Even the Abstract have Colour:
Consensus in Word–Colour Associations

Saif M. Mohammad
Institute for Information Technology
National Research Council Canada.
Ottawa, Ontario, Canada, K1A 0R6
`saif.mohammad@nrc-cnrc.gc.ca`

Abstract

Colour is a key component in the success-ful dissemination of information. Since many real-world concepts are associated with colour, for example *danger* with red, linguistic information is often complemented with the use of appropriate colours in information vi-sualization and product marketing. Yet, there is no comprehensive resource that captures concept–colour associations. We present a method to create a large word–colour asso-ciation lexicon by crowdsourcing. A word-choice question was used to obtain sense-level annotations and to ensure data quality. We fo-cus especially on abstract concepts and emo-tions to show that even they tend to have strong colour associations. Thus, using the right colours can not only improve semantic coherence, but also inspire the desired emo-tional response.

1 Introduction

Colour is a vital component in the successful deliv-ery of information, whether it is in marketing a com-mercial product (Sable and Akcay, 2010), in web design (Meier, 1988; Pribadi et al., 1990), or in in-formation visualization (Christ, 1975; Card et al., 1999). Since real-world concepts have associations with certain colour categories (for example, *danger* with red, and *softness* with pink), complementing linguistic and non-linguistic information with appro-priate colours has a number of benefits, including: (1) strengthening the message (improving semantic coherence), (2) easing cognitive load on the receiver, (3) conveying the message quickly, and (4) evoking the desired emotional response. Consider, for exam-ple, the use of red in stop signs. Drivers are able to recognize the sign faster, and it evokes a subliminal emotion pertaining to possible danger, which is en-tirely appropriate in the context. The use of red to show areas of high crime rate in a visualization is another example of good use of colour to draw emo-tional response. On the other hand, improper use of colour can be more detrimental to understanding than using no colour (Marcus, 1982; Meier, 1988).

A word has strong association with a colour when the colour is a salient feature of the concept the word refers to, or because the word is related to a such a concept. Many concept–colour associa-tions, such as *swan* with white and *vegetables* with green, involve physical entities. However, even ab-stract notions and emotions may have colour as-sociations (*honesty*–white, *danger*–red, *joy*–yellow, *anger*–red). Further, many associations are culture-specific (Gage, 1969; Chen, 2005). For example, *prosperity* is associated with red in much of Asia.

Unfortunately, there exists no lexicon with any significant coverage that captures these concept–colour associations, and a number of questions re-main unanswered, such as, the extent to which hu-mans agree with each other on these associations, and whether physical concepts are more likely to have a colour association than abstract ones.

In this paper, we describe how we created a large word–colour lexicon by crowdsourcing with effec-tive quality control measures (Section 3), as well as experiments and analyses to show that:

- More than 30% of the terms have a strong colour association (Sections 4).

368

- About 33% of thesaurus categories have strong colour associations (Section 5).
- Abstract terms have colour associations almost as often as physical entities do (Section 6).
- There is a strong association between different emotions and colours (Section 7).

Thus, using the right colours can not only improve semantic coherence, but also inspire the desired emotional response.

2 Related Work

The relation between language and cognition has received considerable attention over the years, mainly on answering whether language impacts thought, and if so, to what extent. Experiments with colour categories have been used both to show that language has an effect on thought (Brown and Lenneberg, 1954; Ratner, 1989) and that it does not (Bornstein, 1985). However, that line of work does not explicitly deal with word–colour associations. In fact, we did not find any other academic work that gathered large word–colour associations. There is, however, a commercial endeavor—Cymbolism[1].

Child et al. (1968), Ou et al. (2011), and others show that people of different ages and genders have different colour preferences. (See also the online study by Joe Hallock[2].) In this work, we are interested in identifying words that have a strong association with a colour due to their meaning; associations that are not affected by age and gender preferences.

There is substantial work on inferring the emotions evoked by colour (Luscher, 1969; Kaya, 2004). Strapparava and Ozbal (2010) compute corpus-based semantic similarity between emotions and colours. We combine a word–colour and a word–emotion lexicon to determine the association between emotion words and colours.

