8

Colour blindness and its contribution to colour vocabulary

A replication and extension

Emma Meeussen[1,2],

Supervisors: Dr. D. Dediu[2], and Prof. A. Majid[1,2]

[1] *Radboud University, Nijmegen*

[2] *Max Planck Institute for Psycholinguistics, Nijmegen*

October 22, 2014

Emma Meeussen

4079507

Abstract

Although often studied from a universalist perspective, colour terminology shows cross-linguistic variation. Several causes have been put forward for these differences, including environmental and cultural factors. The current study aims to further investigate the contribution of physiology to colour naming by assessing the relation between colour blindness and colour terminology and the possible mechanisms underlying such a relation. We collated a database of worldwide colour blindness incidence, UV-B irradiance and linguistic information in order to replicate a previous geographical review demonstrating that Daltonism (red-green colour blindness) was related to latitude and that a distinction between the terms for 'green' and 'blue' was correlated with Daltonism incidence (Brown and Lindsey, 2004). We extended previous work by performing regression analyses and a mediation analysis assessing the contribution of UV-B irradiance to colour vocabulary. We found that previous results stood up in a different, partly overlapping database using different statistical models and that UV-B plays a significant role in the relation between geographical location and colour blindness. Together, the studies suggest human physiology as a likely contributor to colour vocabulary.

1 Introduction

One of the most well-studied aspects of semantics is the the question whether meaning arises indepently of the words we use or because of them. What is the origin of meaning and how is it reflected in linguistic categories? How is the link between words and their meaning established? A view that used to be pervasive in the domain of cognitive science but is not commonly held anymore is that linguistic categories are universal (Brown and Lenneberg, 1954). According to this view, the terms we use to describe the world are a reflection of universal language-independent categories (but see Malt and Majid (2013) for a discussion on this topic) and category boundaries are assumed to reflect universal categories that are 'inherent' because they are innate, present in the environment or present in our perceptual input. This view is particularly salient in the case of the categorisation of parts of the body, in which it seems almost natural to assume that these categories are determined by the visual boundaries provided by the joints (e.g. Lakoff, 1987).

Despite these observations, we know that languages vary considerably with respect to the way they categorise the perceptual spectrum: for example, the language for smell differs greatly between some groups of languages (Majid and Levinson, 2011). Research into what underlies variation in semantic domains has focused mostly on the visual modality and has involved linguistic categories in action verbs (Hauk et al., 2004; Willems et al., 2010), categorisation of body parts (Majid, 2010) and spatial terminology (Pederson et al., 1998). Possible explanations for variability include the use of visual boundaries as cues for linguistic division of the human body and nonlinguistic spatial conceptualisation for spatial vocabulary. Whatever the cause for the observed variation, it is clear that there is no single underlying factor, innate or perceptual, that can explain the patterns we see across languages. As such, the idea that languages are somehow formed on the basis of a single universal blueprint can be considered unlikely (Evans and Levinson, 2009). Indeed, cross-linguistic inventarisation shows there are many ways to divide up each semantic domain, making the idea that semantic structure is only shaped by universal categories unfeasible (Malt and Majid, 2013). Other factors must also have an effect on the establishment of semantic structure.

The colour domain is one of the particular areas in which research into meaning and categorisation has focused. Colour terminology is well-suited for the study of semantics because of the many different putative contributors and the considerable variation across languages. Different suggestions have been made as to what underlies the commonalities and differences in colour vocabulary. Perhaps they arise because of the structure of our innate capacity for viewing colour although the observed variation doesn't point to such a universal human trait as a likely cause. Alternatively, the different surroundings that each population is subjected to during their members' lifetime or differences in genetic makeup might contribute to colour naming.

The variation in colour terminology within languages across the world can be measured by the number of colour terms but is also reflected in the way they divide up the colour spectrum. Some languages pattern like English, Dutch and German which contain eleven basic colour terms and have category boundaries at similar locations. Others divide up the colour spectrum in a different way: the Berinmo language for example uses three seperate terms for 'dark', 'light' and 'red' (Davidoff et al., 1999). Data from the World Atlas of Language Structures (Dryer and Haspelmath, 2013) show that from the total of 119 languages documented, 46 make use of less than 6 colour terms (Kay and Maffi, 2013). The reason for this variation has been investigated thoroughly over the past decades and has resulted in the emergence of several notable ideas. Because what is available in the perceptual input affects the development of a particular perceptual skill (Goldstone, 1998), it is likely that both the sociocultural environment and our physical surroundings play a role and a corresponding explanation has been put forward: societies that at some point developed dyeing techniques would have had a greater need of additional colour terms (Berlin and Kay, 1991).

In their work on colour naming across the world, Berlin and Kay (1991) divided the world's languages into six distinct stages according to the number of basic colour terms they possess. This inventarisation showed that blue and green do not always fall into different semantic categories. In light of the previously proposed explanations for variation it is relevant to review whether the presence of such a distinction as a reflection of linguistic variability can be used to assess the contribution of perceptual, genetic or environmental factors. One might speculate that since both blue and green are among the most prominent colours in our natural environment the use of their corresponding names is less likely to be under influence of technical advancements and might not be as susceptible to sociocultural factors as the use of other terms.

The cross-linguistic variation in colour vocabulary has been widely investigated and several causes have been proposed which can be divided up into sociocultural aspects (e.g. Ember, 1978, environmental effects on physiology (e.g. Lindsey and Brown, 2002) and genetic factors (e.g. Jameson and Komarova, 2009a,b). All of these factors would initiate variation in the system by altering the perceptual input: either by adding colours, thereby necessitating additional vocabulary, or by decreasing the distinctions that can be made, thereby altering the shape of the colour naming space. But additional, more cognitively oriented claims regarding colour vocabulary have also been made. For example, Regier et al. (2009) demonstrated that a single model based on the shape of the colour space and the position of focal colours therein, could be in concordance with both universal and variable patterns in colour naming across the world, although their explanation does not account for the variation in some languages like Waorani and Karajá. The presence of a distinction between 'blue' and 'green' can be studied as a representative of the variation found in colour terminology in general.

Because it is clear that colour vocabulary does not contain a predetermined or innate set of terms, it has been suggested that the observed worldwide variation corresponds to cultural factors. The introduction of dyeing techniques and the hypothetical resulting effect on colour inventory might account for the worldwide patterning we see with respect to colour naming systems (Berlin and Kay, 1991). This suggestion was not tested directly and was therefore not corroborated with statistical evidence but Ember (1978) did demonstrate that the size of the colour lexicon is affected by cultural complexity, though only in populations with lighter eye pigmentation. Cultural complexity apparently plays a role in the structure of the colour system. That it is not the only factor to be considered is reflected in Ember's conclusion that both cultural and biological factors should be taken into account when studying the structure of the colour vocabulary.

