A systematic investigation of conceptual colour associations.

Diana Su Yun Tham[1]

Paul T. Sowden[2,3]

Alexandra Grandison[3]

Anna Franklin[4]

Anna Kai Win Lee[3]

Michelle Ng[3]

Juhyun Park[1]

Weiguo Pang[5]

Jingwen Zhao[5]

[1]Lancaster University

[2]University of Winchester

[3]University of Surrey

[4]University of Sussex

[5]East China Normal University

Correspondence: Paul T. Sowden, Department of Psychology, University of Winchester, Winchester, SO22 4NR, UK; 01962 827406; paul.sowden@winchester.ac.uk

A systematic investigation of conceptual colour associations.

Associations with colours are a rich source of meaning and there has been considerable interest in understanding the capacity of colour to shape our functioning and behaviour as a result of colour associations. However, abstract conceptual colour associations have not been comprehensively investigated and many of the effects of colour on psychological functioning reported in the literature are therefore reliant on *ad hoc* rationalisations of conceptual associations with colour (e.g. blue – openness) to explain effects. In the present work we conduct a systematic, cross-cultural, mapping of conceptual colour associations using the full set of hues from the World Colour Survey (WCS). In Experiments 1a and 1b we explored the conceptual associations that English monolingual, Chinese bilingual and Chinese monolingual speaking adults have with each of the 11 Basic English Colour Terms (black, white, red, yellow, green, blue, brown, purple, pink, orange, grey). In Experiment 2 we determined which specific physical WCS colours are associated with which concepts in these three language groups. The findings reveal conceptual colour associations that appear to be 'universal' across all cultures (e.g. white – *purity;* blue – *water/sky related;* green – *health;* purple – *regal;* pink – *'female' traits*) as well as culture specific (e.g. red and orange – *enthusiastic* in Chinese; red – *attraction* in English). Importantly, the findings provide a crucial constraint on, and resource for, future work that seeks to understand the effect of colour on cognition and behaviour, enabling stronger *a priori* predictions about universal as well as culturally relative effects of conceptual colour associations on cognition and behaviour to be systematically tested.

Keywords: *colour perception; colour cognition; colour language; colour-concept association; concept-colour association; cross-cultural*

A systematic investigation of conceptual colour associations

For most people colour is an immediate and essential component of their perception of the world and understanding colour phenomena has been an area of highly productive research (e.g., Elliot, Fairchild, & Franklin, 2015; Hardin & Maffi, 1997; Hurlbert & Owen, 2015; Regier & Kay, 2009). The importance of colour extends beyond simple perceptual experience to encompass rich symbolic functions, as emphasised by the many expressions of language specific sentiments or concepts (Hutchings, 1997, 2004; Mohammad, 2011) such as, in the English language, *green with envy* (jealous), *out of the blue* (randomly), *red hot* (exciting), *white as a ghost* (fear) and in Chinese, *bright white* (to understand), *red fire* (flourish, good luck) or *grey heart* (discouraged). However, at present, we lack a systematic and comprehensive understanding of the extent to which these abstract *conceptual,* colour associations are shared by members of a culture, differ across culture, and are linked to specific points in colorimetric space.

Further, whilst abstract conceptual colour associations are important to understand in their own right, the importance of building a systematic understanding of conceptual colour associations is further emphasised when we consider their use in other areas of psychological research. One topic of enquiry that has attracted substantial interest in recent years has been the impact of colour on psychological functioning (Elliot, 2015; Elliot & Maier, 2014 for a review). Colour has been found to affect performance on IQ, anagram, and numeracy tests (e.g. Elliot, Maier, Moller, Friedman, & Meinhardt, 2007, Brooker & Franklin, 2016), creativity tasks (e.g. Mehta & Zhu, 2009), challenging cognitive tasks (e.g. Kumi, Conway, Limayem & Goyal, 2013; Yamazaki 2010), physical strength tasks (e.g. Elliot & Aarts, 2011), and in combat sports (e.g. Hill & Barton, 2005). Notably, explanations of many of these effects have been provided *ad hoc* through reference to conceptual colour associations,

which may arise ontogenetically through repeat pairings (Crozier, 1996; Mehta & Zhu, 2009).

Importantly, a number of the reported effects of colour on cognition and behaviour have been inconsistent across studies or have failed to replicate (e.g. Langguth *et al.*, 2009; Larsson & Von Stumm, 2015; Lynn, Giebelhausen, Garcia, Li & Patumanon, 2016; O'Connor, 2010; Peperkoorn, Roberts & Pollet, 2016). This variability in findings, coupled with the *ad hoc* reasoning that much previous research on the effects of colour on cognition and behaviour has had to rely upon, suggests that for further progress to be made in this field it is essential to build a more systematic understanding of conceptual colour associations.

Thus, the central motivations for the present work are (1) to address the lack of a comprehensive and systematic understanding of conceptual colour associations and (2) through identifying conceptual colour associations for English, monolingual and bilingual Chinese speaking participants, to enable future work on the effects of colour on cognition and behaviour to make systematic *a priori* predictions that are both universal as well as culturally relative (e.g., Davies & Corbett, 1997; Grandison, Davies, & Sowden, 2014; Taylor, Clifford, & Franklin, 2013).

Varieties of colour association

Colour associations may take different forms including those that are bound to specific objects as well as non-object related emotional and abstract conceptual colour associations. Usually, colour is viewed on an object within a particular psychological context and, further, the colour's meaning and valence can be influenced by the object on which the colour is viewed. Palmer and Schloss (2010) conducted an extensive investigation of the objects associated with a range of colours resulting in their proposal of an Ecological Valence

Theory (EVT) of colour preference (see also Strauss, Schloss, & Palmer, 2013; Taylor & Franklin, 2012). This argues that colour preferences result from people's combined liking/disliking (valences) of all associated coloured objects. For example, strongly disliked dark greenish yellow is associated with strongly disliked entities such as dirty water, bile, and rotten food, whereas strongly liked saturated blue is associated with strongly liked entities such as clear sky and clean water. The EVT suggests that the relevant factors determining one's preference for a particular colour include an object's importance in one's life, its perceptual salience, the frequency with which it is experienced, and the strength of the affective response that it causes.

Colour associations that are inspired either by objects or by physical aspects of the environment are likely to emerge cross culturally. For example, green is the colour of vegetation, and so it has been suggested may have generic connotations with growth, life and health (Crozier, 1996). Indeed, in English, idioms containing the word *green* represent a connectedness with nature, such as *having a green thumb* or *green living.* Further, the connotations of the colour green may in turn influence the naming of objects. For instance, an effective antiseptic solution "zelyonka", literally meaning "green thing", is widely used in Russia to treat and prevent infections, giving an example of a conceptual link between the colour green and health in Russian. Similarly, object based associations may be culturally specific and give rise to new conceptual associations. For example, in Chinese there is a saying: "wearing a green hat", which symbolises unfaithfulness. This saying dates back to the Yuan dynasty where the relatives of prostitutes were forced to wear green hats — a link that further reflects the grounding of some colour associations in physical objects (see Taylor and Franlkin, 2012).

In addition to work on colour-object associations, a further body of work has explored colour-emotion associations. Thorstenson, Elliot, Pazda, Perrett and Xiao (2019) investigated

colour-emotion associations in relation to facial stimuli. Participants were asked to manipulate the colour of neutral faces to best match the emotions of anger, disgust, fear, happiness, sadness and surprise. Increased redness was associated with anger, happiness and surprise, whereas decreased redness was associated with disgust, fear and sadness. Additionally, participants increased yellowness in relation to anger and happiness and decreased yellowness for fear and sadness. The authors highlight the link between increased redness and approach-oriented emotions and decreased redness and avoidance-oriented emotions. This pattern conveys colour-emotion associations that are consistent with physiological signals of emotion, which may be akin to the colour-object associations discussed previously. Further work has investigated colour-emotion associations in other contexts. In a study aiming to create a large colour-word association lexicon via crowdsourcing, Mohammad (2011) found consistent associations with a number of abstract concepts including emotions. For example, red was consistently linked with anger, black with disgust and white with honesty. Additionally, Hupka, Zaleski, Otto, Reidl and Tarabina (1997) employed a closed ended task with American, Russian, German, Mexican, and Polish subjects to derive a list of consistent colour-emotion associations that varied across cultures. Participants were required to indicate the extent to which they associated anger, envy, fear and jealousy with 12 focal colours. Americans associated envy with black, green and red, whilst for Russians it was black, purple, and yellow. Poles connected anger, envy, and jealousy with purple, whilst Germans associated envy and jealousy with yellow. This evidence that emotion-colour associations vary across culture has been further substantiated (Barchard, Grob & Roe, 2016) and this may be due to differences in the expression and perception of emotion across cultures. Thus, this body of research presents a complex pattern of findings with regard to colour-emotion associations.

In contrast to the well-established work on colour-object and additional work on colour-emotion associations, other types of abstract conceptual colour associations, such as the English association of pink with femininity, have been the focus of controversy and debate. For instance, corporations have capitalised on the pink-femininity association through targeted marketing, exploiting what is believed by some to be a harmful stereotype (Del Giudice, 2017). Work that systematically investigates abstract conceptual colour associations of this nature is limited and key questions remain unanswered.

One issue is the extent to which abstract conceptual colour associations vary across culture. For centuries, people from all over the world have used colour to symbolize particular beliefs or events and it is suggested that, depending on the culture or the part of the world that they are from, the symbolism of a colour varies according to a group's habits, traditions and myths (Hemming, 2012). For example, the colour red is used to signal loss/failure/bad/danger in the American culture (Moller, Elliot & Maier, 2009), whereas red is used to signal happiness or prosperity in the Chinese culture (Jacobs, Keown & Worthley, 1991). Further, in recent research by Zhang and Han (2014), they found that in the Chinese population the same colour could have opposite meanings based on professions. In the Chinese stock market, red refers to a rise in stock price and green indicates a decrease, which is the reverse of the typical use of red to denote negative meaning. The authors found that red undermined performance on an IQ test among college students from China whereas red enhanced performance on an IQ test among stockbrokers from the same culture. Thus, colour effects may manifest through culturally learnt associations that are specific to country, geographical region or even cultural group. Consequently, there is a clear need to systematically assess conceptual colour associations in multiple language groups using the same set of colour stimuli to determine the extent to which conceptual colour associations are universal or culturally relative.

To date, two types of task have been established in order to understand conceptual colour associations. First, closed-ended tasks that require participants to choose colours that are associated with a limited and predefined list of words. For example, in three closed-ended tasks using 9 – 16 concepts, stereotypical colour associations have been found in American (US), Yunnan Chinese, and Hong Kong (HK) Chinese participants (Bergum & Bergum, 1981; Courtney 1986; Chan & Courtney, 2001). For example, the association between *stop* and red was shown by 100% of US participants, 66.4% of HK participants and only 48.5% of Yunnan participants. Additionally, the association between *go* and green was shown by 99.2% of US participants, 62.6% of HK participants and 44.7% of Yunnan. Whereas *safe* and green were associated by 62.2% of Yunnan participants, 61.4% of US participants, and only 38.2% of HK participants. While closed-ended tasks have the advantage of obtaining elevated colour-concept consensus (almost all had associations with >50% consensus), they have the disadvantage of not eliciting knowledge of a considerable number of other concepts associated with a colour.

Second, open-ended tasks (e.g., Bonnardel, Beniwal, Mayukhini, & Davies, 2014; Prado-Leon, Avila-Chaurand, & Rosales-Cinco, 2006) require participants to list words that are associated with a limited number of colour stimuli (e.g., 8 to 16). For example, Prado-Leon and colleagues (2006) instructed Mexican students to choose concepts from a list of 60 words that were evoked by each of 16 colours and to list any additional associations that were not on the initial list. Although the open ended nature of the latter part of the procedure resulted in a larger number of associations (399 concepts), it had the disadvantage of obtaining fairly low consensus (criterion of 20% minimum). 'Well-determined' associations of more than 20% consensus were *sexuality* – red, *death* – black, *innocence* – white, *nature* – green, *dirtiness* – brown, whereas less-determined associations with consensus of 10% to

20% were *sadness* – grey, *femininity* – pink, *brightness* – yellow, *passivity* – blue, *activity* – orange.

Because of the higher consensus found in closed than open ended tasks, inconsistent apparent patterns of colour association may emerge across these different types of study. For this reason, unlike previous research, in the present work we combine these approaches and use them to make cross-cultural comparisons between two or more language groups. Further, it is not clear whether colour effects are mediated by categorical representations of colour (e.g., the category of red colours), by individual colour exemplars (e.g., a specific red) or how they are mediated by colorimetric dimensions such as lightness. For instance, whilst many English speakers might agree that red is associated with danger, is this association with the whole range of colours categorised as red or with more specific exemplars of red? No study has mapped associations onto colour space systematically to inform these questions.

To address the limitations of the sparse nature of previous work on abstract conceptual colour associations the present work seeks to comprehensively and systematically assess colour-concept-colour associations using a large number of (330) colour stimuli from the World Colour Survey (WCS stimulus grid; Kay, Berlin, Maffi, Merrifield & Cook, 2009; see Figure 1 for an example of the WCS stimulus grid). Further, this will allow us to assess whether conceptual associations link to whole colour categories, to particular exemplars, or to particular colorimetric dimensions.

Figure 1. The stimulus grid taken from the World Colour Survey (Kay, Berlin, Maffi, Merrifield, Cook, 2009), which samples stimuli from the Munsell colour system at regular hue intervals, at various lightness levels and at high saturation for a given hue.
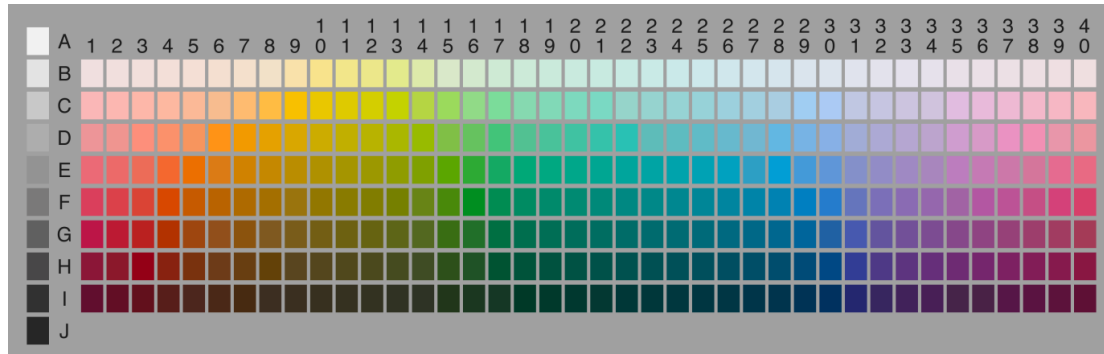
The effect of context

A further factor that may influence conceptual colour associations is the context in which they are invoked. As noted earlier, the pattern of findings about the effect of colour on psychological functioning is not consistent across the field (e.g., Langguth et al., 2009; O'Connor, 2010). Different colours seem to impede or facilitate different types of tasks and the same colour can have different effects on different tasks and in different situations. For instance, the red colour of a traffic light will be rapidly understood by drivers as communicating the possibility of danger and the need to stop, whereas it is argued that in the context of a romantic dinner the same colour could signal love or attraction. Elliot and Maier's (2012) 'colour-in-context' theory suggests that colours trigger different types of motivation, which can inhibit or enhance performance, depending on the context. For example, whilst in achievement contexts red seems to promote avoidance motivation, which impedes cognitive performance. Conversely, in mating contexts red may trigger an approach motivation.