Berlin and Kay (1969), and later Kay and Maffi (1999), showed that often colour terms appeared in languages in certain groups. If a language has only two colour terms, then they are white and black. If a language has three colour terms, then they tend to be white, black, and red. Such groupings are seen for up to eleven colours, and based on these groupings, colours can be ranked as follows:

$$\begin{aligned}&\text{1. white, 2. black, 3. red, 4. green, 5. yel-}\\&\text{low, 6. blue, 7. brown, 8. pink, 9. purple,}\\&\text{10. orange, 11. grey}\end{aligned} \quad (1)$$

There are hundreds of different words for colours.[3] To make our task feasible, we chose to use the eleven basic colour words of Berlin and Kay (1969).

The MRC Psycholinguistic Database (Coltheart, 1981) has, among other information, the *imageability ratings* for 9240 words.[4] The imageability rating is a score given by human judges that reflects how easy it is to visualize the concept. It is a scale from 100 (very hard to visualize) to 700 (very easy to visualize). We use the ratings in our experiments to determine whether there is a correlation between imageability and strength of colour association.

3 Crowdsourcing

We used the *Macquarie Thesaurus* (Bernard, 1986) as the source for terms to be annotated by people on Mechanical Turk.[5] Thesauri, such as the *Roget's* and *Macquarie*, group related words into categories. These categories can be thought of as coarse senses (Yarowsky, 1992; Mohammad and Hirst, 2006). If a word is ambiguous, then it is listed in more than one category. Since we were additionally interested in determining colour signatures for emotions (Section 7), we chose to annotate all of the 10,170 word–sense pairs that Mohammad and Turney (2010) used to create their word–emotion lexicon. Below is an example questionnaire:

Q1. Which word is closest in meaning to *sleep*?
- *car*
- *tree*
- *nap*
- *olive*

Q2. What colour is associated with *sleep*?
- black
- green
- purple
- white
- blue
- grey
- pink
- yellow
- brown
- orange
- red

Q1 is a word choice question generated automatically by taking a near-synonym from the thesaurus and random distractors. If an annotator answers this question incorrectly, then we discard information from both Q1 and Q2. The near-synonym also guides the annotator to the desired sense of the word. Further, it encourages the annotator to think clearly

---

[1] http://www.cymbolism.com/about
[2] http://www.joehallock.com/edu/COM498/preferences.html

[3] See http://en.wikipedia.org/wiki/List_of_colors
[4] http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm
[5] Mechanical Turk: www.mturk.com

	white	black	red	green	yellow	blue	brown	pink	purple	orange	grey
overall	11.9	12.2	11.7	12.0	11.0	9.4	9.6	8.6	4.2	4.2	4.6
voted	22.7	18.4	13.4	12.1	10.0	6.4	6.3	5.3	2.1	1.5	1.3

Table 1: Percentage of terms marked as being associated with each colour.

about the target word's meaning; we believe this improves the quality of the annotations in Q2.

The colour options in Q2 were presented in random order. We do not provide a "not associated with any colour" option to encourage colour selection even if the association is weak. If there is no association between a word and a colour, then we expect low agreement for that term. We requested annotations from five different people for each term.

The annotators on Mechanical Turk, by design, are anonymous. However, we requested annotations from US residents only.

4 Word–Colour Association

About 10% of the annotations had an incorrect answer to Q1. Since, for these instances, the annotator did not know the meaning of the target word, we discarded the corresponding colour association response. Terms with less than three valid annotations were discarded from further analysis. Each of the remaining terms has, on average, 4.45 distinct annotations. The information from multiple annotators was combined by taking the majority vote, resulting in a lexicon with 8,813 entries. Each entry contains a unique word–synonym pair, majority voted colour(s), and a confidence score—number of votes for the colour / number of total votes. (For the analyses in Sections 5, 6, and 7, ties were broken by picking one colour at random.) A separate version of the lexicon that includes entries for all of the valid annotations by each of the annotators is also available.[6]

The first row in Table 1 shows the percentage of times different colours were associated with the target term. The second row shows percentages after taking a majority vote of the annotators. Even though the colour options were presented in random order, the order of the most frequently associated colours is identical to the Berlin and Kay order (Section 2:(1)).