A clue as to which biological environmental factors culture might interact with is provided by Ember's demonstration that eye colour predicts the number of colour terms in interaction with societal complexity. This does not necessarily mean that there is a direct causal role for iris colour in either colour naming or colour perception, but it has been demonstrated that iris colour is related to macular pigment optical density ('the yellow spot') (Hammond et al., 1996). The physiological structure of the eye itself thus seems to bear a relation to colour vocabulary. More recently, it has indeed been demonstrated that acquired lens brunenscence, a discoloration of the lens under influence of exposure to UV-B radiation, affects the way colours in the short wavelength part of the spectrum are perceived (Lindsey and Brown, 2002). It is also possible that the damage that UV-B irradiance causes to the short-wave sensitive cones plays a role. All of this can result in tritanopic vision, which is characterised by the inability to distinguish colours in the green-blue-yellow range: green and blue in particular are perceived as a similar hue and yellow is not distinctly perceivable. This aberant perception and its putative effect on colour naming could affect the way colours are named in society as a whole. Lindsey and Brown (2002) investigated the possibility that exposure to UV-B could affect colour perception and influence the way languages are shaped with respect to colour terminology. It turned out that when viewing colours through a brunescent lens, participants named them as though there was no visible distinction between green and blue.

The effect of sunlight exposure on colour perception has further been demonstrated in a study by Laeng et al. (2007), who investigated colour vision deficiencies using the Farnsworth Munsell 100-Hue Test, one of the few tests sensitive to tritan defects, and showed that latitude of birth had an effect on the ability to perceive differences in the yellow-green-blue range of the spectrum. Although the direction of this effect was opposite to that proposed by the tritanopia hypothesis, it is an indication of the relation between sunlight exposure and colour vision.

While the results so far have demonstrated that the relationship between size of the colour lexicon and particular types of defective colour vision holds for acquired defects caused by UV-B irradiance, the possibilities of such an effect might not be limited to those. A different line of research has pursued investigations into the possibility of a relation between genetically determined colour blindness and colour naming both at the individual and population level (Bonnardel et al., 2006; Jameson and Komarova, 2009a,b; Komarova et al., 2007).

Hereditary colour blindness, although in lay terms often summarised under a single name, is actually an umbrella term for different defects with a range of effects on perception. Human colour perception is due to three types of cone cells on the retina that are sensitive to different wavelengths. They are coded for by two genes on chromosome 7 and one on the X chromosome. The genes coding for long and midwave sensitive cones have emerged from a single sequence of DNA that has diverged to form two different genes. This evolutionary path is still evidenced by the relatively small difference in spectral sensitivity between red and green sensitive opsins (the molecules in the retina affected by incoming light). Their proximity on the genome further means that these genes are susceptible to unequal recombination. Combined with the fact that they are situated on the X chromosome this significantly increases the likelihood of red-green colour blindness as opposed to colour vision deficits in the blue-yellow end of the spectrum (Deeb, 2005). Mutations in these genes result either in dichromacy (absence of one type of cones) or monochromacy (absence of two types of cones) while people suffering from achromatopsia lack retinal cones altogether. Within these basic types of colour blindness one can distinguish between complete absence of a particular type of cones or their reduced expression in the retina on the one hand - deuteranopia, protanopia and tritanopia - and a change in spectral sensitivity in either of the cones, resulting in anomalous colour vision, on the other hand. Studies on variation in human colour vision demonstrate that even between individuals with normal colour vision, there is variation in spectral sensitivity of the retinal cones (Deeb, 2006). Anopic colour vision deficits result in an inability to distinguish particular shades: red, orange and green in case of protanopia and deuteranopia; blue, green and yellow in case of tritanopia. The effect of aberrant colour vision is most evident in colour sorting tasks (Bonnardel et al., 2006).

The precise effects of colour vision abnormalities on language are difficult to examine experimentally: animal models are not suited for obvious reasons and tracing the mechanisms in humans would require either unethical manipulation or the collection of detailed longitudinal data. However, the issue has been investigated within the domain of computational modeling and *in silica* experiments have shown that the presence of agents with aberrant colour vision affects the colour system used by a population of agents as a whole, even when they appear in 'low, real-world-like numbers' (Jameson and Komarova, 2009a). The theoretical mechanism by which this effect comes about is that a change in colour discrimination abilities in colour blind agents affects their optimal partitioning of the colour space. The resulting change in categories within individual agents in turn would affect the optimal partitioning for a population as a whole, introducing divergent categories depending on the number of aberrant agents.

Although both correlational and *in silica* studies thus demonstrate a possible effect of colour vision deficits on colour vocabulary it is also known that in spite of colour vision deficits, even colour blind people can use colour terms adequately in day-to-day communication (Bonnardel et al., 2006; Cohen and Matthen, 2010; Cole et al., 2006; Shepard and Cooper, 1992) by making use of contextual information or brightness cues provided by the rods (Montag and Boynton, 1987; Paramei et al., 1998) or a residual cone mechanism (Nagy and Boynton, 1979). However, since it has not been investigated whether colour vision deficient people will also opt to use the colour terms which lie in their defective range during natural communication, this does not mitigate the possible effect of colour vision deficiencies.

Interestingly, Neitz et al. (2002) demonstrated the plasticity of the neural substrate of colour vision by letting subjects wear colour filters and determining their effect on colour perception during and after the experiment. They showed that the perceptual system actively uses input from the environment to adjust the weights of the chromatic channels. Human susceptibility to environmental input thus might provide a mechanism by which colour terminology could be influenced by environmental factors: people in the same environment will share a similar colour perceptual system, which in turn could affect their common colour vocabulary.

Another way in which colour vocabulary can become established under the influence of colour blind individuals is suggested by Lindsey and Brown (2002): if a small part of a population is unable to make a distinction between particular colours, communication about these colours would be affected in any conversation involving two colour blind people. The effect of even a small number of them would be greater in conversations involving three or more individuals. This way, an overarching colour category such as 'grue' could become the preferred term for a population, even though some members are able to distinguish between two categories.

An additional line of research has contributed to this issue, tying together the influence of genetics and environment. Brown and Lindsey (2004) investigated the possibility that acquired colour blindness is evolutionarily linked to hereditary colour blindness and that it facilitates colour vocabulary variation. They furthermore hypothesised a specific direction of the correlation between geographical location and colour terminology: because of the increased exposure to UV-B light, populations around the equator suffer more from damage to short-wavelength-sensitive cones and lens brunescence. They suggested that this in turn might mean that for these populations, good colour perception and discrimination in the red-green area of the spectrum are more important for survival than for populations further removed from the equator. Thus, there would be more selection pressure against colour blindness in the mid- and long-wave end of the spectrum resulting in a lower prevalence of red-green colour blindness for these populations. Based on a survey of Daltonism and dictionary data for a set of 118 languages they concluded that both genetics and the physical environment are related to linguistic expression of colours.

The question we aim to answer in the current study is whether the observed correlation between Daltonism and the presence of a distinction in the blue-green range will hold in a replication and whether the model proposed by Brown and Lindsey (2004) stands up under additional analyses. Replicating previous investigations provides an opportunity to establish corroborating evidence, and strengthen existing theories by demonstrating their validity on different data sets. Importantly, the analyses performed in Brown and Lindsey (2004) have several limitations, including collinearity between two of the variables under investigation. The current study addresses this issue by performing different regression analyses. If - as expected based on the corroborating evidence from linguistics and computational modeling - the effects remain, it is possible to review the mechanisms that could underlie them. In particular, we further test the model put forward by Brown and Lindsey (2004) claiming that the effect of latitude on colour blindness is mediated by UV-B irradiance. If the model withstands more stringent analyses, this provides strong corroborative evidence. In case of the current topic, we can further expand the implications toward the larger debate on the cultural and innate effects on language. Moreover, the mechanisms at play in the colour domain may shed light on those involved in different semantic domains and can have implications for the nature and time scale of language evolution in general: what are the factors involved and how quickly can they lead to language change?