For the present research, an implication of these findings is that colour associations may similarly be context sensitive. Here we focus on establishing generic conceptual colour

associations outside of any specific context and on developing the methods to systematically and comprehensively map concepts to particular points in colorimetric space. We hope that this will pave the way for future work to explore, less dominant, context specific colour associations by conducting similar mapping exercises to that used in the present research on a context by context basis. This should facilitate the formulation of robust, context specific, *a priori* predictions in future work in addition to the more dominant, non-context specific, effects facilitated by the present work.

The present study

Thus, our main objective is to explore and compare the range of conceptual colour associations that English (monolingual) language speakers from the UK, and 2 groups of Chinese (English/Chinese bilingual and Chinese monolingual) language speakers from China have with the 11 'basic colour terms' (BCT's) of the English language. A considerable amount of research has explored the way that colour terms are represented in language. In 1969, Berlin and Kay suggested that the basic colour terms in any language, including English, are those that are "psychologically salient". By this, they meant that the BCT's are used with high consensus and consistency by all competent speakers of the language. Berlin and Kay (1969, 1991) used a set of stimulus materials developed by Lenneberg and Roberts (1956) to assess the basic colour terms of 98 languages. They concluded that there are universals in the semantics of colour in (probably) all languages, and a hierarchy of development of the BCT's:

[BLACK – WHITE] → RED → [YELLOW – GREEN] → BLUE → BROWN →
[PURPLE – PINK – ORANGE – GREY]

There are six positions in the hierarchy shared by eleven colour categories such that colour categories grouped in brackets share a common position in the series. Although there are hundreds of individual words for colours (Boynton & Olsen, 1990; Lindsey & Brown, 2014; Sturges & Whitefield, 1995), we are choosing to focus on basic colour terms because these are the most salient. The English language has all 11 of Berlin and Kay's BCTs whereas Mandarin/Cantonese language has 9 basic colour terms (excluding brown and orange; Gao & Sutrop, 2014). Gao and Sutrop predict that the full set of 11 basic colour terms will appear in Mandarin suggesting that orange is the next candidate for basic status in Mandarin and brown will also ultimately become basic.

The descriptions of colour stimuli and comparisons across studies in other areas of colour research have been aided by the development of colour order systems that relate the colour stimulus to the dimensions of colour such as hue (e.g. red, green, blue), chroma/saturation (the intensity or amount of pigment in a colour; Fairchild, 2005), and lightness. However, a key limitation of much previous work on colour and psychological functioning is inadequate stimulus control for saturation and lightness, and often a predominant focus on just *hue* (e.g., see Elliot & Maier, 2012, for a review). As a result of this limitation, two clear problems are highlighted. First, it makes replication impossible – replication requires the availability of precise and accurate information about the properties of the colour stimuli being used, including hue, lightness, and saturation. Second, it is difficult to make predictions about effects of colour on cognition and behaviour because real-world colours typically vary on more than just the hue dimension. Further, much work on conceptual colour associations has not presented a large variety of physical colour stimuli to enable any accurate predictions on these relationships. Consequently, in the present work we make use of accurately specified stimuli described in the Munsell colour system[1]. Through using a systematic and standardised colour system, we will be able to identify if it is any red

hue, for example, (regardless of lightness and saturation) or a specific red (with a specific hue, lightness, and saturation) that evokes a particular association.

Therefore, in a series of studies, we firstly explore the associations that English and Chinese speakers have with the 11 basic colour terms, presented as physical colour arrays drawn from the WCS stimulus grid, through an open-ended task (Experiments 1a & 1b). We then investigate which specific physical colours are associated with these meanings through a closed-ended task using the full WCS stimulus grid (Experiment 2). Our findings will yield a comprehensive mapping of conceptual colour associations in two distinct languages and have important implications for predicting how particular colours might have reliable effects on human cognition and behaviour via their associations.

Experiment 1: identifying concept→colour associations

Experiment 1 was designed to explore the set of associations that participants have with each of the 11 basic colour categories. In particular, we sought to identify those associations that participants have that are conceptual rather than grounded in specific object-colour associations. Whilst it might be expected that participants have many different associations with colours, our objective was to identify those associations about which there is agreement across participants (i.e. consensus) as distinct from those that are idiosyncratic to particular individuals. In doing this, we aimed to reveal the *basic* conceptual colour associations of the English and Chinese languages.

To reveal these basic conceptual colour associations, we undertook three steps. First, in an open-ended task (Experiment 1a), we asked participants to list non-object words that they associated with each of the 11 basic colour categories (concept→colour association). Second, in Experiment 1b, we asked a further set of participants to categorise the words from

Experiment 1a into different types of association (object, emotions, quality of things, actions and symbolic). Having categorised the words into different types of association we eliminated any object associations. Finally, we then looked for words that were essentially synonymous within culture in order to produce a final list of concepts that could be evaluated for: (1) *consensus* (level of agreement about the concept→colour association across participants; and (2) the *specificity* of the association to a particular colour category. At the end of this process we were therefore able to identify those associations that had the greatest consensus and specificity to a colour category. These were the candidate concepts for the basic conceptual colour associations of English and/or Chinese languages.

Experiment 1a: Open-ended list task

Experiment 1a was an open-ended list task whereby participants were asked to list as many non-object words (concepts) as possible associated with arrays of colours representing the 11 basic colour categories.

Method

Participants. There were 80 participants comprised of 23 monolingual English speaking participants from the UK (15 females), 23 bilingual Chinese (and English) speaking participants from China residing in the UK (18 females), and 34 monolingual Chinese speaking participants from China residing in China (31 females). Their ages ranged between 18 and 25 years. Monolingual English speaking participants and bilingual Chinese speaking participants were undergraduate students at the University of Surrey in England whereas monolingual Chinese speaking participants were undergraduate students at East China

Normal University in Shanghai, China. Although our bilingual Chinese participants reported that they spoke Mandarin and/or Cantonese, we considered them as bilinguals due to the English language proficiency pre-requisite in order to come to study at the University of Surrey and undertake a university qualification, which is also taught in English. Our monolingual Chinese participants on the other hand reported that they spoke Mandarin (see Supplementary Information **Table 1** for more information on Chinese speaking participants' province of origin). Monolingual English speaking participants had a mean age of 21.7 years (SD = 1.77), bilingual Chinese speaking participants had a mean age of 21.5 years (SD = 1.34), and monolingual Chinese speaking participants had a mean age of 20.5 years (SD = 1.57). All participants were screened for colour vision deficiencies using the City Colour Vision Test (Fletcher, 1981) and Ishihara Test for Colour Blindness (Ishihara, 2009). Participants from the University of Surrey were paid £5 for their participation whereas participants from East China Normal University were paid ¥40 for their participation. The protocol was carried out in accordance with the ethical standards of the Research Ethics Committee of the University of Surrey.

Stimuli and Experimental Set-up. Munsell colour tiles (measuring 2cm x 1cm) from the Munsell Book of Colour[2] were used to reproduce the colour stimuli of the World Colour Survey (WCS stimulus grid; Cook, Kay & Regier, 2005). The colour category membership of the colours in the WCS stimulus grid were determined using Berlin and Kay's English colour naming data (1969; 1991). We chose to use one language's colour naming data so that we can make consistent comparison between the two cultures. These data show the colours in the WCS stimulus grid that fall into each of the 11 basic English colour categories (see **Figure 2**). The colour tiles that fell into each of the 11 basic colour categories were arranged on 11 separate A4-size boards, one colour category per board. The position of each colour tile was

as set out as in the WCS stimulus grid with 0.8 inch spacing in between tiles. All boards were painted with Munsell N5 grey to give a background matching the original WCS stimulus grid background. Each board was presented with a piece of clear acrylic sheet (1cm thickness) on top of the stimulus grid to protect the colour tiles.

CIE (1931) Yxy measurements made using a Minolta Colorimeter showed that with the clear acrylic sheet placed over the stimuli readings changed by reducing Y by .58 cd/m^2, whilst leaving x & y unchanged. A fluorescent light tube that simulated standard European natural daylight (D65 simulator) was used to illuminate the Munsell-WCS stimulus grid in an otherwise dark room.

Figure 2. Colour terms and mapping for (American) English. Only these tiles were used in Experiment 1a. White dots indicate focal colours. Reconstructed from *Basic Color Terms* (pg. 119), by B. Berlin & P. Kay, 1969, Berkeley: University of California Press. Copyright 1969 by CSLI Publications. Adapted with permission.
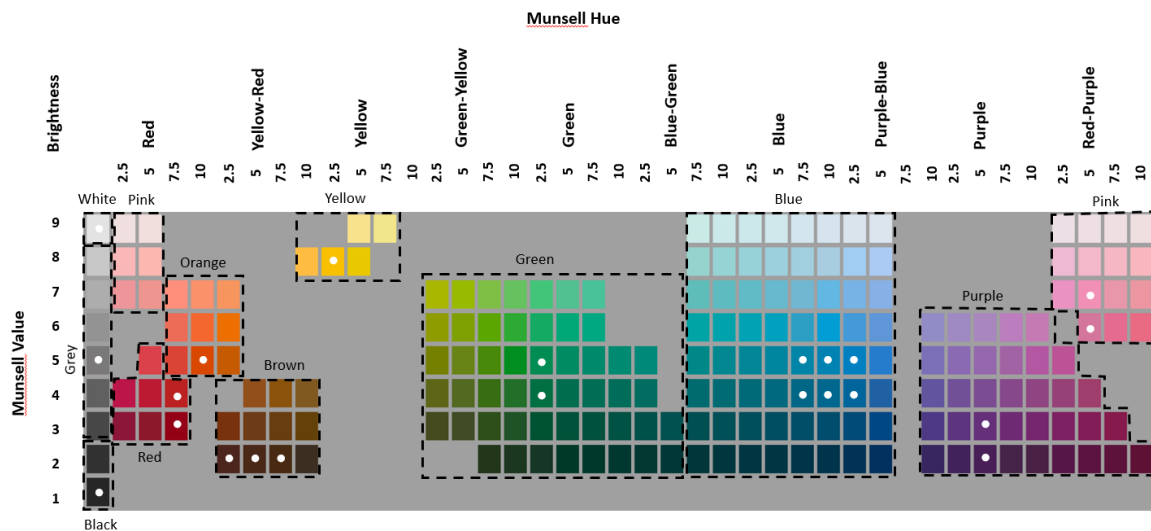
Procedure. The arrays of stimuli representing the 11 basic colour categories were presented one at a time in a random order. Depending on the native language spoken, participants were asked either in English, Cantonese, or Mandarin to list concepts (in English

or Chinese respectively) that they associate with each of the colour categories that are not specific objects. They were provided with the following instructions (either in English, Traditional Chinese for Cantonese speaking participants, or Simplified Chinese for Mandarin speaking participants):

*"Here is an array of colours. Tell me any associations that you have with any of the colours before you that **are not** associations with specific objects. For instance, certainly, many people might associate green with grass. However, we would like you to think of symbolic associations (e.g. silence), emotions (e.g. happy), qualities of things (e.g. slippery), or actions (e.g. rolling) that you associate with the colours. Take as long as you need to list all the associations you can think of."*

The instructions in Chinese were back translated to English by three multi-lingual native Chinese speakers and checked by three English-speaking experts to ensure validity. The instructions were left in front of participants for reference. All participants wrote down the concepts they associated with each colour category on separate 10cm x 15cm index cards. They were also allowed to return to previous colour arrays and cards to list further associations when they came to mind.

Results and Discussion

The list task was designed to reveal conceptual associations with 11 basic colour categories. Overall, there were a total of 1672 (English) words from English monolinguals, 2357 (Chinese) words from Chinese bilinguals, and 4220 (Chinese) words from Chinese monolinguals associated with the 11 categories, although some of these words were specified to more than one colour category. For example, '*depressing*' was listed in the colour

categories black and grey. After taking account of these duplications, there were 1076 unique English words, 1329 unique Chinese words (from bilinguals), and 2170 unique Chinese words (from monolinguals), which will be used in the word categorization task. The raw data with the full list of associations produced by each participant are available at the Open Science Framework (https://osf.io/).

Experiment 1b: Word categorisation task

Experiment 1b was designed to refine the list of unique English and Chinese associations by removing any object-based associations and by identifying concepts that were essentially synonymous and that could therefore be 'collapsed' into a single term. Further, we assessed the extent of agreement about each concept→colour association across participants (*consensus*) and the *specificity* of the association to a particular colour category. From this process we aimed to extract a final list of words that were clearly conceptually associated with a specific colour category and that had well-determined agreement across participants.

Method

Participants. There were 56 participants comprised of 11 monolingual English participants (10 females; mean age = 20.5 years, SD = 2.07), 11 bilingual Chinese participants (8 females; mean age = 20.4 years, SD = 1.4), and 34 monolingual Chinese participants (29 females; mean age = 20.9 years, SD = 2.54). None had participated in Experiment 1a but their demographic was the same as participants in Experiment 1a. Participants from the University of Surrey were paid £20 for their participation whereas participants from East China Normal University were paid ¥120 for their participation.

Materials and Procedure. The stimuli used in this task were the total list of English words ($n = 1076$), Chinese words from bilinguals ($n = 1329$), and Chinese words from monolinguals ($n = 2170$) produced in the list task after taking account of duplications where the same word was associated with more than one colour category.

Participants read through the randomised word list appropriate for their language and identified if each word belonged to one of five categories informed by previous research (Hemmings, 2012; Ludwig & Simner, 2013; Mohammad, 2011) and identified in a pilot test (object, quality of things, actions, emotions, symbolic word). Participants were encouraged to choose one category per word but they were also allowed to choose more than one category if necessary. They were also allowed to choose a sixth category, 'unsure', if they could not otherwise categorise the word. The following instructions were provided to the participants (either in English, Traditional Chinese for Cantonese speaking participants, or Simplified Chinese for Mandarin speaking participants):

"Here is a list of words that people associate with one or more colours. We would like you to categorise each word into one or more of the following categories by ticking the appropriate choice on your response sheet. 'Object' includes actual things such as a post-box or the sea. 'Quality of things' includes perceptual qualities of things that you would perceive using your senses such as shiny or slippery. 'Actions' includes physical actions such as bouncing or rolling. 'Emotions' includes things that you feel or experience such as angry or happy. 'Symbolic associations' includes conceptual associations such as spirituality or danger. If you feel that a word doesn't fit any of these categories you can select the unsure response. Please try and think carefully before you make your selection. Where possible you should choose the single category that you think is best for a particular word. However, sometimes

you may feel that a word falls into more than one of the categories, and in this case, you may choose more than one category if necessary. Once you have completed a page, you can have a 2 minutes break before moving on to the next page."

We included certain example words in the instructions based on a pilot test that showed 100% agreement that those particular words belonged in the respective individual category. As in Experiment 1a, the Chinese versions were checked via back translation. The instructions were left in front of participants for reference. Participants were allowed to take as long as they needed to choose the appropriate category for each word.

Results and Discussion

Word categorisation. Any words that were primarily categorised as object-based associations (defined as more than 50% agreement across participants) were removed from the list. While we instructed Experiment 1a participants not to list object words, it would seem as though the majority of our participants (if not all) found it challenging to completely avoid this, in line with the importance of colour-object associations found in previous research (Palmer & Schloss, 2010). The number of object words listed by our monolingual English participants ranged between 1 and 15 words ($M = 4.28$, $SD = 3.46$), bilingual Chinese participants ranged between 1 and 24 words ($M = 7.14$, $SD = 6.07$), and monolingual Chinese participants ranged between 1 and 24 words ($M = 5.19$, $SD = 6.35$). Further, any words for which there was not good agreement, defined as less than 50% agreement across participants, about the category of association were removed from the list. This categorisation process was undertaken to ensure that: (1) we identified those words that were clearly conceptual

associations; and (2) we retained the maximum number of words for further reduction (collapsing synonymous words).