The number of ambiguous words annotated was 2924. 1654 (57%) of these words had senses that

[6]Please contact the author to obtain a copy of the lexicon.

target	sense	colour
bunk	nonsense	grey
bunk	furniture	brown
compatriot	nation	red
compatriot	partner	white
frustrated	hindrance	red
frustrated	disenchantment	black
glimmer	idea	white
glimmer	light	yellow
stimulate	allure	red
stimulate	encouragement	green

Table 2: Example target words that have senses associated with different colours.

majority class size						
one	two	three	four	five	$\geq$ two	$\geq$ three
15.1	52.9	22.4	7.3	2.1	84.9	32.0

Table 3: Percentage of terms in different majority classes.

were associated with at least two different colours. Table 4 gives a few examples.

Table 4 shows how often the majority class in colour associations is 1, 2, 3, 4, and 5, respectively. If we assume independence, then the chance that none of the 5 annotators agrees with each other (majority class size of 1) is $1 \times 10/11 \times 9/11 \times 8/11 \times 7/11 = 0.344$. Thus, if there was no correlation among any of the terms and colours, then 34.4% of the time none of the annotators would have agreed with each other. However, this happens only 15.1% of the time. A large number of terms have a majority class size ≥ 2 (84.9%), and thus have more than chance association with a colour. One can argue that terms with a majority class size ≥ 3 (32%) have *strong* colour associations.

Below are some reasons why agreement values are much lower than certain other tasks, for example, part of speech tagging:

- The annotators were not given a "not associated with any colour" option. Low agreement for certain instances is an indicator that these words have weak, if any, colour association. Therefore, inter-annotator agreement does not correlate with quality of annotation.
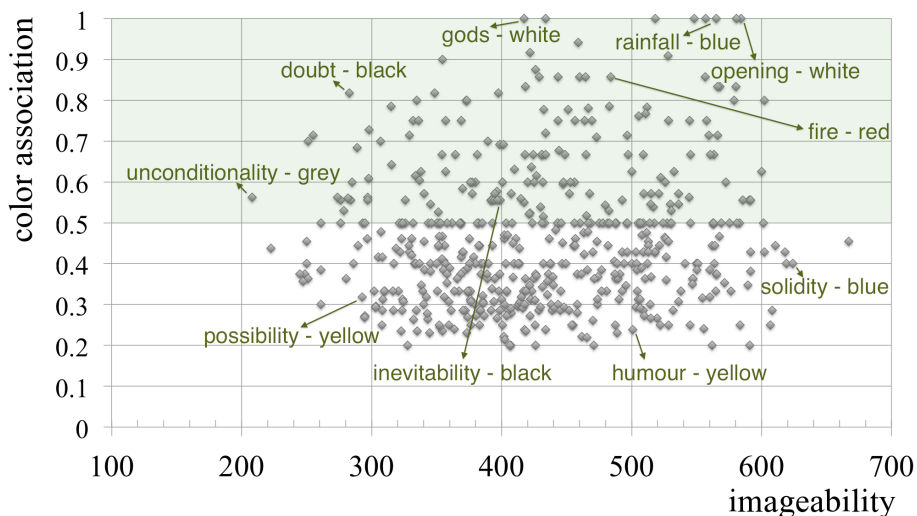
370

Figure 1: Scatter plot of thesaurus categories. The area of high colour association is shaded. Some points are labeled.

- Words are associated with colours to different degrees. Some words may be associated with more than one colour by comparable degrees, and there might be higher disagreement.
- The target word–sense pair is presented out of context. We expect higher agreement if we provided words in context, but words can occur in innumerable contexts, and annotating too many instances of the same word is costly.

Nonetheless, the lexicon is useful for downstream applications because any of the following strategies may be employed: (1) choosing colour associations from only those instances with high agreement, (2) assuming low-agreement terms have no colour association, (3) determining colour association of a category through information from many words, as described in the next section.

5 Category–Colour Association

Different words within a thesaurus category may not be strongly associated with any colour, or they may be associated with many different colours. We now determine whether there exist categories where the semantic coherence carries over to a strong common association with one colour.