The current study thus aims to investigate whether the effects found by Brown and Lindsey (2004) hold in a larger, partly overlapping sample. It also attempts to discern whether the assumptions made with respect to causality in the original study are valid. If so, the aim is to review whether the conclusions the authors draw are also in accordance with other theories regarding the causal mechanisms underlying colour terminology across the world.

We hypothesize that if indeed the findings of Brown and Lindsey (2004) are a true representation of the real world situation and their model is correct, the effects found will stand up in a replication on a different and larger sample. Indeed, we hypothesize that the effects will remain visible in a combination of replications of correlational analyses and linear modeling as well as additional regression analyses. We further hypothesize that if the model proposed by Brown and Lindsey (2004) is correct, a mediation analysis will demonstrate that the contribution of geographical location to colour blindness is due to UV-B irradiance.

2 Method

2.1 Materials

In order to assess a possible relation between worldwide colour blindness incidence and colour terminology, a database was compiled containing information on colour blindness incidence with corresponding geographic location and linguistic properties for a set of 154 languages. After the discovery that a similar database existed (Brown and Lindsey, 2004) and had been used for an analysis assessing the relation between Daltonism and the presence of a term for 'blue' across languages, we were generously granted access to Brown and Lindsey (2004)'s database. It was therefore possible to replicate this geographical investigation and compare the results of our analyses to the previous study. It is important to note that the databases were similar but they were not exact copies. Firstly, our dataset contained more languages (154 compared to 118). Although it may look like this would yield an extension of the prior research, our database was not a superset of the earlier database. In fact, in some cases where the same source was used, different incidence percentages were chosen as more informative or our judgement differed from the authors' with respect to linguistic or ethnologic classification, resulting in slight differences between the databases. We used UV-B data from the same year as the original database.

2.1.1 Data on colour blindness

The first step of data acquisition involved collecting the incidence of colour vision defects in male population across the world. These data were obtained from 85 different references, which were mostly general reports of anthropological expeditions in which colour blindness was assessed as part of a battery of tests, reports geared towards assessing colour blindness in a particular population, for example in medical students or school children and reports geared towards assessing the incidence of colour blindness in specific ethnic groups. We selected sources that had made use of any of the following standardized tests for colour blindness: the Ishihara test (Ishihara, 1917), the anomaloscope (Knoll, 1968), the Holmgren-Thomson wool test (Thomson, 1880) and the Hardy-Rand-Rittler pseudoisochromatic plate test (Hardy et al., 1954).

Where the original sources permitted, data on incidence of colour blindness were initially recorded in different variables: protanopia, protanomaly, deuteranopia and deuteranomaly incidence or the overarching categories deuterodefectivity and proterodefectivity incidence if no distinction was made between anopic and anomalous individuals. Additional categories were tritanopic and 'unclassified'. Tritanopia is not detectable by the use of Ishihara test plates - the most common type of test - and was reported in 5 cases. Because replication required data on incidence of Daltonism (deuteranomaly or deuteranopia) in particular, the different color blindness categories were subsequently recoded into a single variable as follows: if Daltonism or deuteranopia and deuteranomaly incidence was available, this was recorded as the Daltonism variable. If it was not, but tritanopia incidence was disposable, Daltonism was calculated as the difference between overall incidence and tritanopia incidence. If only overall prevalence of color blindness was available, this was included as Daltonism in the database. This classification is warranted as most of the sources made use of tests specifically geared towards assessing red-green color blindness and so the overall incidence likely reflects Daltonism incidence.

We excluded all colour blindness prevalence data based on samples smaller than $n = 50$ and all data for which no subdivision between male and female incidence was made. The complete list of references can be found in Appendix B.

2.1.2 Geographic data

Geographic data consisted of the country and region of origin of the populations studied in the primary sources as well as the corresponding geographic coordinates (latitude and longitude). Where possible these data were acquired from the original source. If no detailed geographical data were available, an informed estimate was made based on available ethnic and/or linguistic information. Geographic coordinates were obtained via Google Maps (GoogleMaps, 2014).

2.1.3 Linguistic data

Information on the language spoken by the population under investigation was obtained from the original source. If no detailed linguistic information was available, an informed decision was made based on available geographical and ethnic cues in the original source combined with language maps from Ethnologue (Lewis et al., 2014). For each entry in the database, the corresponding native language (L1) and its ISO 639-3 three letter language identifier, obtained from Ethnologue, were added. If applicable, a second language (L2) was also included.

In an extension of the database, each language was recorded in a seperate entry, which included name, ISO code, all colour terms for the basic colour catagories (Berlin and Kay, 1991) and a variable indicating whether the language had a seperate term for 'blue'. In order to collect the colour terms per language, dictionary data for each language were obtained from written or online dictionaries, online word lists and the Brown and Lindsey (2004) database. In case no such list or dictionary was available, dictionary data from the second language or a closely related language were used. In two cases, data from a protolanguage were used. The linguistic resources can be found in Appendix A.

2.1.4 UV-B irradiance

Using the geographic coordinates corresponding to each population in our primary database, we obtained data on UV-B irradiance and sunlight incidence from NASA (NASA, 2012). For each entry in the database, average monthly UV-B irradiance over the year 1998 corresponding to the geographical location was calculated from the raw NASA data.

2.2 Design

The database collated genetic, demographic and linguistic data for 210 linguistic groups corresponding to 268 ethnic groups and 242 geographical groups from 77 countries. From this database, 154 languages were studied. The variables under investigation in this study were Latitude, UV-B irradiance, Daltonism incidence and Presence of a term for 'blue'. All of the variables except Presence of a term for 'blue', which was dichotomous, took on numeric values.

2.3 Procedure

To ensure that a replication on the database would yield valid conclusions, all analyses were performed on two sets of data: a subset of the current database containing data from the same sources as the original database (Subset database), and our entire database (Full database). The first ensured that the methods used for analysis were similar to the original and that the results could be reproduced on a different database. The final replication served to investigate whether the results and conclusions would also hold for a larger set of data. One of the limitations of both the current study and the one by Brown and Lindsey (2004) is that the distribution of languages in the database is uneven with respect to geographical location and possibly in relation to language family. This means that any effect could partly be the result of this distribution rather than of a correlation. Performing additional regression analyses as well as a mediation analysis on both databases goes one step in the direction of eliminating such limitations.