Of the list of 1076 English words, 791 words were categorized as concepts (non-object) with more than 50% agreement. Three hundred concepts were categorised as colour associations to do with the quality of things, 190 as actions, 206 as emotions, and 95 as symbolic associations. Of the list of 1329 Chinese words from bilinguals, 803 were categorised as concepts. More than 50% agreement on a single category was achieved for 276 quality of things, 231 actions, 130 emotions, and 166 symbolic words. Finally, of the list of 2170 Chinese words from monolinguals, 1586 were categorised as concepts with more than 50% agreement: 701 quality of things, 488 actions, 153 emotions, and 244 symbolic words.

In the list of concepts, there were also occasions where participants chose slightly different concepts that meant essentially the same thing (e.g. *joy* and *joyful*). Two steps were taken to aggregate synonymous concepts. First, concepts were combined provided that they shared the same root word (e.g. *boredom* and *bored*), were produced for the same colour category and were produced by different people. This last stipulation was because we assumed that if the same person produced two related concepts in response to the same colour category (e.g. a participant listed anger and angry for the colour category red) he/she presumably meant somewhat different things by them. Second, concepts were collapsed only if they were categorised in the same conceptual category. For example, if *bored* and *boredom* were both categorised as 'emotions' then they were collapsed. However, if *bored* was categorised as 'emotions' whereas *boredom* as 'actions', then both concepts were retained separately.

Assessing consensus and specificity of concept→colour associations. Finally, in addition to filtering our list to remove any object associations and to collapse concepts that

had the same meaning, we also wanted to determine the extent to which each association listed from Experiment 1a was agreed upon by participants (consensus), and specific to a particular colour category (specificity). The concept→colour association consensus was defined as the percentage of participants from Experiment 1a that produced the same concept for a given colour term. The concept→colour association specificity on the other hand refers to the specificity of any given concept to one or more colour terms (ranging from 1 to 11 BCT). For example, if participants produced the concept danger 10 times, 7 times associated with the colour red, and 3 times associated with the colour black, then the specificity of the concept danger to the colour red is 70% while the specificity of the same concept to the colour black is 30%.

For the purpose of creating a manageable candidate list of concepts that showed the clearest associations with specific colour terms, inclusion criteria were introduced following each of the two calculations. Therefore, the final list of concepts for each language only included concepts with consensus ≥ 17% (listed by 4 or more participants from the University of Surrey (out of 23 participants) or listed by 6 or more participants from East China Normal University (out of 34 participants)), and specificity ≥ 31% to a particular colour term. Applying these criteria resulted in a final list of 59 English concepts, 30 Chinese concepts from bilinguals, and 33 Chinese concepts from monolinguals.

We ran a hierarchical cluster analysis based on proportional counts for the data from Experiment 1 using correlation-based distance to explore participants' words listed for each colour category. This started with each concept in its own cluster and then, at each stage relaxed the criterion for separation so that the next most similar concept to the existing cluster was linked until all of the concepts were joined in a complete classification tree. Eleven, ten and 9 cluster solutions were selected for the English, bilingual Chinese and monolingual Chinese samples respectively[3]. These are visualised in the Dendrograms shown in **Figures 3-**

5 (full details of the associated colour category, consensus and specificity are listed in

Supplementary Information **Tables 2-4**). It can be seen that each cluster contains related

concepts that are primarily associated with one or a very limited range of colour categories.

Figure 3. Dendrogram of English concept → colour associations produced in Experiment 1a and meeting minimum criteria of consensus and specificity for inclusion in Experiment 2.
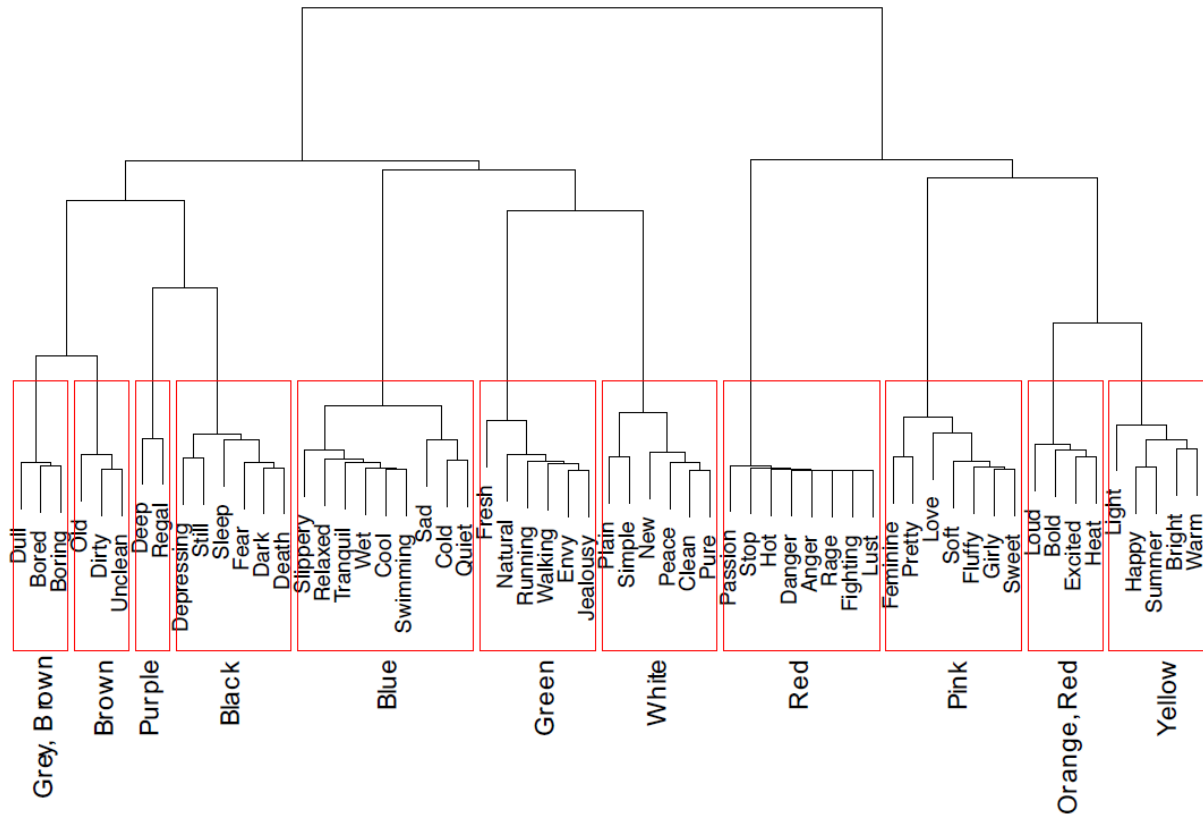
Figure 4. Dendrogram of bilingual Chinese concept → colour associations produced in Experiment 1a and meeting minimum criteria of consensus and specificity for inclusion in Experiment 2.

Figure 5. Dendrogram of monolingual Chinese concept → colour associations produced in Experiment 1a and meeting minimum criteria of consensus and specificity for inclusion in Experiment 2.
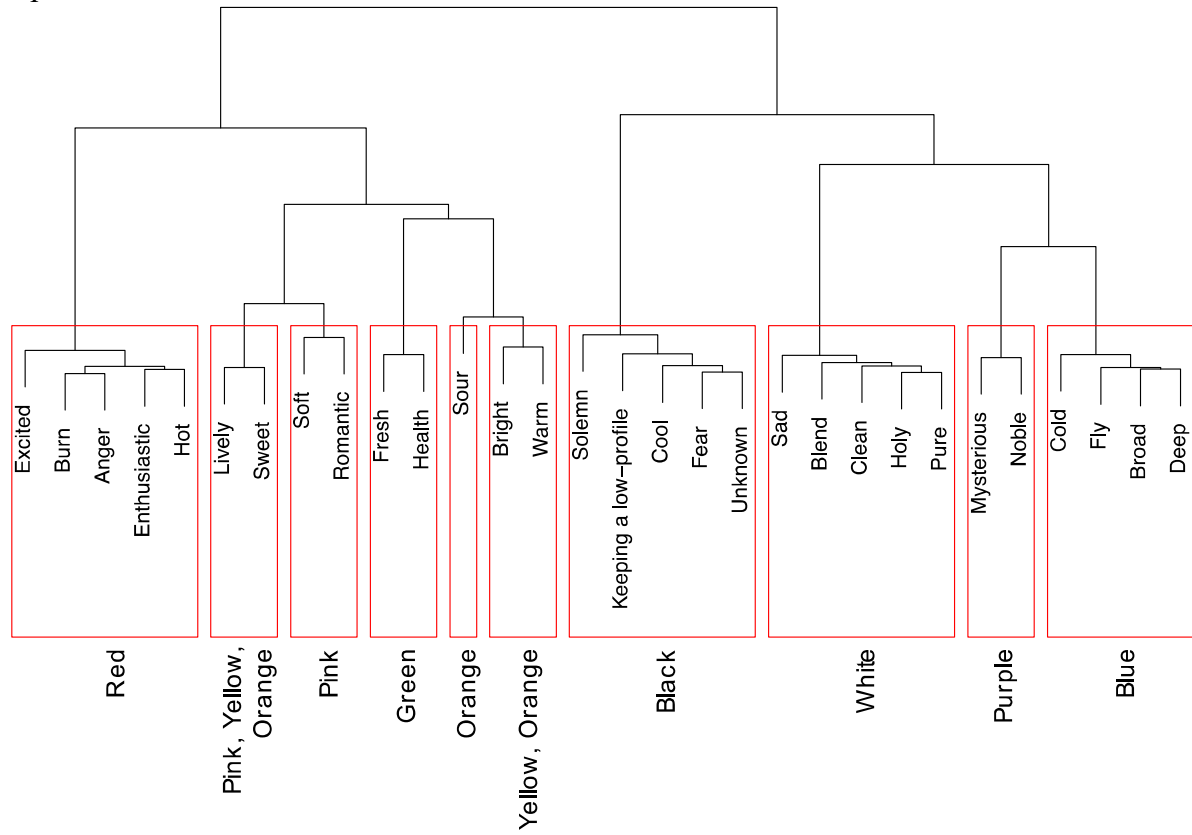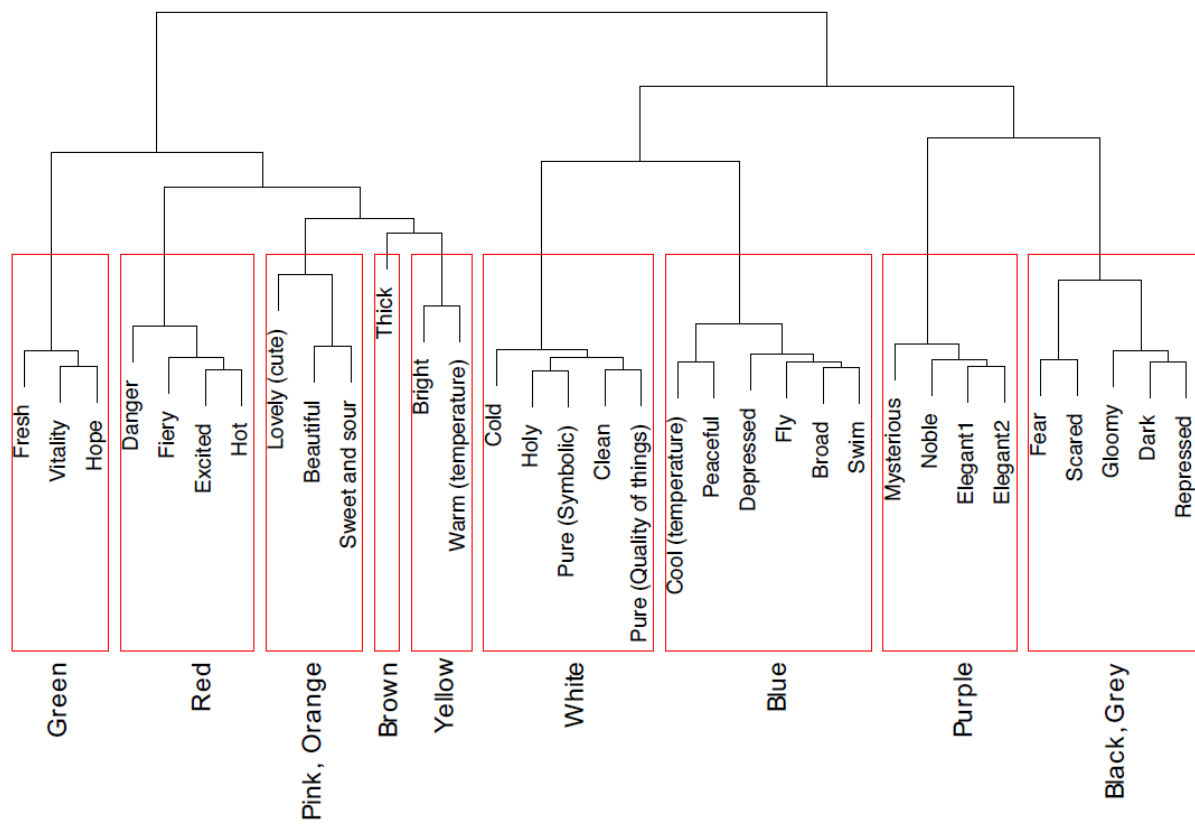
Six English concepts (italicised) exceeded the consensus and specificity thresholds for more than one colour category (*depressing* – black, grey; *love* – red, pink; *loud* – red, orange; *bright* – yellow, orange; *pretty, feminine* – purple, pink), only one Chinese concept from bilinguals exceeded the same thresholds (*fresh* – green, blue), and four Chinese concepts from monolinguals exceeded the same thresholds (*dark* – black, grey; *clean* – white, blue; *cool (temperature)* – green, blue; *hot* – red, orange). However, the English concepts *depressing* and *feminine*, Chinese concept *fresh* from bilinguals, and Chinese concept *hot* from monolinguals were still strongly specified (≥50%) to the respective colours, black, pink, green, and red. In contrast, the remaining four jointly specified English concepts (*love, loud, bright, pretty*) and three Chinese concepts from monolinguals (*dark, clean, cool*) were not as strongly specified to a particular colour (< 50%).

There were also similarities between cultures in the final list of concepts. Eight concepts (hereafter referred to as 'universal' concepts) that were listed by the English (E), Chinese bilingual (CBil) and Chinese monolingual (CMon) participants in the same word category were *excited, fear, clean, fresh, warm, hot, bright, and cold*. The remaining concepts were either produced by at least two cultures (*anger, sad, sweet, soft, deep* – CBil & E; *pure (*symbolic*), holy, fly, broad, mysterious, noble* – CBil & CMon; *depressing(ed), pure (*quality of things*), cool, dark, danger, swim(ing)* – CMon & E), or were culture specific. Although there was universality in concept→colour association (*fear* – black; *clean* – white; *fresh* – green; *warm* – yellow; *hot* - red; *bright* -yellow), two of the universal concepts were listed for a different colour category across languages (*excited* – orange (E), red (CBil & CMon); *cold* – white (CMon), blue (CBil & E), suggesting cross-cultural differences in colours associated with concepts that have similar meaning.

Experiment 2: identifying colour→concept associations

Experiment 2 sought to explore whether there were specific physical colours that were consistently associated with the final list of concepts identified in Experiment 1. Whilst participants might agree that, for instance, *feminine* is associated with the colour category pink, we were interested to see whether participants agreed on exactly which pink(s) was/were *feminine*. This allowed us to see how associations are mapped onto the colour space and explore: (1) whether specific associations are mapped to specific points of the colour space or particular values on colour dimensions (e.g., 'good' concepts lighter colours; 'bad' concepts darker colour); and (2) whether associations relate to entire basic colour categories or specific exemplars within colour categories. Finally, the colour→concept associations from Experiment 2 were checked against the concept→colour associations from Experiment 1 in order to demonstrate associations that are reliable across tasks.

