28. Simner J et al. 2006 Synaesthesia: the prevalence of atypical cross-modal experiences. *Perception* **35**, 1024–1033. (doi:10.1068/p5469)
29. Carreiras M, Duñabeitia JA, Molinaro N. 2009 Consonants and vowels contribute differently to visual word recognition: ERPs of relative position priming. *Cereb. Cortex* **19**, 2659–2670. (doi:10.1093/cercor/bhp019)
30. Carreiras M, Gillon-Dowens M, Vergara M, Perea M. 2008 Are vowels and consonants processed differently? Event-related potential evidence with a delayed letter paradigm. *J. Cogn. Neurosci.* **21**, 275–288. (doi:10.1162/jocn.2008.21023)
31. Duñabeitia JA, Carreiras M. 2011 The relative position priming effect depends on whether letters are vowels or consonants. *J. Exp. Psychol. Learn. Memory Cogn.* **37**, 1143–1163. (doi:10.1037/a0023577)
32. Mankin JL, Simner J. In preparation. Synaesthetic colours for simple and complex words.
33. Simner J, Holenstein E. 2007 Ordinal linguistic personification as a variant of synesthesia. *J. Cogn. Neurosci.* **19**, 694–703. (doi:10.1162/jocn.2007.19.4.694)
34. Smilek D, Malcolmsom KA, Carriere JS, Eller M, Kwan D, Reynolds M. 2007 When ‘3’ is a jerk and ‘E’ is a king: personifying inanimate objects in synesthesia. *J. Cogn. Neurosci.* **19**, 981–992. (doi:10.1162/jocn.2007.19.6.981)
35. Root NB, Rouw R, Asano M, Kim CY, Melero H, Yokosawa K, Ramachandran VS. 2018 Why is the synesthete’s ‘A’ red? Using a five-language dataset to disentangle the effects of shape, sound, semantics, and ordinality on inducer-concurrent relationships in grapheme-color synesthesia. *Cortex* **99**, 375–389. (doi:10.1016/J.CORTEX.2017.12.003)
36. Rouw R, Case L, Gosavi R, Ramachandran VS. 2014 Color associations for days and letters across different languages. *Front. Psychol.* **5**, 369. (doi:10.3389/fpsyg.2014.00369)
37. Simner J, Ward J, Lanz M, Jansari A, Noonan K, Glover L, Oakley DA. 2005 Non-random associations of graphemes to colours in synaesthetic and non-synaesthetic populations. *Cogn. Neuropsychol.* **22**, 1069–1085. (doi:10.1080/02643290500200122)
38. Simner J. 2013 The rules of synaesthesia. In *The Oxford handbook of synesthesia* (eds J Simner, EM

- Hubbard), pp. 147–164. Oxford, UK: Oxford University Press.
39. Mankin JL. 2017 The psycholinguistics of synaesthesia: what can we learn from studying synaesthetes? *PsyPAG Q.* **103**, 6–9.
 40. Simner J. 2007 Beyond perception: synaesthesia as a psycholinguistic phenomenon. *Trends Cogn. Sci.* **11**, 23–29. (doi:10.1016/j.tics.2006.10.010)
 41. Taft M, Xu J, Li S. 2017 Letter coding in visual word recognition: the impact of embedded words. *J. Memory Lang.* **92**, 14–25. (doi:10.1016/j.jml.2016.05.002)
 42. Witthoft N, Winawer J, Eagleman DM. 2015 Prevalence of learned grapheme-color pairings in a large online sample of synesthetes. *PLoS ONE* **10**, e0118996. (doi:10.1371/journal.pone.0118996)
 43. Ward J, Simner J. 2003 Lexical-gustatory synaesthesia: linguistic and conceptual factors. *Cognition* **89**, 237–261. (doi:10.1016/S0010-0277(03)00122-7)
 44. Hubbard EM, Ramachandran VS. 2005 Neurocognitive mechanisms of synesthesia. *Neuron* **48**, 509–520. (doi:10.1016/j.neuron.2005.10.012)
 45. Nunn JA *et al.* 2002 Functional magnetic resonance imaging of synesthesia: activation of V4/V8 by spoken words. *Nat. Neurosci.* **5**, 371–375. (doi:10.1038/nn818)
 46. Smilek D, Dixon MJ. 2002 Towards a synergistic understanding of synaesthesia: combining current experimental findings with synaesthetes' subjective descriptions. *Psyche* **8**.
 47. Grimes DR, Bauch CT, Ioannidis JPA. 2018 Modelling science trustworthiness under publish or perish pressure. *R. Soc. open sci.* **5**, 171511. (doi:10.1098/rsos.171511)