2.4 Data analysis

All the statistical analyses were performed using R (R Core Team, 2014). In order to test the hypothesis put forward by Brown and Lindsey (2004), namely that the colour term system of a language is affected by latitude through UV-B irradiance, we replicated all the statistical analyses from the original article. First, we analysed the correlations between the all of the relevant variables: Daltonism, Cosine of latitude, Presence of a term for 'blue' and UV-B irradiance. The cosine of latitude was used as a measure of distance to the equator, rather than latitude itself. Secondly, two linear multiple-regression analyses were performed with Daltonism as dependent variable and Cosine of latitude (model 1) or UV-B irradiance (model 2) and Presence of a term for 'blue' as factors. We also performed a logistic regression predicting Presence of a term for 'blue' from Daltonism and either Cosine of latitude (model 1) or UV-B irradiance (model 2).

We expanded on previous work by carrying out a mediation analysis that served to investigate whether UV-B irradiance could be considered as a mediating factor in the relation between latitude and Daltonism, and how much of an effect was due to UV-B.

3 Results

In correspondence with the previous investigations we found that Daltonism prevalence and the cosine of Latitude were significantly correlated, both in the Subset database ($r = -.36$, $p = .0002$) and the Full database ($r = -.41$, $p = 1 \times 10^{-7}$). Moreover, Daltonism was significantly correlated with UV-B irradiance in both the Subset ($r = -.42$, $p = 2 \times 10^{-5}$) and the Full dataset ($r = -.46$,

$p = 2 \times 10^{-9}$), as was expected based on the similarly strong correlations we found between UV-B and cosine of latitude ($r = .97$, $p = 2 \times 10^{-16}$ on both databases).

To facilitate comparisons, the results from the correlation analyses per dataset are included in Table 1. The table demonstrates that the effects found in the Subset database are significant and correspond in magnitude to the effects found by Brown and Lindsey (2004), indicating that our data are similar enough to warrant further analyses. Moreover, the effects found in the Full database are very similar to the ones in both the Brown and Lindsey (2004) and the Subset database - in some cases even stronger - indicating that both the methods used and distribution of the Full database are solid. Table 1 further includes point-biserial correlations that demonstrate that the direction of the relation between 'blue' and the other variables is as expected based on the evolutionary hypothesis: 'blue' is more prevalent where UV-B is lower, where Daltonism is higher and at higher latitudes.

[insert Table 1 about here]

The relation between latitude, UV-B and Daltonism prevalence is further illustrated in Figures 1 and 2. Figure 3 summarises the relation between latitude and UV-B insolation, separated into languages with a term for 'blue' and those without. The relationship is approximated by a binomial and shows that languages without a term for 'blue' are more common around the equator. As it has been demonstrated that the effects found by Brown and Lindsey (2004) can be reproduced on the Full database, all of the figures included are based on that dataset.

[insert Figure 1; 2; 3 about here]

The linear multiple regression analysis for the Full database concerned all four variables under investigation (Daltonism, UV-B, Cosine of latitude and Presence of a term for 'blue'). It demonstrated that in a model predicting Daltonism prevalence from the factors of UV-B, Cosine of latitude and Presence of a term for 'blue', only UV-B ($t = -2.48$, $p = .015$) and 'blue' ($t = 2.60$, $p = .010$) were significant predictors, while Cosine of latitude was not ($t = 1.27$, $p = .21$). In the smaller database, only 'blue' approached significance ($t = 1.92$, $p = .058$) but not UV-B ($t = -1.69$, $p = .09$) or Cosine of latitude ($t = 0.94$, $p = .35$). A logistic regression revealed that on the integral database the presence of a term for 'blue' can be predicted by the factors UV-B ($z = -2.51$, $p = .012$) and Daltonism ($z = 3.36$, $p = .0008$) but not by the cosine of latitude ($p = .12$) while on the smaller database all factors proved to be significant contributors (Cosine of latitude: $z = 2.62$, $p = .009$; UV-B: $z = -3.16$, $p = .002$; Daltonism: $z = 3.03$, $p = .002$).

However, because two of the variables - UV-B irradiance and Cosine of latitude - in the analysis were highly correlated, multicollinearity may have affected the estimation of the coefficients of the predictors in the regression analyses. We therefore repeated the logistic and linear regression analyses by creating two separate models that included either UV-B or latitude as a predictor. The results show that both the Cosine of latitude ($t = -3.85$, $p = .0002$) and UV-B ($t = -4.45$, $p = 2 \times 10^{-5}$) are significant predictors of Daltonism in a model including Presence of a term for 'blue' ($t = 3.384$, $p = 0.0009$ in the Latitude model; $t = 2.834$, $p = .005$ in the UV-B model) as the only other factor. An analysis of variance of the two models demonstrated that only the model with UV-B irradiance as a predictor performed significantly better than the Brown and Lindsey (2004) study in predicting Daltonism incidence ($p = 0.014$ vs. $p = .21$) indicating that UV-B plays a stronger role in predicting Daltonism.

For the logistic regression model, the cosine of Latitude remained a significant predictor of Presence of a term for 'blue' in the absence of UV-B ($z = -2.87$, $p = .004$) together with Daltonism ($z = 3.67$, $p = .0002$). On the other hand, UV-B was a significant predictor of 'blue' in a model including only UV-B ($z = -3.45$, $p = .0006$) and Daltonism ($z = 3.36$, $p = .0008$). A likelihood ratio test comparing both of the smaller models to the earlier ones revealed that only the model including cosine of Latitude performed significantly better than the original in predicting 'blue' ($p = .009$). These results indicate that Latitude is a stronger predictor for 'blue' than UV-B. However, since these two variables are also significantly correlated, this is likely to be a matter of degree.

The effects of latitude on Daltonism and 'blue' can be graphically compared in the following figures. Figure 4 is a density plot showing that for languages without a term for 'blue', the prevalence of Daltonism is concentrated at lower numbers than for languages with a seperate term for 'blue'. The proportion of groups with a seperate term for 'blue', with a Daltonism prevalence over .032 (the median) and over .035 (the prevalence at which the number of languages with a term for 'blue' is equal to that without) is shown as a function of latitude (Figure 5) and UV-B dose (Figure 6). The final two figures demonstrate high correspondence in magnitude and shape of the influence of both latitude on Daltonism and 'blue' (Figure 5) and UV-B on Daltonism and 'blue' (Figure 6), indicating Daltonism and Presence of a term for 'blue' are similarly distributed.

[insert Figure 4; 5; 6 about here]

The model that Brown and Lindsey (2004) advocated and which ties together the variables under investigation posits that latitude and Daltonism are correlated and causally related through UV-B irradiance: locations at lower latitudes receive higher doses of UV-B, resulting in increased damage to the retina, in particular in the short wavelength end of the colour spectrum. The resulting higher incidence of acquired tritanopia puts selective pressure on the population, hypothetically resulting in lower prevalence of Daltonism. The significant correlations between each of the variables as well as the significant effect of latitude in the linear regression demonstrated that they are indeed linked. However, by performing a mediation analysis it is also possible to assess whether latitude contributes significantly to the prevalence of Daltonism through its link with UV-B irradiance or whether other - unmeasured - variables are at play (MacKinnon et al., 2007). Using this method it is also possible to assess the proportion of the effect that is mediated by UV-B. We performed such an analysis using the nonparametric bootstrapping method (Tingley et al., 2013) based on 1000 simulations. As expected due to the strong correlation between the cosine of latitude and UV-B, most of the effect of cosine of latitude on Daltonism was in fact mediated by UV-B: the average causal mediation effect was significantly different from zero ($ACME = -17.42$, $p < .001$, 95%, CI $[-34.13, -5.44]$) while the average direct effect was not ($ADE = 9.40$, $p = .18$, 95%, CI $[-2.58, 24.58]$). Moreover, it becomes evident that the mediated effect is opposite in sign to the direct effect, a type of mediation known as inconsistent mediation (MacKinnon et al., 2007). In this case the mediator acts as a suppressor variable. Because of the inconsistent mediation, the proportion of the effect that is mediated - 2.17 - was larger than one. Mediation analyses can also be performed using the Sobel test (Wang., 2014). Although this test also reveals that the ACME was significantly different from zero ($z = -3.36$, $p = 7.86 \times 10^{-4}$) the bootstrapping method is less conservative and therefore more powerful (MacKinnon et al., 1995).