In order for colour to have a common influence on cognition and behaviour there needs to be a demonstrable, specific, mapping between the relevant association and either a specific point in colour space (tile), specific values on a dimension of colour space (e.g. the lightness dimension) or a specific colour category. Furthermore, this mapping needs to be consistent across observers. As a starting point, we chose to present the WCS stimulus grid, which has 320 chromatic colour chips that vary on hue and lightness dimensions and an additional 10 achromatic chips. Thus, in this experiment we required participants to select those colours from the WCS stimulus grid that represented each of the concepts identified in Experiment 1. We were interested in whether these associations were tile specific, categorical or driven by particular values on colorimetric dimensions such as lightness.

Two types of response from the WCS stimulus grid were recorded: *unlimited colour choice* and *best exemplar* for each concept. The unlimited colour choice task required

participants to select as many colour tiles as they associated with each of the concepts. This provided a range of colour→concept associations, showing all the specific points in colour space that participants associated with each of the concepts from Experiment 1. In the best exemplar task, participants were required to choose the single colour tile that they most strongly associated with each of the concepts. This enabled cross checking across experiments to explore whether participants made the same concept→colour and colour→concept associations when their colour choices were restricted and unlimited.

Method

Participants

There were 120 participants comprised of 40 monolingual English speaking participants from the UK (28 females; mean age 21 years, $SD = 1.93$), 40 bilingual Chinese speaking participants from China, resident in the UK (29 females; mean age 21.3 years, $SD = 2.29$), and 40 monolingual Chinese speaking participants resident in China (32 females; mean age 21.4 years, $SD = 1.97$). None had participated in the previous experiments but they were drawn from the same demographic described in Experiment 1. Participants from the University of Surrey were paid £10 for their participation whereas participants from East China University were paid ¥40 for their participation. Participants were also screened for colour vision deficiencies using the City Colour Vision Test (Fletcher, 1981) and Ishihara Test for Colour Blindness (Ishihara, 2009).

Stimuli Displays and Set-Up

Munsell colour tiles were used to reproduce the colour stimuli of the WCS stimulus grid on a Munsell N5 grey board, which was also illuminated by a D65 illuminant. A complete reproduction of the Munsell-WCS stimulus grid (**Figure 1**) was produced on a large board measuring 120cm x 50cm. A clear acrylic sheet with the same thickness as for Experiment 1a was attached above the large Munsell-WCS stimulus grid to protect the tiles. Counters (measuring 0.8cm x 1.8cm) painted in Munsell N5 were used as markers of the colour tiles associated with a particular concept.

The final list of 59 English concepts, 30 Chinese concepts from bilinguals, and 33 concepts from monolinguals from Experiment 1b was presented to the English and Chinese participants respectively.

Procedure

The concepts were presented in a random order. Two different responses were recorded. Firstly, participants were told to freely choose any colour tile(s) from the large Munsell-WCS stimulus grid (using counters spread just below the stimulus grid) that they associated with that concept. As in Experiment 1a and 1b, depending on the native language spoken, participants were provided with verbal instructions either in English, Cantonese, or Mandarin. The translated version (in Cantonese and Mandarin) was back translated to ensure that instructions did not differ across culture. Participants were read out the following instructions:

"Here is an array of colours. I will read a series of words out to you one at a time that you may associate with some of these colours. I would like you to use the counters just below the

array to mark on the clear acrylic sheet any individual colours that you associate with each word I read out. If there is more than one colour that you associate the word with, I would like you to mark them one by one using these counters. You may take as long as you need."

Once participants were satisfied with each of their unlimited colour choices, the experimenter then asked them to choose the best exemplar for the same concept, "*Which colour best represents this word*". Participants pointed to the colour within the stimulus grid that best represented the concept. Their responses were coded during the experiment using Microsoft Excel.

Results and Discussion

First, we establish whether participants systematically map concepts to specific points in the colour space defined by co-ordinates on the dimension of hue, saturation, and lightness: tile-concept association. Second, we establish the extent to which particular concepts were associated with particular lightness ranges (lightness-concept association) and colour categories (category-concept association). We did not analyse saturation-concept association because the WCS stimulus grid has deliberately high levels of saturation for all hues. Finally, we compare data from Experiment 1 and Experiment 2 to assess consistency between listing concepts associated with colours (concept→colour association) and choosing colours associated with concepts (colour→concept association).

As there were two datasets in this experiment (unlimited colour choice and best exemplar for each concept), we use the frequency data from the unlimited colour choice and best exemplar for each concept to explore consistency for tile-concept, lightness-concept, and category-concept associations. Illustrative visualisations for these data can be seen in Figure

6, which shows contour plots for the eight 'Universal' concepts identified in Experiment 1

(see Supplementary Information **Figure 1** for a full set of contour plots of the Experiment 2

data for all words and language groups).

Figure 6. 'Contour maps' for 8 universal concepts (identified in Experiment 1) in English (E: left), bilingual Chinese (CBil: middle), and monolingual Chinese (CMon: right). Contour lines show the frequency with which each colour tile was selected for each concept in the unlimited colour choice. White dots represent the best exemplar for each concept.
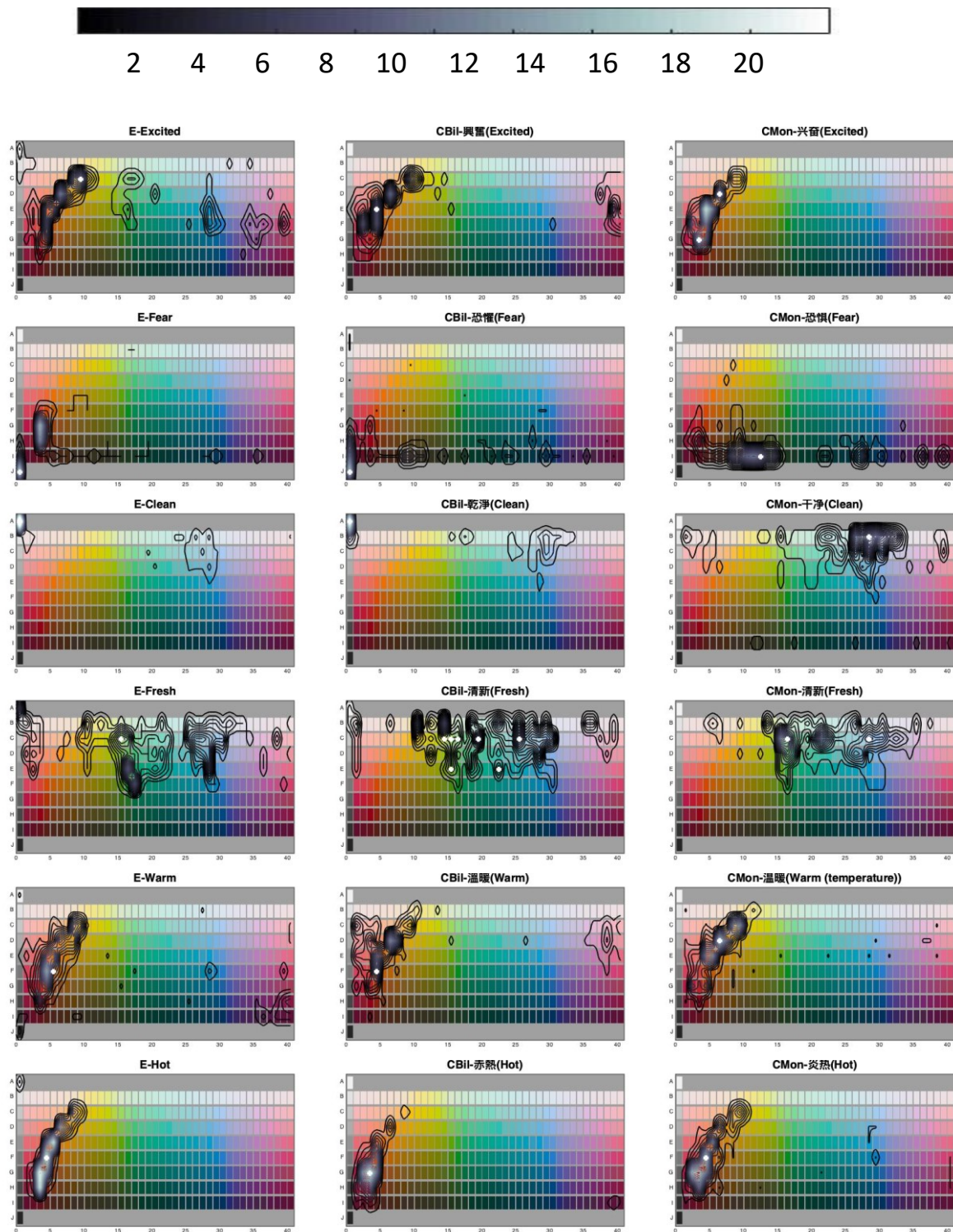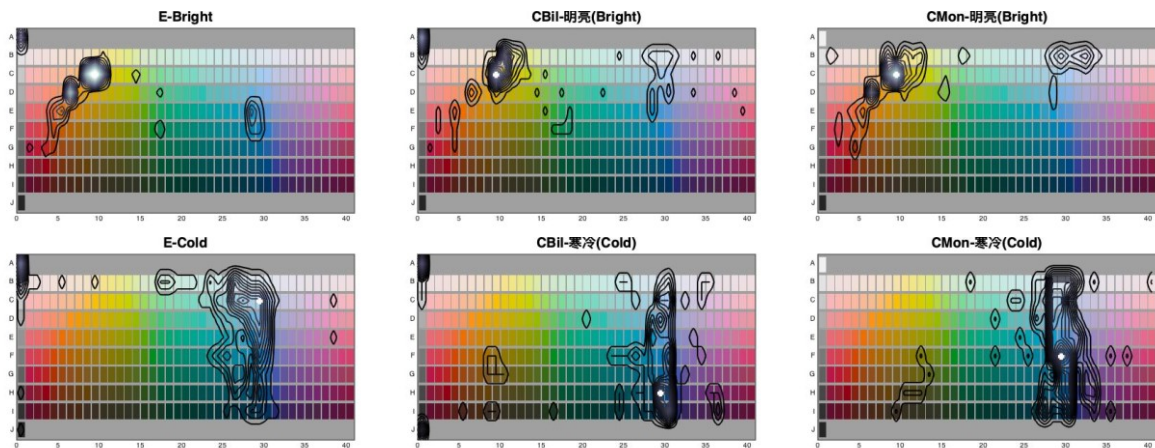
The contour plots show both clear universal tendencies in terms of the colours selected for specific concepts by participants in Experiment 2 as well as clear differences. For example, the contours for the concept '*excited*' were very similar across languages in terms of the frequency with which different tiles were chosen and therefore the colour categories associated with this concept, but it is clear that the best exemplar differed considerably across language group with English participants selecting a much lighter hue than Chinese monolingual participants. In contrast, for the concept '*clean*', the English and Chinese bilingual participants selected very different hues as the best exemplar (English & Chinese bilingual – white; Chinese monolingual – blue), but all groups selected a similarly light hue. Finally, for some concepts all three groups show similar contours and selected the same best exemplar, as seen for the concept 'bright'. We next explore these tile specific, lightness specific and categorical associations with colour in more detail.

Tile-concept association

In the unlimited colour choice response, the 40 English speaking participants made a total of 3881 colour tile selections associated with 59 concepts, the 40 bilingual Chinese-speaking participants made a total of 1618 colour tile selections associated with 30 concepts, whereas the 40 monolingual speaking participants made a total of 2076 colour tile selections

associated with 33 concepts. We wanted to know whether participants systematically mapped colour tiles to concepts and whether this mapping was consistent across participants in terms of the unlimited colour tiles selected and in terms of the best exemplar for each concept.

There were differences in the strength of tile-concept associations when English and Chinese speaking participants were allowed unlimited choice and a best tile choice for each concept. When participants were allowed unlimited choice for each concept, the English group consistently made more high agreement (≥50% of participants associating a particular colour tile with a given concept) tile-concept associations (0J – *dark, death;* 0A – *clean, plain, pure, simple;* 3G – *danger, anger;* 3H – *fighting, rage*; 9C – *bright*) than the bilingual Chinese group (0A – *holy, clean;* 3G – *burn*) and the monolingual Chinese group (3G – *fiery*). Across the groups these high agreement associations were limited to black, white, red, and yellow tiles. When participants were asked to choose one tile that best represents that concept, only black and white tiles were consistently chosen (≥50% of the time) for the English concepts *dark* and *death* (0J), and *clean* (0A). However, none were consistently chosen for any of the Chinese concepts.

Overall, 29 English concepts (*Dark, Death, Depressing, Fear, Clean, Pure, Plain, New, Simple, Rage, Anger, Danger, Fighting, Lust, Passion, Happy, Bright, Swimming, Deep, Still, Sleep, Peace, Sweet, Fluffy, Bold, Heat, Dull, Boring, Bored*), 11 Chinese concepts from bilinguals (*Solemn, Fear, Holy, Clean, Pure, Blend, Burn, Hot, Fly, Cool/ruthless, Noble*), and 10 Chinese concepts from monolinguals (*Gloomy, Fiery, Danger, Bright, Holy, Pure* (quality of things), *Cold*, *Swim, Lovely* (cute), *Warm* (temperature) showed agreement between the colour tile chosen with greatest consensus when *unlimited colour choice* was allowed and the tile chosen as the *best exemplar* for the concept (see the tile column in Supplementary Information **Tables 5, 6 and 7** for the associated tile in each case).

Lightness-concept association

For the lightness-concept association analysis, the tiles were collated into three lightness ranges by dividing the WCS stimulus grid (**Figure 1**) into three sections. The first three rows (A, B, and C) were classified as 'Light' (Munsell values 8, 9, 9.5), the middle four rows (D, E, F, and G) were classified as 'Medium' (Munsell values 4, 5, 6, 7) and the remaining three rows (H, I, and J) were classified as 'Dark' (Munsell values 1.5, 2, 3). In this analysis, we were interested in which particular concepts were associated with a range of colour tiles that are consistently categorised as one of the three lightness range in both unlimited choice and best exemplar dataset. In other words, which concepts are specified to light, medium, and dark across both tasks? We chose to focus on concepts with less than 50% consensus to a specific tile in order to ensure that the majority of the associations were not just with one tile (i.e., *dark and death* – 0J; *clean* – 0A) that happens to be dark or light. For the English data, 56 English concepts met the criterion whereas for the Chinese data all 30 Chinese concepts (from bilingual Chinese) and all 33 Chinese concepts (from monolingual Chinese) met the criterion. Of these, 41 (or 73%) English concepts ('light' lightness category - *Light, Simple, Plain, Quiet, Pure, Soft, Still, Fluffy, New, Peace, Fresh, Tranquil*; 'medium' lightness category – *Swimming, Wet, Running, Love, Feminine, Jealousy, Hot, Pretty, Excited, Girly, Heat, Loud, Warm, Envy, Passion, Slippery, Bold, Relaxed, Walking*; 'dark' lightness category – *Depressing, Deep, Dirty, Unclean, Fear, Fighting, Sad, Old, Anger*), 21 (or 70%) Chinese concepts from bilinguals ('light' lightness category – *Pure, Holy, Clean, Fresh, Bright, Soft*; 'medium' lightness category - *Enthusiastic, Excited, Warm, Romantic, Burn, Sweet, Hot, Lively, Broad, Fly*; 'dark' lightness category – *Deep, Fear, Solemn, Mysterious, Sad*) and 29 (or 88%) Chinese concepts from monolinguals ('light' lightness category – *Pure* (symbolic), *Pure* (quality of things), *Clean, Bright, Fresh, Lovely* (cute), *Holy*; 'medium' lightness category – *Excited, Swim, Hot, Fiery, Warm* (temperature), *Broad,*

Beautiful, Vitality, Elegant1, Cool (temperature)*, Sweet & Sour, Fly, Peaceful, Hope*; 'dark'

lightness category – *Gloomy, Fear, Dark, Repressed, Scared, Thick, Mysterious, Danger*) had

clear association to a particular lightness range (i.e. ≥50% of tiles selected in both datasets

fell in the same lightness range; see Supplementary Information **Table 8**). It seems clear that

both the English and Chinese speaking participants tended to associate positive concepts with

lighter colours and negative concepts with darker colours.