We determine the strength of colour association of a category by first determining the colour c most associated with the terms in it, and then calculating the ratio of the number of times a word from the category is associated with c to the number of words in the category associated with any colour. Only cate-

gories that had at least four words that also appear in the word–colour lexicon were considered; 535 of the 812 categories from *Macquarie Thesaurus* met this condition. If a category has exactly four words that appear in the colour lexicon, and if all four words are associated with different colours, then the category has the lowest possible strength of colour association—0.25 (1/4). 19 categories had a score of 0.25. No category had a score less than 0.25. Any score above 0.25 shows more than random chance association with a colour. There were 516 such categories (96.5%). 177 categories (33.1%) had a score 0.5 or above, that is, half or more of the words in these categories are associated with one colour. We consider these to be strong associations.

6 Imageability

It is natural for physical entities of a certain colour to be associated with that colour. However, abstract concepts such as *danger* and *excitability* are also associated with colours—red and orange, respectively. Figure 1 displays an experiment to determine whether there is a correlation between imageability and association with colour.

We define imageability of a thesaurus category to be the average of the imageability ratings of words in it. We calculated imageability for the 535 categories described in the previous section using only the words that appear in the colour lexicon. Figure 1 shows the scatter plot of these categories on the imageability and strength of colour association axes. If

	white	black	red	green	yellow	blue	brown	pink	purple	orange	grey
anger words	2.1	**30.7**	**32.4**	5.0	5.0	2.4	6.6	0.5	2.3	2.5	9.9
anticipation words	**16.2**	7.5	11.5	**16.2**	10.7	9.5	5.7	5.9	3.1	4.9	8.4
disgust words	2.0	**33.7**	**24.9**	4.8	5.5	1.9	9.7	1.1	1.8	3.5	10.5
fear words	4.5	**31.8**	**25.0**	3.5	6.9	3.0	6.1	1.3	2.3	3.3	11.8
joy words	**21.8**	2.2	7.4	**14.1**	13.4	11.3	3.1	11.1	6.3	5.8	2.8
sadness words	3.0	**36.0**	**18.6**	3.4	5.4	5.8	7.1	0.5	1.4	2.1	16.1
surprise words	11.0	13.4	**21.0**	8.3	**13.5**	5.2	3.4	5.2	4.1	5.6	8.8
trust words	**22.0**	6.3	8.4	14.2	8.3	**14.4**	5.9	5.5	4.9	3.8	5.8

Table 4: Colour signature of emotive terms: percentage of terms associated with each colour. For example, 32.4% of the anger terms are associated with red. The two most associated colours are shown in bold.

	white	black	red	green	yellow	blue	brown	pink	purple	orange	grey
negative	2.9	**28.3**	**21.6**	4.7	6.9	4.1	9.4	1.2	2.5	3.8	14.1
positive	**20.1**	3.9	8.0	**15.5**	10.8	12.0	4.8	7.8	5.7	5.4	5.7

Table 5: Colour signature of positive and negative terms: percentage terms associated with each colour. For example, 28.3% of the negative terms are associated with black. The two most associated colours are shown in bold.

higher imageability correlated with greater tendency to have a colour association, then we would see most of the points along the diagonal moving up from left to right. Instead, we observe that the strongly associated categories are spread all across the imageability axis, implying that there is only weak, if any, correlation. Imageability and colour association have a Pearson's product moment correlation of 0.116, and a Spearman's rank order correlation of 0.102.

7 The Colour of Emotion Words

Emotions such as joy, sadness, and anger are abstract concepts dealing with one's psychological state. As pointed out in Section 2, there is prior work on emotions evoked by colours. In contrast, here we investigate the colours associated with emotion words. We combine the word–emotion association lexicon compiled by Mohammad and Turney (2010; 2011) and our word–colour lexicon to determine the colour signature of emotions—the rows in Table 4. Notably, we see that all of the emotions have strong associations with certain colours. Observe that anger is associated most with red. Other negative emotions—disgust, fear, sadness—go strongest with black. Among the positive emotions: anticipation is most frequently associated with white and green; joy with white, green, and yellow; and trust with white, blue, and green. Table 4 shows the colour signature for terms marked positive and negative (these include terms that may not be associated with the eight basic emotions). Observe that the neg-

ative terms are strongly associated with black and red, whereas the positive terms are strongly associated with white and green. Thus, colour can add to the potency of emotional concepts, yielding even more effective visualizations.