The mediation analysis demonstrates two things: firstly, the directions of both the ADE and ACME are as expected based on results from the correlational and regression analyses. Secondly, the sizes of the ACME and ADE show that the effect of latitude on Daltonism is likely to be mediated by UV-B.

4 Discussion

The goal of the current study was to replicate prior research by Brown and Lindsey (2004) regarding the relation between Daltonism and colour terminology in a cross-linguistic study. They had demonstrated a correlation for Daltonism and the presence of a distinction between terms for 'green' and 'blue' across languages. Moreover, they had shown that the variables in the model which they assume to underlie such a correlation - latitude and UV-B irradiance - were correlated with Daltonism and 'blue' in their own right, thus demonstrating that their evolutionary model is in concordance with real world data.

We showed that the correlations found in the original study stood up in a replication on a larger, partly overlapping database. The effects were similar in size and in most cases the corresponding p-values were lower. Moreover, our mediation analysis showed that the effect of latitude on Daltonism is indeed largely mediated by UV-B irradiance. As the influence of UV-B on the macula and lens had previously been shown to affect both colour vision and naming (Lindsey and Brown, 2002), this points towards the role of a link between tritanopia and colour vocabulary in the relation between Daltonism and 'blue'. While a direct relation between tritanopia and Daltonism, UV-B and colour naming is difficult to demonstrate due to the lack of data on tritanopia incidence, the strong correlation between UV-B and Daltonism together with the mechanisms that are known to underlie the relation between UV-B irradiance and colour vision defects suggests that the mechanism of selection pressure proposed by Lindsey and Brown is a likely explanation.

Colour vocabulary is thus demonstrated not only to be influenced by environmental and cultural factors, but also by human physiology. As the influence of the yellow-blue perceptual system stands in relation to the genetic profile of a population and has been demonstrated to be influenced by environmental, biological input, interlanguage variability in semantics could be explained not just by cultural (Ember, 1978) or physiological (e.g. Jameson and Komarova, 2009b) differences, but also by innate properties. That such a mechanism is demonstrated to be a possible contributor here means that the makeup of the human perceptual system could play a role in the structure of language, a claim that has been the center of a long-standing debate on the role of the human cognitive makeup in semantic categories (e.g. Brown and Lenneberg, 1954; Jackendoff, 1983; Majid, 2014).

There are three different ways in which UV-B affects colour vision: lens brunescence, macula discolouration and destruction of short-wave-sensitive cones (Lindsey and Brown, 2002). Between them, these acquired effects have been demonstrated to affect colour perception in the short-wave end of the spectrum. Exactly which of these effects plays a role in the current relationship is of limited importance: the resulting colour vision deficiencies each correspond to a tritan defect, resulting in confusion pairs in the blue-yellow region. More important is the way in which the effect of UV-B corresponds to the genetic profile of a population. A large part of the literature on the relationship between colour blindness and colour vocabulary notably focuses on the direct putative effects of hereditary colour vision deficiencies.

Computational studies have demonstrated that presence of colour deficient agents in a population can cause shifts in colour vocabulary. Just as red-green colour blind individuals could affect the language of a population, yellow-blue colour blinds might influence the colour vocabulary. It has previously been demonstrated that such a change in vocabulary is the case in practice (Lindsey and Brown, 2002), and the current study provides corroborating evidence.

The mediation analysis performed as part of this study is valid on the assumption of a causal relationship between latitude and Daltonism. While causality cannot be demonstrated based on the correlation analysis alone, both the results of the regression analysis and the theoretical underpinnings suggest that the direction of the effect is as proposed here. An alternative explanation for the link between Daltonism and latitude would be that life at higher latitudes is easier for populations with high Daltonism incidence because of the reduced chance of acquiring tritanopic defects. This would not oppose the mechanism of selective pressure that Lindsey and Brown propose. Another alternative explanation of the data would be that as tritan defects can be caused by higher macular density, and this is shown to be related to eye colour (i.e. pigmentation of the iris), the effect of latitude is mediated through genetically determined dose of melanin in the eye rather than through the effects of UV-B irradiance. However, the mediation analysis demonstrated that most of the effect of latitude on Daltonism was mediated by the factor of UV-B irradiance.

Most computational studies investigating the relation between colour blindness and colour vocabulary have focused on a direct effect of Daltonism - the most common type of colour blindness - on language. Currrent theories propose that a few colour aberrant agents modify both the number of categories and the location of category boundaries in other agents through conversation, despite the fact that red and green lie on opposite sides of the confusion circle. While the results that have been obtained in Daltonism research are not at odds with our findings and may provide a general mechanism by which colour categories could become established in language, the hypothesis in Brown and Lindsey (2004) is particular about the cause and the direction of the effect. One to one comparison of the results is therefore not feasible. Moreover, our current database is not sufficient to test the claims put forward in computational studies as it does not contain information on the exact location of category boundaries. It would be possible to test for a correlation between Daltonism and number of categories, but as previous literature has made no precise claims regarding vocabulary size this does not seem feasible. In conclusion, a direct effect of Daltonism on colour terminology does not aid us in finding an explanation for the results found here in the blue-green range of the spectrum.

The study by Laeng et al. (2007) investigating the effect of UV-B exposure in arctic regions demonstrated an effect of sunlight that is inverse to the relation shown here: they showed that at higher latitudes, individuals were more confused as to the distinctions between green and blue. This effect is the opposite of what is presupposed to underlie our correlation, but it should be noted that it is not necessarily the result of (limited) exposure to UV-B during a lifetime that facilitates the effect in the arctic, but rather the amount of natural versus artificial lighting that reaches the eye during childhood.

A limitation of the current hypothesis is that it has not been conclusively demonstrated that Daltonism is negatively selected in humans. However, McNeill (1972) put forward the close link between the colours of ripe (red) and unripe (green) fruits that might make for an evolutionary advantage. Indeed it has been demonstrated that trichromacy is advantageous for detecting fruit in a background of leaves (Osorio and Vorobyev, 1996) and for detecting young red leaves in green foliage (Dominy and Lucas, 2001), making the selection pressure hypothesis a likely scenario.