The remaining 17 English, 9 bilingual Chinese and 4 monolingual Chinese concepts

had less than 50% association to any one of the three-lightness ranges.

Colour category-concept association

We used the colour naming data from two separate studies to label each of the 330

tiles in the WCS stimulus grid with one of the 11 basic colour terms (BCT). As for

Experiment 1, Berlin and Kay's (1969, 1991) data were used to determine the name of each

tile associated with a BCT (see **Figure 2**). Non-BCT tiles were defined as those tiles not

belonging to any of the 11 BCT categories in Berlin and Kay's English colour naming data.

Although most of the tiles selected by English, bilingual Chinese, and monolingual Chinese

participants fell within one of Berlin and Kay's BCTs (73%, 75%, and 67.5% respectively),

some BCT tiles were selected only marginally more often than non-BCT tiles. This includes

the English concept '*unclean*' (51% of tiles selected were BCT tiles), and monolingual

Chinese concepts, *'sweet and sour'* (52%), *'bright'* (51%), and *fresh'* (50%). Next, we

extended the colour categories by using Lindsey and Brown's (2014) colour naming data to

identify the colour category of all those tiles in the WCS that were unnamed in Berlin and

Kay's data. Thus, all tiles in the WCS were categorised into one of the 11 basic colour

categories. This classification system was used to explore the frequency of colour-concept

associations in relation to these broader colour categories. In this analysis, we were

interested in (1) whether the choice of colour category (colour category → concept associations) was similar for the unlimited choice dataset and the best exemplar dataset; and (2) whether the most consistent concept-colour category association (Experiment 1) is similar to the most consistent colour category-concept association (Experiment 2).

To examine whether the participants' choice of colour category associated with a concept was similar or different across the unlimited choice and best exemplar datasets we conducted a likelihood ratio chi-square test (LRT) for each word on a two-way contingency table of colour category by dataset. The LRT assumes that colour responses to each word follow a multinomial distribution. We used a LRT, rather than a chi-square test of independence, as there were cells with zeros (i.e. for a given word there were some colours that were never associated with that word). There was no evidence of independence of the colour category choices for the best exemplar and unlimited choice datasets for any concept for any language group, thus, suggesting that there is consistency in the colour category-concept association across both datasets. As both the unlimited choice dataset and the best exemplar dataset were not independent, we decided to use data from the best exemplar task to illustrate any universal, common, and cultural specific themes emerging from the category-concept association.

Emerging themes in category-concept associations. To explore participants' patterns of responding on the 11 colour categories, as for Experiment 1, we ran a hierarchical cluster analysis based on proportional counts for the best exemplar dataset using correlation based distance. This started with each concept in its own cluster and then, using participants' best exemplar for each concept response, at each stage relaxed the criterion for separation so that the next most similar concept to the existing cluster was linked until all of the concepts were joined in a complete classification tree. An eleven clusters solution was suggested for

the English participants, a ten clusters solution was suggested for the bilingual Chinese

participants, whereas a 9 clusters solution was suggested for the monolingual Chinese

participants (see **Figures 7, 8,** and **9** for colour→concept dendrograms for each population;

see also Supplementary Information **Tables 5, 6,** and **7** for full details)[3].

Figure 7. Dendrograms of English colour → concept associations derived from the best exemplar tile data collected in Experiment 2.
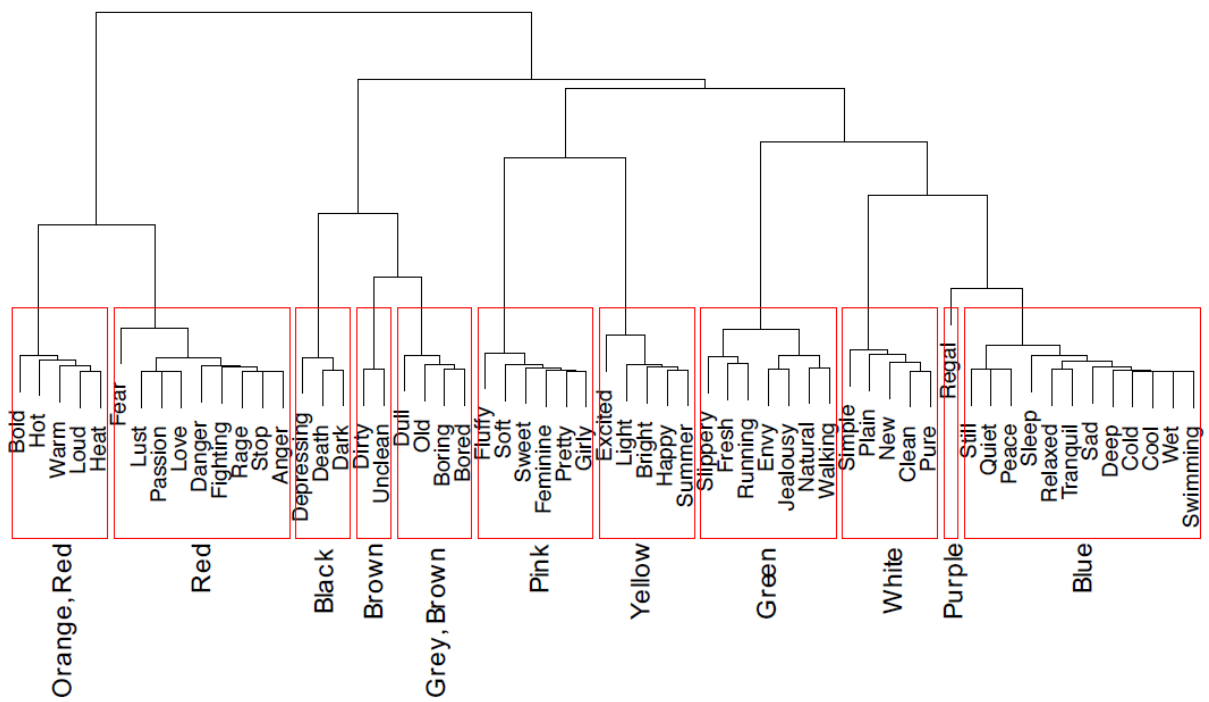
Figure 8. Dendrograms of Chinese bilingual colour → concept associations derived from the best exemplar tile data collected in Experiment 2.
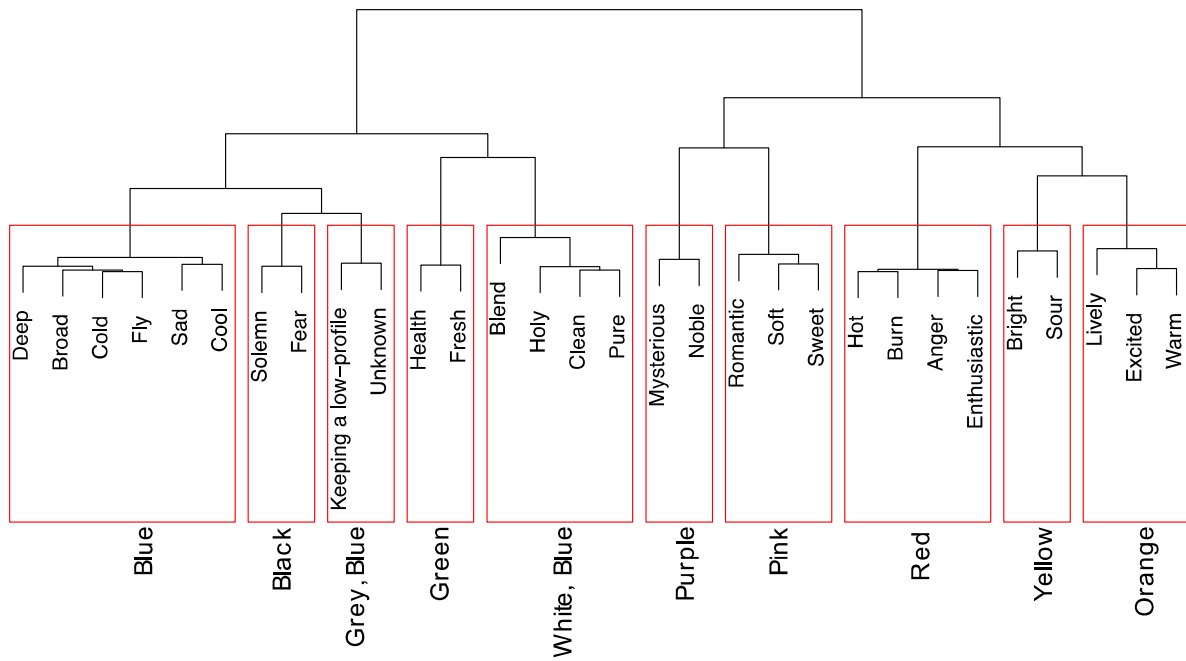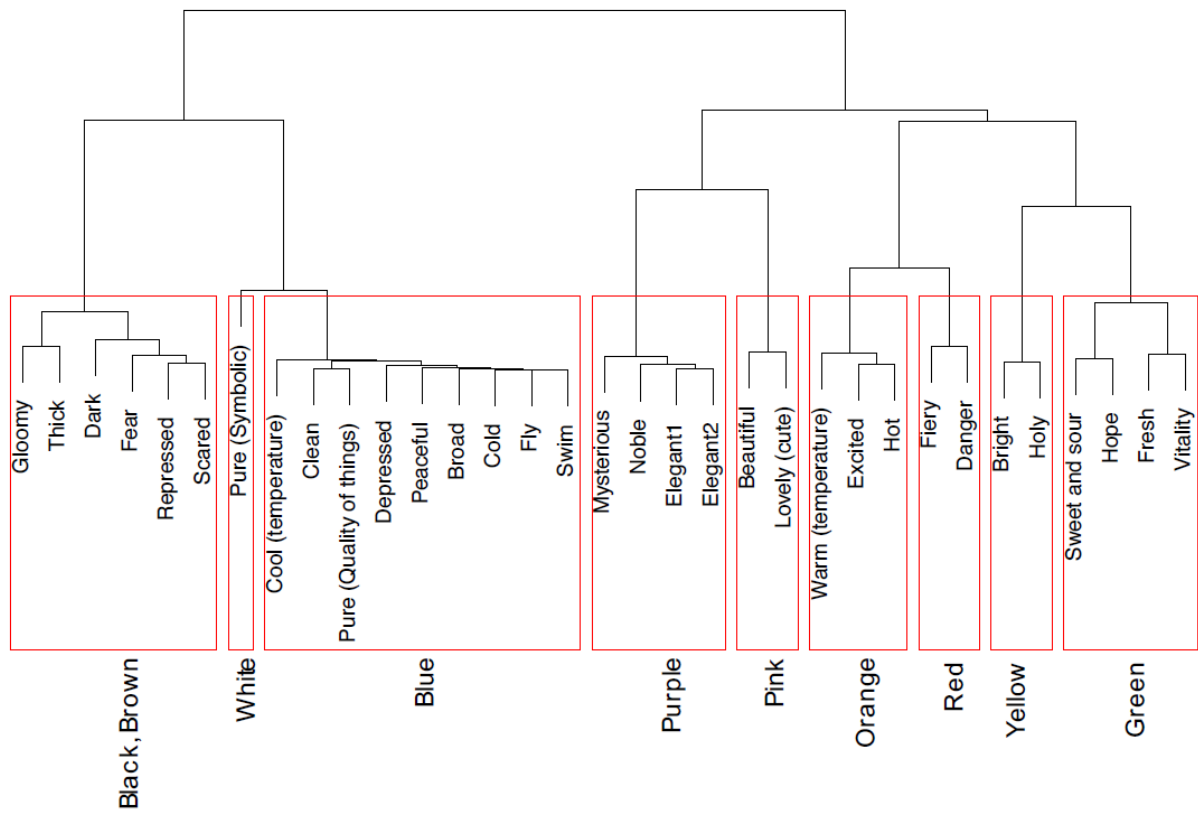
Figure 9. Dendrograms of Chinese monolingual colour → concept associations derived from the best exemplar tile data collected in Experiment 2.

Although there were 11 clusters for 59 English concepts, 10 clusters for 30 Chinese concepts (from bilinguals), and 9 clusters for 33 Chinese concepts (from monolinguals), each associated with particular colour categories, these could be further split into themes within each colour category. Some themes were universal whereas others were culture specific. Universal colour category themes across all three cultures were: *negative concepts* – black; *fresh, health* – green; *purity* – white; *water / sky related* – blue; *regal* – purple; *'female' traits* – pink; *bright* – yellow (see Supplementary Information **Tables 5-7** for full set of themes). There were also culture specific themes when associating colours with concepts (e.g. English, *boring / old* – grey, brown; Chinese bilingual, *keeping a low profile* – blue, grey).

Further emphasising the important role of lightness in conceptual colour associations, words within a colour category could also be grouped around light, medium, and dark colours, particularly, for the colour categories blue, pink, red, yellow, purple, and green. For the English concepts, groupings were dark blue (DBlue) –*sad, deep;* medium blue (MBlue) - *swimming, cool, wet, relaxed, slippery;* light blue (LBlue) – *still, quiet, peace, tranquil;* medium pink (MPink) – *girly, feminine, sweet, pretty;* light pink (LPink) – *soft, fluffy;* dark red (DRed) – *fighting, fear;* medium red (MRed) – *lust, passion, love;* medium green (MGreen) – *envy, jealousy, walking, running, slippery*; light green (LGreen) – *fresh;* medium yellow (MYellow) - *excited*; light yellow (LYellow) – *light*.

For bilingual Chinese concepts: DBLue – *deep, sad;* MBlue – *broad, cool (temperature), fly;* MPink – *romantic, sweet;* LPink – *soft;* DRed – *anger;* MRed – *enthusiastic, burn;* MGreen – *health;* LGreen – *fresh.*

For monolingual Chinese concepts: MBlue – *cool (temperature), peaceful, broad, fly, swim;* LBlue – *clean, pure (quality)*; medium pink (MPink) – *beautiful;* light pink (LPink) – *lovely (cute)*; DRed – *danger;* MRed – *fiery;* MGreen – *hope, vitality, sweet and sour;* LGreen – *fresh;* dark purple (DPurple) – *mysterious;* medium purple (MPurple) – *elegant1.*

Thus, lightness and hue can be seen to interact when specifying conceptual colour associations.

It is worth noting that these themes are identified in relation to the data from experiment 2 (closed-ended task) where participants were provided with 330 tiles that varied in both hue and lightness. Next, we explore themes that are consistent between Experiments 1 and 2.

Comparison between open-ended (Experiment 1) and closed-ended (Experiment 2) tasks.

To determine whether the most consistent concept→colour (Experiment 1: open-ended task) and colour→concept (Experiment 2: closed ended task) associations were similar or different, we compared the data from Experiment 1 with the best exemplar data (colour category of the best exemplar) from Experiment 2 (see Supplementary Information **Tables 5, 6, and 7** for respective associations between colour category and concept in Experiments 1 & 2, for English, bilingual Chinese, and monolingual speakers respectively). For each word we conducted a likelihood ratio chi-square test (LRT) on a two-way contingency table of colour category by experiment. Given that multiple tests were being conducted for each language group, we used Bonferroni corrected p values to assess statistical significance. Overall, the results suggested that for most concepts there was no significant evidence that experiment 1 and 2 responses were independent. The only exceptions were '*cold*' ($X^2(10) = 45.89\ p < .001$), '*gloomy*' ($X^2(10) = 32.65\ p < .001$), and '*scared*' ($X^2(10) = 29.61\ p = .001$) for Chinese monolinguals. It is plausible that these concepts in monolingual Chinese participants were mostly associated with dark in the lightness dimension (in Experiment 2) rather than just relying on the hue dimension (in Experiment 1). For example, for the concept *gloomy,* monolingual Chinese participants in experiment 1 strongly associated *gloomy* with grey

whereas participants in experiment 2 strongly associated *gloomy* with blue and brown, and were choosing tiles that were dark in the lightness dimension.