8 Conclusions and Future Work

We created a large word–colour association lexicon by crowdsourcing. A word-choice question was used to guide the annotator to the desired sense of the target word, and to ensure data quality. We observed that abstract concepts, emotions in particular, have strong colour associations. Thus, using the right colours in tasks such as information visualization, product marketing, and web development, can not only improve semantic coherence, but also inspire the desired psychological response. Interestingly, we found that frequencies of colour choice in associations follow the same order in which colour terms occur in language (Berlin and Kay, 1969). Future work includes developing automatic corpus-based methods to determine the strength of word–colour association, and the extent to which strong word–colour associations manifest themselves as more-than-random chance co-occurrence in text.

References

Brent Berlin and Paul Kay. 1969. *Basic Color Terms: Their Universality and Evolution.* Berkeley: University of California Press.

J.R.L. Bernard, editor. 1986. *The Macquarie Thesaurus.* Macquarie Library, Sydney, Australia.

Marc H. Bornstein. 1985. On the development of color naming in young children: Data and theory. *Brain and Language*, 26(1):72–93.

Roger W. Brown and Eric H. Lenneberg. 1954. A study in language and cognition. *Journal of Abnormal Psychology*, 49(3):454–462.

Stuart K. Card, Jock D. Mackinlay, and Ben Shneiderman, editors. 1999. *Readings in information visualization: using vision to think.* Morgan Kaufmann Publishers Inc., San Francisco, CA.

Wei-bin Chen. 2005. Comparative studies on cultural meaning difference of colors between china and western societies. *Journal of Fujian Institute of Socialism.*

Irvin L. Child, Jens A. Hansen, and Frederick W. Hornbeck. 1968. Age and sex differences in children's color preferences. *Child Development*, 39(1):237–247.

Richard E. Christ. 1975. Review and analysis of color coding research for visual displays. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 17:542–570.

Max Coltheart. 1981. The mrc psycholinguistic database. *Quarterly Journal of Experimental Psychology*, 33A:497–505.

John Gage. 1969. *Color and Culture: Practice and Meaning from Antiquity to Abstraction.* University of California Press, Ewing, NJ.

Paul Kay and Luisa Maffi. 1999. Color appearance and the emergence and evolution of basic color lexicons. *American Anthropologist*, 101:743–760.

Naz Kaya. 2004. Relationship between color and emotion: a study of college students. *College Student Journal*, pages 396–405.

Max Luscher. 1969. *The Luscher Color Test.* Random House, New York, New York.

Aaron Marcus. 1982. Color: a tool for computer graphics communication. *The Computer Image*, pages 76–90.

Barbara J. Meier. 1988. Ace: a color expert system for user interface design. In *Proceedings of the 1st annual ACM SIGGRAPH symposium on User Interface Software*, UIST '88, pages 117–128, New York, NY, USA. ACM.

Saif Mohammad and Graeme Hirst. 2006. Distributional measures of concept-distance: A task-oriented evaluation. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Sydney, Australia.

Saif Mohammad and Peter Turney. 2010. Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. In *Proceedings of the NAACL-HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, LA, California.

Saif M. Mohammad and Peter D. Turney. 2011. Crowdsourcing a word–emotion association lexicon. *In Submission.*

Li-Chen Ou, M. Ronnier Luo, Pei-Li Sun, Neng-Chung Hu, and Hung-Shing Chen. 2011. Age effects on colour emotion, preference, and harmony. *Color Research and Application*, pages n/a–n/a.

Norma S. Pribadi, Maria G. Wadlow, and Daniel Boyarski. 1990. The use of color in computer interfaces: Preliminary research.

Carl Ratner. 1989. A sociohistorical critique of naturalistic theories of color perception. *Journal of Mind and Behavior*, 10(4):361–373.

Paul Sable and Okan Akcay. 2010. Color: Cross cultural marketing perspectves as to what governs our response to it. pages 950–954, Las vegas, CA.

Carlo Strapparava and Gozde Ozbal, 2010. *The Color of Emotions in Texts*, pages 28–32. Coling 2010 Organizing Committee.

David Yarowsky. 1992. Word-sense disambiguation using statistical models of Roget's categories trained on large corpora. In *Proceedings of the 14th International Conference on Computational Linguistics (COLING-92)*, pages 454–460, Nantes, France.