Although the correlation demonstrated was robust, both the original study and our replication do not take into account genetic and linguistic relatedness. The fact that the results remained on a larger and different set of languages suggests that they might not be a side effect of the clustering of languages, but the effect could be modulated when controlling for relatedness. Further extensions will control for this to ensure that the effect is generalisable taking these variables into consideration. Furthermore, the current database of 154 languages can further be expanded to include a more diverse or representative sample.

Additionally, while the results up to now indicate that the shape of the colour vocabulary (i.e. number of colour terms and location of category boundaries) is caused by a combination of physiological, cultural and genetic factors which have individually been demonstrated to contribute, their relative contribution as well as the timescale remains open. It would therefore be interesting to see how the three interact, for example by collating a database including not just genetic and physiological variables but also an indication of technical advancement as a cultural contributor. By replicating the analyses on such an extended database, it would be possible to assess the amount of variation in colour vocabulary that is explained by each of these three factors.

If defective perception can be shown to affect the structure of a semantic domain, individual differences in human cognition in general might also affect language. For example, the variation in the human perceptual system, which was shown in rare cases to be tetrachromatic, particularly in women (Deeb, 2005), might lead to different colour categories in populations with more female influence. In a similar fashion, the makeup of the healthy cognitive system might explain not the differences but some of the more 'universal' properties of language - systematicities that we see in larger groups of languages and that have evolved over time (Baronchelli et al., 2010). While this is feasible, care should be taken not to confuse the effect of human genetics, physiology and culture on existing categories with their role in the emergence of categories within a semantic domain.

Actual historical evolutionary data on colour terminology is sparse (Komarova et al., 2007) so any conclusions with respect to timescale are drawn based on computational models. However, it is difficult to map what happens in a simulated population to the real world, particularly as the rate of change is heavily dependent on the chosen parameters and not all models include cross-generational communication.

Here we have demonstrated that real world data are consistent with the hypothesis that the presence of a seperate term for 'blue' is determined by UV-B irradiance and correlated with Daltonism incidence, showing these factors play a role in shaping the structure of colour vocabulary.

This provides an indication that individual cognitive differences have an effect on cross-linguistic variation in the colour domain.

5 Acknowledgements

I am indebted to Delwin Lindsey and Angela Brown for making available their original database. I thank my supervisors, Dan Dediu and Asifa Majid, for their supervision and feedback. In particular, I thank Asifa Majid for her advice regarding linguistic matters and Dan Dediu for his help in programming and statistical analyses. I also express my gratitude for their patience and unrelenting optimism.

References

Baronchelli, A., Gong, T., Puglisi, A., and Loreto, V. (2010). Modeling the emergence of universality in color naming patterns. *Proceedings of the National Academy of Sciences*, 107(6):2403 –2407.

Berlin, B. and Kay, P. (1991). *Basic color terms: their universality and evolution.* University of California Press.

Bonnardel, V. et al. (2006). Color naming and categorization in inherited color vision deficiencies. *Visual Neuroscience*, 23(3/4):637.

Brown, A. M. and Lindsey, D. T. (2004). Color and language: Worldwide distribution of Daltonism and distinct words for 'blue'. *Visual Neuroscience*, 21(03):409–412.

Brown, R. W. and Lenneberg, E. H. (1954). A study in language and cognition. *The Journal of Abnormal and Social Psychology*, 49(3):454.

Cohen, J. D. and Matthen, M. (2010). *Color ontology and color science.* MIT Press.

Cole, B. L., Lian, K.-Y., Sharpe, K., and Lakkis, C. (2006). Categorical color naming of surface color codes by people with abnormal color vision. *Optometry and Vision Science*, 83(12):879–886.

Davidoff, J., Davies, I., and Roberson, D. (1999). Colour categories in a stone-age tribe. *Nature*, 398(6724):203–204.

Deeb, S. (2005). The molecular basis of variation in human color vision. *Clinical Genetics*, 67(5):369–377.

Deeb, S. S. (2006). Genetics of variation in human color vision and the retinal cone mosaic. *Current Opinion in Genetics & Development*, 16(3):301–307.

Dominy, N. J. and Lucas, P. W. (2001). Ecological importance of trichromatic vision to primates. *Nature*, 410(6826):363–366.

Dryer, M. S. and Haspelmath, M., editors (2013). *WALS Online.* Max Planck Institute for Evolutionary Anthropology, Leipzig. (Available online at `http://wals.info`, Accessed on 2014-10-04).

Ember, M. (1978). Size of color lexicon: Interaction of cultural and biological factors. *American Anthropologist*, 80(2):364–367.

Evans, N. and Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, 32(05):429–448.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, 49(1):585–612.

GoogleMaps (2014). Google maps. Retrieved July and August, 2012, from `https://maps.google.com/`.

Hammond, J., Billy, R., Fuld, K., and Snodderly, M. (1996). Iris color and macular pigment optical density. *Experimental Eye Research*, 62(3):293–298.

Hardy, L. H., Rand, G., and Rittler, M. C. (1954). HRR polychromatic plates. *JOSA*, 44(7):509–521.

Hauk, O., Johnsrude, I., and Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, 41(2):301–307.

Ishihara, S. (1917). Test for colour-blindness. *Tokyo: Hongo Harukicho*.

Jackendoff, R. (1983). *Semantics and Cognition*, volume 8. MIT press.

Jameson, K. A. and Komarova, N. L. (2009a). Evolutionary models of color categorization. I. population categorization systems based on normal and dichromat observers. *JOSA A*, 26(6):1414–1423.

Jameson, K. A. and Komarova, N. L. (2009b). Evolutionary models of color categorization. II. realistic observer models and population heterogeneity. *JOSA A*, 26(6):1424–1436.

Kay, P. and Maffi, L. (2013). *Number of Basic Colour Categories*. Max Planck Institute for Evolutionary Anthropology, Leipzig.

Knoll, H. A. (1968). Anomaloscope for testing color vision. US Patent 3,382,025.

Komarova, N., Jameson, K., and Narens, L. (2007). Evolutionary models of color categorization based on discrimination. *Journal of Mathematical Psychology*, 51(6):359–382.

Laeng, B., Brennen, T., Elden, ., Gaare Paulsen, H., Banerjee, A., and Lipton, R. (2007). Latitude-of-birth and season-of-birth effects on human color vision in the Arctic. *Vision Research*, 47(12):1595–1607.

Lakoff, G. (1987). *Women, fire, and dangereous things.* Chicago: Chicago University Press.

Lewis, M. P., Simons, G. F., and Fennig, C. D. (2014). *Ethnologue: Languages of the world, Seventeenth edition.* SIL international Dallas, TX.

Lindsey, D. T. and Brown, A. M. (2002). Color naming and the phototoxic effects of sunlight on the eye. *Psychological Science*, 13(6):506–512.

MacKinnon, D. P., Fairchild, A. J., and Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, 58:593.

MacKinnon, D. P., Warsi, G., and Dwyer, J. H. (1995). A simulation study of mediated effect measures. *Multivariate Behavioral Research*, 30(1):41–62.

Majid, A. (2010). Words for parts of the body. *Words and the mind: How words capture human experience*, pages 58–71.

Majid, A. (2014). Comparing lexicons cross-linguistically. In Oxford Handbooks Online. doi:10.1093/oxfordhb/9780199641604.013.020.