Although the likelihood ratio test showed that the majority of concept→colour category associations generated in Experiment 1 were consistent with the colour category→concept associations generated in Experiment 2, the cluster analyses suggested that there were still notable differences between the open-ended task used in Experiment 1 and the closed-ended task used in Experiment 2. Comparing the cluster analysis solutions generated for each language group and experiment (Experiment 1, figures 4–6; Experiment 2, figures 7–9), it is clear that some words were consistently clustered together across both experiments (e.g. English, *dull, boring, bored* – grey, brown) whereas others were not (e.g. Chinese monolingual *sweet and sour* – Experiment 1 pink, orange, Experiment 2, green).

Next, using the themes identified following the cluster analysis of data from Experiment 2, we explored which of these themes could also be applied in a consistent way to the cluster solutions from Experiment 1, with the additional constraint of only including concepts that were most strongly associated with the same colour across experiments (see Supplementary Information **Tables 5-7)**. **Table 1** summarises theme-colour associations that were consistent across Experiment 1 and Experiment 2 for all language groups ('universal' themes), across two language groups ('common' themes), or specific to one language group. Overall, the proportion of associations that were consistently clustered together in the same colour category across both experiments was higher in the English group (84.7%) than the Chinese bilinguals (73.3%) and the lowest in the Chinese monolinguals (63.6%).

Table 1. Theme-colour associations for each culture that were consistent between Experiments 1 and 2 (actual concepts are in brackets). Themes are grouped according to universal (across three culture), common (across two culture), and culture specific. Dominant colours are listed first.

		Colour category (words) for each culture		
		English	Chinese Bilingual	Chinese Monolingual
Universal themes	Clean, pure	White (clean, pure, new)	White (clean, pure, holy)	White (pure (symbolic)) & blue (clean)
	Negative concepts	Black (death, dark)	Black (solemn, fear)	Black (scared, dark)
	Water/sky related	Blue (slippery, cold, cool, wet, swimming)	Blue (deep, broad, fly)	Blue (cool, broad, fly, swim)
	Fresh, health, nature	Green (fresh, running, natural, walking)	Green (health) & blue (fresh)	Green (vitality, fresh)
	'Female' traits	Pink (soft, sweet, feminine, pretty, girly)	Pink (soft, sweet)	Pink (lovely)
	Heat	Orange (heat)	Red (burn, hot)	Red (fiery), orange (warm (temperature))
	Bright	Yellow (light, bright)	Yellow (bright)	Yellow (bright)
	Regal	Purple (regal)	Purple (noble)	Purple (elegant1, elegant2, noble)
Common themes	Anger	Red (anger, fighting, rage)	Red (anger)	
	Stop / danger	Red (stop, danger)		Red (danger)
	Enthusiastic / excited		Red (enthusiastic)	Orange & red (excited)
	Depressed	Black (depressing)		Blue (depressed)
	Relaxed	Blue (quiet, relaxed, tranquil)		Blue (peaceful)
	Mysterious		Purple (mysterious)	Purple (mysterious)
Culture specific themes	Attraction	Red (lust, passion)		
	Bold	Orange (bold, loud)		
	Summer	Yellow (summer, happy)		
	Sad	Blue (sad)		
	Old, boring	Grey & brown (dull, old, boring, bored)		
	Plain	White (plain, simple)		
	Dirty	Brown (dirty, unclean)		
	Envy	Green (envy, jealousy)		
	Fluffy	Pink (fluffy)		
	Blend		White (blend)	
	Romantic		Pink (romantic)	
	Hope			Green (hope)
	Thick			Brown (thick)

In addition to concepts clustering according to colour dimension, concepts within a colour category were also grouped around positive, neutral, and negative concepts. Specifically, this was found for the colour categories green, red, and blue. For the English concepts groupings were green positive (GP) – *fresh, running, walking;* green neutral (GNeu) –*natural;* green negative (GNeg) – *jealousy, envy;* red positive (RP) – *passion, lust*; red negative (RNeg) – *fighting, anger, rage, stop, danger;* blue neutral (BNeu) – *slippery, cold, cool, wet, swimming*; blue negative (BNeg) – *sad.* This is also found in bilingual Chinese concepts: RP – *enthusiastic;* RNeg – *anger, hot, burn;* GP – *fresh, health.* Whereas in monolingual Chinese concepts, they were RP – *excited;* RNeg – *danger, fiery, hot;* GP – *fresh, vitality;* blue positive (BP) – *peaceful, broad;* BNeu – *cool (temperature), broad, fly, swim,* BNeg - *depressed.* This suggests that colour associations in particular, for the colour category red, green, and blue may relate to positive, neutral, and negative concepts in English and Chinese culture.

Finally, for those associations that were inconsistent between Experiment 1 and Experiment 2, it is likely that those concepts are loosely associated with more than one colour category or that those concepts are more associated with a range on the lightness dimension instead of the colour category dimension. For example, for the former, the concept *lively* (CBil) is associated with pink and yellow in Experiment 1 but orange and green in Experiment 2, and the concept *beautiful* (CMon) is associated with orange in Experiment 1 but pink in Experiment 2. As for the latter, the concept *peace* (E) is perhaps associated more with the light colours (Experiment 1 – white; Experiment 2 - light blue and white) than a particular colour, and the concept *unknown* (CBil) is perhaps more associated with the dark colours (Experiment 1 – black; Experiment 2 - dark grey, dark blue) than a particular colour.

General Discussion

This systematic investigation aimed to explore and compare the range of colour associations for concepts in English and/or Chinese speaking adults. In addition, we were also interested in whether these associations are categorical, specific to particular exemplars of colours or driven by lightness. By understanding the associations made with different colours, we can then develop systematic predictions about the effects of a wider range of colours on cognition and behaviour.

The range of colour associations with concepts, meeting minimum criteria for consensus and specificity, was greater for English speaking adults (59 concepts) than for bilingual Chinese speaking adults (30 concepts) and monolingual Chinese speaking adults (33 concepts). Across open-ended and closed-ended tasks, there were 'universal' (across all culture), 'common' (across two culture), and culturally specific associations evidence from the present data and those reported by previous research (e.g., Bonnardel et al., 2014; Prado-Leon et al., 2006). The 'universal' themes associated with specific colour categories were: *purity* – white; *negative* concepts – black; *fresh, health, nature* – green; *water/sky related* – blue; *regal* – purple; *'female' traits* – pink; *bright* – yellow; *heat* – orange, red. The 'common' themes associated with specific colours were: *depressed* – black & blue; *anger, stop/danger, enthusiastic/excited* – red; *relaxed* – blue; *mysterious* – purple. These associations are consistent with previous findings on conceptual colour pairings (e.g., Dael et al., 2016; Frank & Gilovich, 1988; Hanafy & Reham, 2016; Ludwig & Simner, 2013; Mohammad, 2011; Palmer et al., 2013; Soldat, Sinclair, & Mark, 1997). In contrast, there were also culturally specific themes/concepts shown by the Chinese bilingual participants that were not seen in either the Chinese monolingual or the English participants and vice

versa. For example, only the Chinese bilingual group associated pink with *romantic* and white with *blend*.

Some concepts were found to be associated to a particular colour category (of the English Basic Colour Terms) that was consistent across open-ended and closed-ended tasks, suggesting strongly established conceptual colour associations. For the English population, all 11 colour categories had strong association with at least one concept whereas for both the Chinese bilingual and Chinese monolingual there were eight (all bar grey, orange and brown) and 10 (all bar grey) colour categories respectively with strong association to at least one concept. Interestingly, orange and brown, which had no consistently associated concepts across tasks in the bilingual group, are not basic colour terms in Mandarin or Cantonese (Gao & Sutrop, 2014) and grey is amongst those in the last group of colours to emerge in Berlin & Kay's hierarchy. Thus, it may be that less established colour terms are also used less consistently (if at all) with respect to conceptual colour associations.

Critically, when participants were provided with a large array of colour tiles to choose from, we found that they made choices that were not always related to whole colour categories. First, we found that associations can also be to particular points in colour space. For example, tiles 0J (black), 0A (white), 2G (red), 3G (red), 3H (red), 9C (yellow) and so forth were shown to be specifically associated to certain concepts (e.g. 3G - *anger*). These specific tile mappings were largely to colours from the top positions in Berlin and Kay's hierarchy of colour terms (black, white, and red) particularly in the English speaking and Chinese bilingual samples. Instead of relying on just hue category information in making predictions about colour-behaviour relationships (e.g., red-*danger*), the current work, which used a systematic and standardised colour system, demonstrated that there is a point in colour space (e.g., tile 3G) that is most strongly associated with a particular concept (e.g., *danger*). A likely reason as to why specific black and white tiles were not consistently selected by the

Chinese monolingual sample is that this group may be relying on other dimensions of colour (e.g., lightness) when making their tile choices. For example, as a whole Chinese monolingual participants made proportionally more associations that were aligned to one of the three lightness categories (91.3%) than the English participants (66.1%) and the Chinese bilingual participants (70%).

Second, we found that particular groups of concepts were associated with particular lightness ranges such that negative concepts were associated with the darker colours whereas positive concepts were associated with the lighter colours. There is considerable evidence to show that one particular dimension of colour, namely lightness, has an intrinsic emotional meaning (Palmer, Schloss, Xu, & Prado-Leon, 2013). Meier and colleagues showed that in a categorisation task positive (vs. negative) words are evaluated faster and more accurately when they are presented in white (vs. black) font (Meier, Robinson, Clore, 2004). In a large-scale cross-cultural study on affective meanings of colour terms (Adams & Osgood, 1973), individuals from 23 cultures rated eight colour terms (red, black, grey, red, yellow, blue, green, or white) using affective adjectives. In all cultures, white was evaluated positively whereas black was evaluated negatively. Other experimental studies confirm these valence-based associations with lightness using adjectives relating to happiness (Wright & Rainwater, 1962; Dael et al., 2016), affective states (Hemphill, 1996), and positive versus negative ratings (Valdez & Mehrabian, 1994) even at an early age (Stabler & Johnson, 1972).

Whilst our study supports the notion of light-positive and dark-negative associations, previous work suggests that such opposite meanings may also be carried by colour categories themselves. For example, given that red is paired with *danger* in daily life, and given the importance of avoiding negative possibilities (Baumeister et al., 2001), red may carry negative meaning whereas green has more general positive associations and carries positive meaning (Moller, Elliot, & Maier, 2009). Further, consistent with this notion, red has been

shown to enhance memory for negative words, whereas green increases it for positive ones (Kuhbandner & Pekrun, 2013). However, concepts in our study seem to cluster around positive, neutral, and negative themes <u>within</u> the colour categories green and red suggesting that the picture is not as simple as suggesting red is associated with negative meaning and green positive. For instance, for English concepts, there were positive, neutral, and negative themes associated with green, and positive and negative themes associated with red. Although we found similar positive and negative themes for the colour category red in Chinese bilinguals and Chinese monolinguals, green was only associated with a positive theme. Interestingly, the findings for the Chinese sample on these colours correspond with a recent finding on Indian participants. Namely, Bonnardel and colleagues (2014) demonstrated that Indian participants elicited concepts that coincide with red-green opponency. Concepts for the colour category red were categorised into positive (e.g. love, euphoria) and negative (e.g. jealousy) emotion words that were high energy, whereas concepts for the colour green were categorised into just low energy, positive emotion words (e.g., fresh, peaceful). This suggests that there may be cultural differences between Eastern and Western cultures in patterns of red/green, positive/negative associations.

Additionally, a relationship between lightness and colour was found for the colour categories red, yellow, green, purple, blue, and pink. For example, whilst for both English participants and monolingual Chinese participants light green was associated with *fresh*, conversely medium green was associated with *envy, jealousy, walking, running* and *slippery* for English participants but with *hope, vitality and sweet & sour* for monolingual Chinese participants. These different patterns of relationship suggest that light, medium, and dark colours can carry specific meaning dependent on the culture. Whilst there was consistency across culture in that for all language groups concepts were clustered around dark and medium lightness levels for the colour category red (dark red was associated with *fighting*

(E), *fear* (E), *anger* (CBil), and *danger* (CMon); medium red was associated with *lust* (E), *passion* (E), *enthusiastic* (CBil), and *fiery* (CMon)), and around light and medium lightness levels for the colour categories pink and green, further cultural differences were found in clustering for the colour category blue and purple. Concepts were clustered around light, medium, and dark lightness levels for these colour categories for the English culture whereas they were clustered around medium and dark lightness levels for the Chinese bilingual culture, and light and medium lightness level for the Chinese monolingual culture. As for the colour category purple, only Chinese monolingual participants had concepts around medium and dark lightness level (medium purple – *elegant;* dark purple – *mysterious).*

Overall, our results confirm the particular importance of lightness to positive and negative based associations. Further, they demonstrate that red and green do not necessarily carry opposite meanings but that those colours can carry both positive and negative meaning in a culturally dependent way, and that concepts are not necessarily clustered around the dimension of hue but instead interact with lightness as well.

Implications for effects of colour on psychological functioning

There has been great interest in colour among psychologists, reflecting the fact that colour is an omnipresent feature of our environment and plays a role in information processing. It has been argued that the influences of colour on cognition and behaviour are a result of learned associations (Mehta & Zhu, 2009). Thus, we will discuss the psychological effect of the colours red, blue, and green (as have been widely investigated in this area) in the context of the present generic associations and make predictions for future studies in relation to the current findings.

Red. Previous studies addressing the effect of colour on psychological functioning have mainly focused on the colour red, showing that red heightens attention (Diaz-Roman et al., 2015; Pravossoudovitch, Cury, Young, & Elliot, 2014), undermines performance (Shi, Zhang, & Jiang, 2014), increases aggressiveness and dominance (Bagchi & Cheema, 2013; Fetterman, Liu, & Robinson, 2015; Krenn, 2014), and enhances attraction (Elliot & Niesta, 2008) dependent on its context (see Elliot & Maier, 2014, for a review). It has been suggested that many of these effects are the result of learned colour associations. Our findings demonstrated that in the present generic context, red (especially the tiles 3G and 3H) is strongly associated with concepts that represent aggression (*anger, fighting, rage*), attraction (*lust, passion*), and heighten attention (*stop, danger*) in English speaking participants.

Interestingly, there are clear cross-cultural differences in associations with red. Whilst Chinese bilingual speaking participants also associated concepts to do with aggression (*anger*) with the colour red, the concept that best represents attraction (*romantic*) was associated with the colour pink, and no concepts associated with heightened attention were listed. In contrast, Chinese monolingual speaking participants showed no concepts associated with aggression but one concept associated with heightened attention (*danger*). Like their Chinese bilingual speaking counterparts, Chinese monolingual speaking participants associated the colour pink with the concept that best represents attraction (*beautiful*). In addition, there were culturally specific associations with red that were not found in the English culture, *enthusiasm* (CBil) and *excited* (CMon).

Importantly, there was no specific tile that was consistently associated with aggression or attraction in either Chinese bilingual or Chinese monolingual participants. Although English speaking participants associated specific tiles (3G and 3H) with the concepts representing aggression, Chinese bilingual speaking participants associated a particular lightness dimension (dark) with aggression. For the association with attraction,

Chinese speaking participants associated the concepts *romantic and beautiful* with the colour pink. This was not specified to a particular tile but with tiles of medium lightness. English speaking participants on the other hand associated concepts representing attraction with the colour red (3G and 3H) and with tiles of medium lightness.

Two concepts relating to heightened attention were strongly associated with the colour red for English participants (*stop* and *danger*), but only one concept was found in our Chinese monolingual participants (*danger*) and none was found in our Chinese bilingual participants. The cross-cultural differences for these concepts reflect similar patterns of difference found in previous work (Bergum & Bergum, 1981; Courtney, 1986; Chan & Courtney, 2001) on US, Yunnan Chinese, and Hong Kong Chinese population (*stop* – red (US 100%; Yunnan 48.5%; HK 66.4%); *danger* – red (US 89.8%; Yunnan 64.7%; HK 63%)). S*top* was not strongly associated with the colour red in the Chinese speaking population in comparison to the English speaking population. In fact, in both the Chinese bilingual and Chinese monolingual speaking participants from our study, the concept stop was only listed twice in our list task for each population. In addition, the concept *danger* was listed in more than one colour category in the Chinese bilingual participants which meant that concepts relating to heightened attention are not as strongly associated with the colour red for Chinese speaking participants.

A proposed association between the colour red and failure has been consistently considered to be a major contributing factor to poor cognitive performance in achievement contexts when presented with the colour red. It has been suggested that these effects (and by implication this association) may be highly specific to this task context as argued by Elliot and Maier (2012), a possibility that we elaborate on in our notes of caution below. In line with a potential context dependence of a red-failure association, in the current generic context we did not find any association between red and failure in either population. Neither failure

nor its synonyms (e.g., defeat, unsuccessful) were listed in Experiment 1a. Although previous studies showed evidence indicating poorer test performance when briefly shown the colour red (e.g., Elliot et al., 2007; Lichtenfeld, Maier, Elliot, & Pekrun, 2009; Maier, Elliot, & Lichtenfeld, 2008; Shi et al., 2015), null effects have been reported elsewhere in the literature (e.g., Kuller, Mikellides, & Janssens, 2009; Larsson & von Stumm, 2015). Other recent literature showed opposite effects (enhanced performance) dependent on profession (Zhang & Han, 2014), type of tasks (Mehta & Zhu, 2009), and task difficulty level (Xia et al., 2016). Accordingly, red enhanced performance in a simple proof-reading task as opposed to a difficult proof-reading task or creative tasks (Xia et al., 2016). The inconsistencies regarding the effect of the colour red on cognitive performance may well in part be due to subtle contextual differences that change the salient colour associations as emphasised by the absence of a red failure association in the present generic context.

Finally, it is widely accepted that the colour red has positive connotations in the Chinese culture as it represents celebrations, joyful occasions (He, 2011) and fortune (Huang, 2011). Therefore, it is no surprise the concepts *enthusiastic* and *excited* were listed in our study for Chinese bilingual and monolingual participants. However, the concept *prosperity*, which is commonly associated with red in the Chinese culture, was not evident in the list of concepts elicited in Experiment 1a. Even its synonyms (*wealth, success, fortune*) were not found in our list. As the participants in our study were young adults, it is plausible that the lack of association with prosperity and fortune meant that these associations are not as salient to young student participants as compared to those that are older and have their own wealth.

Based on the current findings and previous studies, we can predict that, i) the effect of red on aggression and dominance is stronger than its effect on cognitive performance in English and Chinese speaking adults; ii) that the effect of red on attraction and heightened attention should be more enhanced in English than Chinese speaking adults; and iii) that there

may be an effect of red on feeling joy, which is more enhanced in Chinese than English speaking adults.

Blue. In contrast to red, studies have yielded evidence suggesting that viewing blue enhances creative performance (Mehta & Zhu, 2009) and more recently, is beneficial for detail-oriented tasks that are difficult (Xia et al., 2016). It has been argued that red and blue have different associations within the cognitive domain such that they induce alternative motivations (Mehta & Zhu, 2009). Accordingly, because red is often associated with danger, it should activate an avoidance motivation in achievement contexts. On the other hand, it is argued that blue is associated with openness, peace, calm, sad, and concepts related to ocean and sky (Kaya & Epps, 2004; Mehta & Zhu, 2009; Palmer et al., 2013; Wexner, 1954) and therefore should activate an approach motivation thus making people behave in a more explorative, risky manner (Friedman & Forster, 2010 for a review). Across both experiments, these passive concepts under the themes water/sky related and relaxed were also consistently associated with the colour category blue in our study by English (e.g., *cool, wet, relaxed,* etc.) Chinese bilingual (e.g., *fly, broad, deep,* etc.), and Chinese monolingual (e.g., *fly, broad, swim, peaceful, etc.*) participants. We found a specific blue tile (28E) that was associated with passivity across all cultures. However, there seem to be cross-cultural differences in the range of lightness selected. While Chinese bilingual participants choose tiles ranging from medium lightness to dark and Chinese monolingual participants choose tiles ranging from light to medium lightness, English participants chose tiles ranging from light to medium and dark to represent concepts associated with passivity. Thus, although passive concepts are consistently associated with the colour blue (28E) across culture and experiments, the present finding suggests that passivity in the English culture encompasses a larger lightness range than in the Chinese culture.

Green. Recently, green has also been found to facilitate creative performance. Lichtenfeld and colleagues (2012) suggested that green is associated with growth (physical and psychological) and may serve as a cue that evokes a mastery-approach in a creative context. In the current study the colour green was associated with concepts such as *fresh* and *health,* but there was no direct association related to growth. As was the case for red and failure associations between green and growth may be context specific. The association observed with health supports previous research on green being associated with increased perceived health (Schuldt, 2013) and positive exercise outcomes (Akers et al., 2012). Conventional associations with green were not as strong as expected. For example, green is conventionally conceived to signal safety. However, in a study on red versus green (Pravossoudovitch et al., 2014), there was strong evidence of an association between red and danger, but weak evidence of an association between green and safety. In support of Pravossoudovitch and colleague's findings, the green-safety association was not found in our study or in previous work (Chan & Courtney, 2001). In addition, we did not find any specific (green) tile that was strongly associated with any concepts in comparison to red and blue. Previous research has shown that green may also at times, serve as an implicit cue that exerts a subtle influence on affect, cognition, and behaviour (Akers et al., 2012; Lichtenfeld et al., 2012; Schuldt, 2013). However, we suspect that the power, breadth, and applicability of green's meaning and influence will be far less than that of red and blue.

Notes of caution

Whilst we are the first to carry out a cross-cultural conceptual colour association study using both open-ended and closed-ended tasks and a full array of actual colour stimuli specified in a colorimetric space, we are also cautious about the interpretation of our findings.

First, whilst we showed that given a general context, people make consistent associations to colour categories, to specific colour tiles and to specific lightness ranges that support some of the colour effects within the literature, there are a number of concepts featured in other relevant literature that were not generated by our participants. For example, the concepts failure and prosperity were not given in response to the colour category red, and the concepts growth and safe did not arise in response to the colour category green. Although these concepts were not listed, this does not necessarily mean that they are never associated with the respective colours. More this finding might imply that these associations are highly context sensitive in contrast to, say, the association between red and danger, which emerges even in the present general context. According to Elliot and Maiers' (2012) colour-in-context theory, the physical and psychological context in which a colour is perceived is thought to influence its meaning and, accordingly, responses to it. Consequently, the general context in which our participants were listing associations might not have encouraged those associations to become salient in comparison to a specific context such as 'taking an exam' or 'opening a new business'. Thus, presenting red in the context of taking an exam might evoke the concept 'failure' but not in the context of opening a new shop. The potential context sensitivity of some colour associations underlines the importance for future work to explore colour associations on a context by context basis using the same kind of systematic approach developed in the present work.

Second, our strict criteria to only include words that are strongly defined as one of the abstract concepts (emotion, action, quality of things, symbolic) meant that we could have missed out on certain words that had high consensus and specificity to a particular colour. For example, the same word listed by Chinese monolingual participants *lovely(cute)(可愛)* was not included in the final list of words in Chinese bilingual participants even when it met the consensus and specificity criterion under the colour category pink (consensus = 71%;

specificity = 86%). This was due to the fact that Chinese bilingual participants did not define the word *lovely(cute)(可愛)* as an abstract concept.

Third, our sample sizes for the studies, especially the English and bilingual Chinese samples are relatively small. This could result in some associations that are present in the population being missed in our sample and in associations that appear robust in our sample failing to generalise to the population.

Fourth, we should be careful not to overgeneralise effects to a whole culture. Jiang and colleagues (2015) found red and green conveying different meanings based on cultural environment such that mainland Chinese participants tended to associate red with up and green with down, whereas Hong Kong Chinese participants associated green with up and red with down suggesting that culturally specific environmental cues could influence behaviour. Similarly, using nine to 16 concepts (Bergum & Bergum, 1981; Courtney, 1986; Chan & Courtney, 2001) such as *safe, danger, caution, cold, hot, go, stop, on, off, potential hazard, radiation hazard, soft, hard, weak, strong,* and *normal*, it was found that Chinese participants from Yunnan and Hong Kong did not yield such clear-cut associations as those found with US participants. Our samples of Chinese participants residing in the UK and China consisted of undergraduate university students from different regions of China. In addition, while the Chinese bilingual participants knew Chinese and English, they also spoke in Cantonese and/or Mandarin whereas the Chinese monolinguals only spoke Mandarin. These differences in dialect and region may have impacted on our findings such that they are specific to particular dialects and regions. Future work should seek to compare across dialects and speakers of the same dialect from different regions to explore these possibilities.

Fifth, we acknowledge that the themes we generated were subjective and that others might interpret them differently as words can have more than one meaning (e.g., *running* and *walking* can be either a positive, neutral or negative word).

Finally, while we were able to identify if it is any colour category, lightness, or specific tile (with a specific hue, lightness, and saturation) that evokes a particular association, we acknowledge a limitation of using the WCS stimulus grid. We were not able to evoke any association based on saturation as the standardised Munsell colour chips in this array consist of chromatic colour chips at only high levels of saturation.

Conclusions

We have created a large dataset of conceptual associations in English and Chinese bilingual and monolingual speaking populations through first identifying concepts that were strongly associated with a colour category and then reversing the process to identify colours that were associated with those concepts. The integration of both methods using the WCS stimulus grid meant that we were able to identify concepts that have the strongest association with a specific colour category, exemplar and lightness range in order to accurately predict the effects of colour on cognition and behaviour. It is clear that colour associations vary with culture and this should be borne in mind when trying to generalise findings of the current paper to other groups of people. Our findings suggest that using the right colours and lightness in a culturally nuanced way in contexts such as product marketing, education, web development, information visualisation (e.g. food labelling, medical information, road signs), urban design and the visual arts may not only improve semantic coherence, but may also better inspire the desired psychological response.

Footnotes

1. Munsell is a standardised colour metric system based on psychophysical judgments (Newhall, Nickerson & Judd, 1943). Munsell colours vary on three dimensions: Hue (from red through yellow, green, blue to purple); Value or lightness scaled from zero (black) to 9 (white); and Chroma (saturation). Stimuli are specified using a notation that consists of three co-ordinates that each relate to one of the three dimensions. The standardised Munsell colour chips consist of 320 chromatic chips representing 40 equally spaced hues at eight levels of lightness (Munsell Value), each at a high level of saturation, and an additional 10 achromatic chips.

2. The Munsell Book of Colour (1976) is based on the Munsell colour system. The colour samples are painted colour tiles systematically arranged in terms of three coordinates: hue, value, and chroma; the tiles possess numerical values that represents uniform perceptual spacing for each coordinate (Romney, 2008).

3. Prior to determining the best solution for clustering, we used the "average silhouette approach" in an earlier analysis to determine how well (or the stability of quality of cluster result) each object (word) lay within its cluster. A higher average silhouette value indicates a better clustering. The optimal number of clusters k is the one that maximises the average silhouette over a range of possible values of k (Kaufman & Rousseeuw, 1990). According to this criterion, the optimal number of clusters for each experiment were: for the English participants E1 = 12, E2 = 10; Chinese bilingual participants, E1 = 8, E2 = 11; Chinese monolingual participants, E1 = 8, E2 = 9. In order to facilitate direct comparison of clusters between experiments we therefore opted for a best 'average' solution across cluster analyses in both

experiments of 11, 10 and 9 cluster solutions for the English, Chinese bilingual and

Chinese monolingual samples respectively.

Author note

A preliminary account of the English and Chinese bilingual data was presented at the European Conference on Visual Perception (Tham, D., Sowden, P. T., Grandison, A., Franklin, A., Lee, A., & Ng, M. (2015). A systematic investigation of colour and concept associations. European Conference on Visual Perception, Glasgow, Scotland. Abstract published in *Perception*, *44*, 149).

The full set of raw data for the colour association data reported in the manuscript are freely available at the Open Science Framework https://osf.io/.

Context of the Research

At the time that this work was conceived the body of research being published in experimental, cognitive, social and educational psychology, reporting that colour could influence a range of aspects of cognition, affect and behaviour, was growing significantly. The explanations of these effects were often couched in terms of symbolic associations with colour priming aspects of cognition that then shaped behaviour. But explanations were typically post hoc. We (Sowden, Grandison and Franklin) were conducting similar work at the time and became aware that, like others, we lacked a systematic analysis of symbolic (as opposed to object) word-colour associations to draw upon in order to be able to make strong *a priori* predictions about how exposure to colour would influence cognition and behaviour. Thus, we embarked on the present work to provide this foundation for future work in the field. As most work was being conducted with English language speakers, we chose to map English symbolic associations with colour. But we also felt that stronger work would be possible if authors approached research into the effects of colour on behaviour cross-culturally because different symbolic word-colour associations across languages would lead to different predicted effects on behaviour, as a function of those different associations. We selected Chinese as it is the world's other most spoken language and one that preliminary work suggested would have a rich range of contrasting colour associations compared to English and hence the programme of work was born.

Acknowledgements

References

Adams, F. M., & Osgood, C. E. (1973). A cross-cultural study of the affective meanings of color. *Journal of Cross-Cultural Psychology, 4*(2), 135-156. http://dx.doi.org/10.1177/002202217300400201

Akers, A., Barton, J. L., Cossey, R., Gainsford, P., Griffin, M., & Micklewright, D. (2012). Visual Colour Perception in Green Exercise: Positive Effects on Mood and Perceived Exertion. *Environmental Science and Technology*, *46*(16), 1–6. http://doi.org/10.1021/Es301685g.

Barchard, K. A., Grob, K. E., & Roe, M. J. (2016). Is sadness blue? The problem of using figurative language for emotions on psychological tests. *Behavior Research Methods, 49,* 443–456.

Bagchi, R., & Cheema, A. (2012). The Effect of Red Background Color on Willingness-to-Pay: The Moderating Role of Selling Mechanism. *Journal of Consumer Research*, *39*, 947–960.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*, 323–370.

Bergum, B.O., Bergum, J.E., 1981. Population stereotypes: an attempt to measure and define. In: *Proceedings of the Human Factors Society 25th Annual Meeting*, Santa Monica, CA: Human Factors Society, pp. 662–665.

Berlin, B., & Kay, P. (1969, 1991). *Basic color terms: Their universality and evolution*. Berkeley, CA: University of California Press.

Bonnardel, V. Dubey, N., Beniwal, S., Mayukhini P., and Davies, C. (2014). Colour

    Associations in a young adult Indian Population. 1-4 April, *New Perspectives on*

    Colour –AISB-50, Goldsmiths, London.

Boynton, R. M., & Olson, C. X. (1990). Salience of chromatic basic color terms confirmed

    by three measures. *Vision Research, 30*, 1311–1317.

Brooker, A. & Franklin, A. (2016). The effect of colour on children's cognitive performance.

    British Journal of Educational Psychology, 86, 241-255.