Majid, A. and Levinson, S. C. (2011). The senses in language and culture. *The Senses and Society*, 6(1):5–18.

Malt, B. C. and Majid, A. (2013). How thought is mapped into words. *Wiley Interdisciplinary Reviews: Cognitive Science*, 4(6):583–597.

McNeill, N. (1972). Colour and colour terminology. *Journal of Linguistics*, 8(01):21–33.

Montag, E. D. and Boynton, R. M. (1987). Rod influence in dichromatic surface color perception. *Vision Research*, 27(12):2153–2162.

Nagy, A. L. and Boynton, R. M. (1979). Large-field color naming of dichromats with rods bleached. *JOSA*, 69(9):1259–1265.

NASA (2012). NASA Total Ozone Mapping Spectrometer (TOMS). http://toms.gsfc.nasa.gov/ery_uv/new_uv/.

Neitz, J., Carroll, J., Yamauchi, Y., Neitz, M., and Williams, D. R. (2002). Color perception is mediated by a plastic neural mechanism that is adjustable in adults. *Neuron*, 35(4):783–792.

Osorio, D. and Vorobyev, M. (1996). Colour vision as an adaptation to frugivory in primates. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 263(1370):593–599.

Paramei, G. V., Bimler, D. L., and Cavonius, C. R. (1998). Effect of luminance on color perception of protanopes. *Vision Research*, 38(21):3397–3401.

Pederson, E., Danziger, E., Wilkins, D., Levinson, S., Kita, S., and Senft, G. (1998). Semantic typology and spatial conceptualization. *Language*, pages 557–589.

R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Regier, T., Kay, P., and Khetarpal, N. (2009). Color naming and the shape of color space. *Language*, 85(4):884–892.

Rizopoulos, D. (2006). ltm: An R package for latent variable modelling and item response theory analyses. *Journal of Statistical Software*, 17(5):1–25.

Shepard, R. N. and Cooper, L. A. (1992). Representation of colors in the blind, color-blind, and normally sighted. *Psychological Science*, 3(2):97–104.

Thomson, W. (1880). An instrument for the detection of color blindness. *Transactions of the American Ophthalmological Society*, 3:142.

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., and Imai, K. (2013). *mediation: R Package for Causal Mediation Analysis.* R package version 4.4. `http://CRAN.R-project.org/package=mediation`.

Wang., B. (2014). *bda: Density Estimation for Binned/Weighted Data.* R package version 3.2.0-3. `http://CRAN.R-project.org/package=bda`.

Willems, R. M., Hagoort, P., and Casasanto, D. (2010). Body-specific representations of action verbs neural evidence from right-and left-handers. *Psychological Science*, 21(1):67–74.

	Brown and Lindsey (2004)	Subset database	Full database
cosine of latitude × UV-B	.82 $^{n.r.}$	.93 (.97***)	.93 (.97***)
cosine of latitude × Daltonism	.27**	.13 (-.37***)	.17 (-.41***)
UV-B × Daltonism	.22**	.18 (-.42***)	.21 (-.46***)
cosine of latitude × 'blue'		-.41***	-.41***
UV-B × 'blue'		-.49***	-.47***
Daltonism × 'blue'		.39***	.39***

Table 1: R^2 values, Pearson correlation coefficients (between brackets) and Point-biserial correlations (in blue) (Rizopoulos, 2006) from the original Brown and Lindsey (2004) study, for the Subset database and for the Full database. Significance ratings for the Point-biserial correlations were calculated using the Pearson method.

0.0001 '***'

0.001 '**'

0.01 '*'

0.05 '.'

0.1 ' '

n.s. Not significant

n.r. Significance not reported

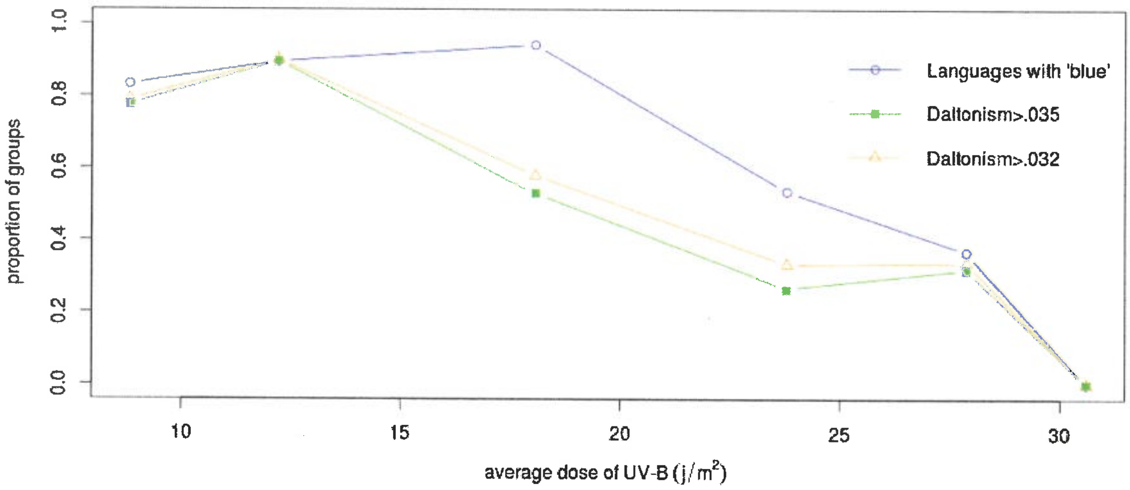

Figure 6: The proportion of linguistic groups using a term for 'blue', with Daltonism prevalence higher than 0.035 and with Daltonism prevalence over 0.032 as a function of UV-B.

Appendix A. Dictionary information

References

Bethany Arnold, Chris Carpenter, Guy Pledger, and Jack Brown. *Diktioneer Seereer-Angeleey; Serere-English Dictionary.* Peace Corps Senegal Volunteers and Trainees, 1st edition, May 2010. URL http://www.google.nl/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CFAQFjAA&url=http%3A%2F%2Fwww.pcsenegal.org%2Ffiles%2F18-diktioneer-seereer-angeleey-serere-english-dictionary%2Fdownload&ei=eJOCUILVKITXOQXk19D_Bg&usg=AFQjCNFfm_Pgnq2Nndx1wGdBrfDwjC5_1w&sig2=8KXSnfODSiceKEpqzIuUUw.

Oumar Bah. Pulaar: the colour spectrum [English translation]. Maneno Matamu, June 2011. Retrieved August 9, 2012, from http://manenomatamu.wordpress.com/2011/12/06/pulaar-the-colour-spectrum-english-translation/.

Brent Berlin and Paul Kay. *Basic Color Terms: Their Universality and Evolution.* University of California Press, Berkeley and Los Angeles, California, 1969. ISBN 9780520076358.

D.F. Bleek. A Bushman dictionary, 1956. Retrieved August 9, 2012, from http://www.worldcat.org/title/bushman-dictionary/oclc/1057812.