Chan, A. H., & Courtney, A. J. (2001). Color associations for Hong Kong Chinese.

    International Journal of Industrial Ergonomics, *28*(3–4), 165–170.

    http://doi.org/10.1016/S0169-8141(01)00029-4.

Cook, R. S., Kay, P. & Regier, T. (2005). The world color survey database: History and use.

    In Cohen, H. & Lefebvre, C. (eds.), *Handbook of Categorisation in the Cognitive*

    Sciences. Amsterdam: Elsevier. Durbin.

Courtney, A.J., 1986. Chinese population stereotypes: Color associations. *Human Factors, 28*

    (1), 97–99.

Crozier, W.R. (1996). The psychology of colour preference. *Review of Progress in*

    Coloration and Related Topics, 26, 63-72.

Dael, N., Perseguers, M., Marchand, C., Mohr, C., Dael, N., Perseguers, M., … Antonietti, J.

    (2016). Put on that colour, it fits your emotion : Colour appropriateness as a function

    of expressed emotion. *The Quarterly Journal of Experimental Psychology, 69,* 1619-

    1630. http://doi.org/10.1080/17470218.2015.1090462.

Davies, I, & Corbett, G. (1997). A cross-cultural study of colour grouping: Evidence for weak linguistic relativity. *British Journal of Psychology, 88,* 493-517.

Del Giudice, M. (2017). Pink, blue, and gender: An update. *Archives of sexual behavior*, *46*(6), 1555-1563.

Díaz-Román, A., Megías, A., Díaz-Piedra, C., Catena, A., & Di Stasi, L. L. (2015). Seeing red on the road. *Perception*, *44*(44), 87–92. http://doi.org/10.1068/p7811.

Elliot, A. J. (2015). Color and psychological functioning: a review of theoretical and empirical work. *Frontiers in Psychology, 6,* 1-8.

Elliot, A. J., & Aarts, H. (2011). Perception of the color red enhances the force and velocity ofmotor output. *Emotion 11*, 445–49.

Elliot, A.J., Fairchild, M.D., & Franklin, A. (2015). *Handbook of Color Psychology.* Cambridge: University Press.

Elliot, A. J., & Maier, M. A. (2012). Color-in-context theory. *Advances in Experimental Social Psychology, 45*, 61–125. doi: 10.1016/B978-0-12-394286-9. 00002-0.

Elliot, A. J., & Maier, M. A. (2014). Color psychology: effects of perceiving color on psychological functioning in humans. *Annual Review of Psychology, 65*, 95–120. doi: 10.1146/annurev-psych-010213-115035.

Elliot, A. J., & Niesta, D. (2008). Romantic red: red enhances men's attraction to women. *Journal of Personality and Social Psychology, 95*, 1150–1164. doi: 10.1037/0022-3514.95.5.1150.

Elliot, A. J., Maier, M. A., Moller, A. C., Friedman, R., & Meinhardt, J. (2007). Color and psychological functioning: the effect of red on performance attainment. *Journal of*

Experimental Psychology: General, 136, 154–168. doi: 10.1037/0096-3445.136. 1.154.

Fairchild, M.D. (2005). *Color Appearance Models (2nd Edition)*. New York:Wiley.

Fetterman, A. K., Liu, T., & Robinson, M. D. (2015). Extending Color Psychology to the Personality Realm: Interpersonal Hostility Varies by Red Preferences and Perceptual Biases. *Journal of Personality*, *81*, 106–116. http://doi.org/10.1111/jopy.12087.

Fletcher, R. (1981). *City Colour Vision Test*. Windsor: Keeler.

Frank, M. G., & Gilovich, T. (1988). The dark side of self and social perception: black uniforms and aggression in professional sports. *Journal of Personality and Social Psychology, 54*, 74–85. doi: 10.1037/0022-3514.54.1.74.

Friedman, R. S., & Förster, J. (2010). Implicit affective cues and attentional tuning: an integrative review. *Psychological Bulletin*, *136*(5), 875–93. http://doi.org/10.1037/a0020495.

Gao, J., & Sutrop, U. (2014). The basic color terms of Mandarin Chinese. *Studies in Language, 38,* 335-359.

Grandison, A., Davies, I.R.L., & Sowden, P.T. (2014). The Evolution of GRUE: Evidence for a new colour term in the language of the Himba. In W. Anderson, C. P. Biggam, C. Hough and C. Kay. *Colour Studies: A broad spectrum*. Amsterdam, NL: Benjamins, pp. 53-66.

Hanafy, I. M., & Reham, S. (2016). Humanites & Social Sciences A Cross-cultural study of emotional Responses on Colours. *Global Journal on Humanities & Social Sciences*, *3*(3), 53–60.

Hardin, C.L., & Maffi, L. (1997). *Color categories in thought and language.* Cambridge: University Press.

He, G. (2011). A Comparative Study of Color Metaphors in English and Chinese. *Theory and Practice in Language Studies*, *1*(12), 1804–1808. http://doi.org/10.4304/tpls.1.12.1804-1808.

Hemming, J. (2012). Red, White, and Black in Symbolic Thought: The Tricolour Folk Motif, Colour Naming, and Trichromatic Vision. *Folklore*, *123*(3), 310–329. http://doi.org/10.1080/0015587X.2012.716599.

Hemphill, M. (1996). A note on adults' color-emotion association. *The Journal of Genetic Psychology*, *157*(3), 275–280.

Hill, R.A., & Barton, R.A. 2005. Red enhances human performance in contests. *Nature 435*, 293.

Huang, Q. (2011). A Study on the Metaphor of " Red " in Chinese Culture. *American International Journal of Contemporary Research*, *1*(3), 99–102.

Hupka, R. B., Zaleski, Z., Otto, J., Reidl, L. & Tarabrina, N. V. (1997) The colours of anger, envy, fear, and jealousy: a cross-cultural study. *Journal of cross-cultural psychology 28*(2), 156–71.

Hurlbert, A.C., Owen, K.A. (2015). Biological, cultural, and developmental influences on color preference. In A.J. Elliott, M.D. Fairchild, & A. Franklin (Eds.), *Handbook of Color Psychology* (pp. 454-480). Cambridge: Cambridge University Press.

Hutchings, J. (1997). Folklore and Symbolism of Green. *Folklore and Symbolism of Green*, *108*, 55–63.

Hutchings, J. (2004). Colour in folklore and tradition? The principles. *Color Research & Application*, *29*(1), 57–66. http://doi.org/10.1002/col.10212

Ishihara, S. (2009) *Ishihara test for colour-blindness*. Tokyo: Kanehara (Original work published 1917).

Jacobs, L., Keown, C., & Worthley, R. (1991). Cross-cultural colour comparisons : Global marketers beware! *International Marketing Review*, *8*(3), 21–30.

Jiang, F., Lu, S., Yao, X., Yue, X., & Au, W.T. (2014). Up or down? How culture and color affect judgments. *Journal of Behavioral Decision Making, 27,* 226-234.

Kay, P., Berlin, B., Maffi, L., Merrifield, W.R., & Cook, R.S. (2009). *The World Color Survey*. Stanford, CA: CSLI Press.

Kaya, N., & Epps, H. H. (2004). Color-emotion associations : Past experience and personal preference. In *AIC 2004 Color and Paints, Interim Meeting of the International Color Association, Proceedings* (pp. 31–34).

Krenn, B. (2014). The impact of uniform color on judging tackles in association football. *Psychology of Sport and Exercise, 15*, 222–225. doi: 10.1016/j.psychsport.2013.11.007.

Kuhbandner, C. & Pekrun, R. (2013). Joint effects of emotion and color on memory. *Emotion, 13,* 375-379.

Kuller, R., Mikellides, B., & Janssens, J. (2009). Color , Arousal , and Performance — A Comparison of Three Experiments Rikard Ku. *Color Research & Application*, *34*(2), 141–152. http://doi.org/10.1002/col.20476.

Kumi R, Conway, C.M., Limayem, M., & Goyal, S. (2013). Learning in color: how color and affect influence learning outcomes. *IEEE Transactions on Professional Communication, 56*, 2–15.

Langguth, B., Eichhammer, P., Pickert, K., Frank, U., Perna, M., Landgrebe, M., … Sand, P. (2009). Stable motor cortex excitability in red and green lighting conditions. *Neuroscience Letters*, *460*(1), 32–35. http://doi.org/10.1016/j.neulet.2009.05.034

Larsson, E., & Stumm, S. V. (2015). Seeing red? The effect of colour on intelligence test performance. *Intelligence, 48,* 133-136.

Lenneberg, E. H., & Roberts, J. M. (1956). *The Language of Experience: A Study in Methodology*. Waverly Press.

Lichtenfeld, S., Elliot, A. J., Maier, M. a, & Pekrun, R. (2012). Fertile green: green facilitates creative performance. *Personality & Social Psychology Bulletin*, *38*(6), 784–97. http://doi.org/10.1177/0146167212436611.

Lichtenfeld, S., Maier, M.A., Elliot, A.J., & Pekrun, R. (2009). The semantic red effect: processing the word red undermines intellectual performance. *Journal of Experimental Social Psychology, 45*, 1273–1276.

Lindsey, D. T., & Brown, A. M. (2014). The color lexicon of American English. *Journal of Vision*, *14*, 1–25. http://doi.org/10.1167/14.2.17.doi.

Ludwig, V. U., & Simner, J. (2013). What colour does that feel? Tactile--visual mapping and the development of cross-modality. *Cortex*, *49*(4), 1089–99. http://doi.org/10.1016/j.cortex.2012.04.004.

Lynn, M., Giebelhausen, M., Garcia, S., Li, Y., & Patumanon, I. (2016). Clothing color and tipping: an attempted replication and extension. *Journal Hospitality & Tourism Research, 40*, 516-524. doi: 10.1177/1096348013504001.

Maier, M. A., Elliot, A. J., & Lichtenfeld, S. (2008). Mediation of the negative effect of red on intellectual performance. *Personality and Social Psychology Bulletin, 34*, 1530–1540.

Mehta, R., & Zhu, R. (2009). Blue or red? Exploring the effect of color on cognitive task performances. *Science 323*, 1226–1229. doi: 10.1126/science. 1169144.

Meier, B. P., Robinson, M. D., & Clore, G. L. (2004). Why good guys wear white— Automatic inferences about stimulus valence based on brightness. *Psychological Science, 15*, 82–87.

Mohammad, S. M. (2011). Even the Abstract have Colour : Consensus in Word – Colour Associations. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies* (pp. 368–373). Portland, OR.

Moller, A. C., Elliot, A. J., & Maier, M. a. (2009). Basic hue-meaning associations. *Emotion*, *9*(6), 898–902. http://doi.org/10.1037/a0017811.

Munsell Color Company, Inc (1976). *Munsell book of color: Matte finish collection*. Baltimore, MD: Munsell Color.

Newhall, S.M., Nickerson, D., & Judd, D.B. (1943). Final report of the O.S.A. subcommittee on the spacing of the Munsell colours. *Journal of the Optical Society of America, 33,* 385-418.

O'Connor, Z. (2010). Colour Harmony Revisited. *Colour Research and Application, 35*(4), 267–273. http://doi.org/10.1002/col.20578

Palmer, S. E., & Schloss, K. B. (2010). An ecological valence theory of human color preference. *Proceedings of the National Academy of Sciences, 107*, 8877–8882. doi:10.1073/pnas.0906172107.

Palmer, S. E., Schloss, K. B., Xu, Z., & Prado-León, L. R. (2013). Music-color associations are mediated by emotion. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(22), 8836–41. http://doi.org/10.1073/pnas.1212562110

Peperkoorn, L.S., Roberts, S.C., & Pollet, T.V. (2016). Revisiting the red effect on attractiveness and sexual receptivity: No effect of the color red on human mate preferences. *Evolutionary Psychology, 14,* 1-13.

Prado-León, L. R., Avila-Chaurand, R., & Rosales-Cinco, R. A. (2006). Color associations in the Mexican university population. In J. Pitchford & C.Biggam (Eds.), *Progress in Colour Studies Volume 2: Psychological aspects* (pp. 57-71). Amsterdam & Philadelphia: John Benjamins.

Pravossoudovitch, K., Cury, F., Young, S. G., & Elliot, A. J. (2014). Is red the colour of danger? Testing an implicit red-danger association. *Ergonomics*, *57*(4), 503–10. http://doi.org/10.1080/00140139.2014.889220.

Regier, T., & Kay, P. (2009). Language, thought, and color: Whorf was half right. *Trends in Cognitive Sciences, 13,* 439-446.

Romney, A. K. (2008). Relating reflectance spectra space to Munsell color appearance space. *Journal of the Optical Society of America*, *25*, 658–666.

Schuldt, J. P. (2013). Does green mean healthy? Nutrition label color affects perceptions of

   healthfulness. *Health Communication*, *28*(May 2015), 814–21.

   http://doi.org/10.1080/10410236.2012.725270.

Shi, J., Zhang, C., & Jiang, F. (2015). Does red undermine individuals' intellectual

   performance? A test in China. *International Journal of Psychology*, *50*(1), 81–84.

   http://doi.org/10.1002/ijop.12076.

Soldat, A. S., Sinclair, R. C., & Mark, M. M. (1997). Color as an environmental processing

   cue: external affective cues can directly affect processing strategy without affecting

   mood. *Social Cognition. 15*, 55–71. doi: 10.1521/soco.1997.15.1.55.

Stabler, J. R., & Johnson, E. E. (1972). The meaning of black and white to children.

   International Journal of Symbology, 3, 11–21.

Strauss, E. D., Schloss, K. B., & Palmer, S. E. (2013). Color preferences change after

   experience with liked/disliked colored objects. *Psychonomic Bulletin & Review, 20*,

   935–943. doi:10.3758/ s13423-013-0423-2.

Sturges, J., & Whitfield, T. W. A. (1995). Locating basic colours in the Munsell space. *Color

   Research and Application, 20*, 364–376.

Sutton, T. M., & Altarriba, J. (2016). Color associations to emotion and emotion-laden

   words: A collection of norms for stimulus construction and selection. *Behavior

   research methods*, *48*(2), 686-728.

Taylor, C., & Franklin, A. (2012). The relationship between color–object associations and

   color preference: Further investigation of ecological valence theory. *Psychonomic

   Bulletin & Review, 19,* 190–197. doi:10.3758/s13423-012-0222-1.

Taylor, C., Clifford, A., & Franklin, A. (2013). Color preferences are not universal. *Journal of Experimental Psychology: General*, *142*, 1015–1027.

Thorstenson, C.A., Elliot, A.J., Pazda, A.D., Perrett, D.I. & Xiao, D. (2019). Color Associations in the Context of the Face. *Emotion, 18*(7), 1032–1042.

Valdez, P., & Mehrabian, A. (1994). Effects of color on emotions. *Journal of Experimental Psychology. General*, *123*(4), 394–409. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7996122.

Wexner, L. B. (1954) 'The degree to which colours (hues) are associated with mood-tones', *Journal of applied psychology 38*: 432–5.

Wright, B., & Rainwater, L. (1962). The meanings of color. *The Journal of General Psychology, 67*, 89–99. doi:10.1080/00221309.1962.9711531.

Xia, T., Song, L., Wang, T. T., Tan, L., & Mo, L. (2016). Exploring the Effect of Red and Blue on Cognitive Task Performances. *Frontiers in Psychology*, *7*, 1–7. http://doi.org/10.3389/fpsyg.2016.00784.

Yamazaki, A. K. (2010). An analysis of background-color effects on the scores of a computer-based English test. *Lecture Notes in Computer Science, 6277,* 630–636. doi: 10.1007/978- 3-642-15390-7_65.

Zhang, T., & Han, B. (2014). Experience reverses the red effect among Chinese stockbrokers. *PloS One*, *9*(2), e89193. http://doi.org/10.1371/journal.pone.0089193.