A. M. Brown and D.T. Lindsey. ICVS Supplemental. Unpublished, n.d.

Raphaël Confiant. Le premier dictionnaire du Créole Martiniquais. Potomitan, 2007. Retrieved July 16, 2012, from http://www.potomitan.info/dictionnaire/francais.php.

Richard Cook, Paul Kay, and Terry Regier. WCS Data Archives, 2003. Retrieved August 9, 2012, from http://www1.icsi.berkeley.edu/wcs/data.html.

Renato B. Figueiredo. Wolof-English and English-Wolof online dictionary. FREELANG, September 2011a. Retrieved July 15, 2012, from http://www.freelang.net/online/wolof.php?lg=gb.

Renato B. Figueiredo. Yupik-English and English-Yupik online dictionary. FREELANG, September 2011b. Retrieved August 14, 2012, from http://www.freelang.net/online/yupik.php?lg=gb.

Renato B. Figueiredo. Balkar-English and English-Balkar online dictionary. FREELANG, 2011c. Retrieved July 13, 2012, from http://www.freelang.net/online/balkar.php.

Renato B. Figueiredo. Kabardian-English and English-Kabardian online dictionary. FREELANG, 2011d. Retrieved July 13, 2012, from http://www.freelang.net/online/kabardian.php.

Renato B. Figueiredo. Sami-English and English-Sami online dictionary. FREELANG, 2011e. Retrieved August 8, 2012, from http://www.freelang.net/online/sami.php?lg=gb.

Renato B. Figueiredo. Sami-English and English-Sami online dictionary. FREELANG, n.d. Retrieved August 8, 2012, from http://www.freelang.net/online/sami.php?lg=gb.

Puttushetra Puttuswarmy Giridhar and Lalita Handoo. *Angami-English-Hindi dictionary.* Central Institute of Indian Languages, 1987. Retrieved July 7, 2014, from http://books.google.nl/books/about/Angami_English_Hindi_dictionary.html?id=ZCdkAAAAMAAJ&redir_esc=y.

Google. Google Translate. URL http://translate.google.com/.

Joseph H. Greenberg and Merritt Ruhlen. *An Amerind Etymological Dictionary.* Library of Congress Cataloging-in-Publication Data, Stanford, 2007. URL www.merrittruhlen.com/files/AED5.pdf.

Edna Romayne Headland. *Diccionario Bilingüe, Uw Cuwa (Tunebo) - Español, Español - Uw Cuwa (Tunebo).* Editorial Buena Semilla, Santafé de Bogotá, Colombia, 1 edition, 1997. ISBN 958-21-0153-9. Retrieved August 8, 2012, from www.sil.org/americas/colombia/pubs/37538.pdf.

Ruchi Jain. Bundeli and its concord pattern. FOSSSIL, 2012. Retrieved July 14, 2012, from http://www.fosssil.in/Essays_ruchi_LISSIM6.htm.

Mahesh Prasad Jaiswal. *A Linguistic Study of Bundeli: A Dialect of Madhyada.* E.J. Brill, 1962. URL http://books.google.nl/books?id=gcUUAAAAIAAJ&printsec=frontcover&hl=nl#v=snippet&q=colour&f=false.

Eliza Jones. *Junior Dictionary for Central Koyukon Athabaskan.* Unknown, 1978. URL http://www.uaf.edu/anla/collections/search/resultDetail.xml?id=K0972J1978i.

KT Kom. *A Description of the Kom Speech with Special Emphasis on Variation.* PhD thesis, North-Eastern Hill University, Shillong-22, 2009.

Bhadriraju Krishnamurti. *The Dravidian Languages.* Cambridge Language Surveys. Cambridge University Press, Cambridge, 2002. URL http://books.google.nl/books?hl=nl&lr=&id=54fV7Lwu3fMC&oi=fnd&pg=PP1&dq=Krishnamurti,+Bhadriraju+(2003).+The+Dravidian+languages+(null+ed.).&ots=k4mtYamqMy&sig=oc5bMMElF1ZaOVn6vZsHFfrjpDO#v=onepage&q=colour&f=false.

Wolf Lustig and Gilbert Ramírez. Ñe'ẽndy: Interactive Guarani Dictionary, 1996. Retrieved July 13, 2012, from http://www.uni-mainz.de/cgi-bin/guarani2/dictionary.pl.

Edward Horace Man. *A Dictionary of the Central Nicobarese Language.* W.H. Allen & Co., London, 1889. URL http://www.scribd.com/doc/77240797/Nicobarese-English-Dictionary.

Parmanand Mewaram. A Sindhi-English dictionary. Digital Dictionaries of South Asia, 1910. Retrieved July 14, 2012, from http://dsal.uchicago.edu/dictionaries/mewaram/.

Victor Palermo. Colour words in many languages. Omniglot. Retrieved August 9, 2012, from http://www.omniglot.com/language/colours/index.php.

M Penny, J Penny, and Pendur Durnath Rao. *Gondi Dictionary, Gondi - English - Hindi - Telugu.* Integrated Tribal Development Agency, Utnoor, Adilabad District, Andhra Pradesh, March 2005. URL http://gondwana.in/ap/dict/lexicon/index.htm.

Vladimir Pericliev. Kaingang and austronesian similarities between geographically distant languages. *Current Issues in Unity and Diversity of Language. Papers selected from 18th Intl. Congr. Ling. (Seoul, 2008)*, 1:875–892, 2009.

J.D. Sapir. Dictionnaire Jóola-Kujamaat (Diola-Fogny), 1994. Retrieved July 17, 2012, from http://people.virginia.edu/~ds8s/Kujamaat-Joola/DIC/Joola-Dic.html.

G Senft. Kilivila color terms. *Studies in Language*, 11(2):313–346, 1987.

Kibny'aanko Seroney. Kalenjin - English Dictionary. AfricanLanguages.com, 2009. Retrieved July 13, 2012, from http://africanlanguages.com/kalenjin/.

Paul Sidwell and Doug Cooper. Mon-Khmer Etymological Dictionary, 2007. Retrieved August 9, 2012, from http://sealang.net/monkhmer/dictionary/.

Ralph Lilley Turner. A Comparative Dictionary of Indo-Aryan Languages, 1962. Retrieved July 17, 2012, from http://dsal.uchicago.edu/dictionaries/soas/.

Rigina Turunen. Die farbbezeichnungen im mokscha-mordwinischen. *Finnisch-Ugrische Forschungen*, 57(1-3):167–194, 2002. URL http://www.doria.fi/handle/10024/20514.

Unknown. Aleut colors. NativeLanguages.org. Retrieved August 11, 2012, from http://www.native-languages.org/aleut_colors.htm.

Unknown. Dakota Sioux Color Words. Native Languages of the Americas website, 1998a. Retrieved August 9, 2012, from http://www.native-languages.org/dakota_colors.htm.

Unknown. Western Apache color words. Native Languages of the Americas website, 1998b. Retrieved August 8, 2012, from http://www.native-languages.org/apache_colors.htm.

Unknown. The Aramaic/Neo-Aramaic vocabulary builder - colors. Learn Assyrian.com, 1998c. Retrieved July 16, 2012, from http://www.learnassyrian.com/aramaic/colors/colors.html.

Unknown. English-Muyuw dictionary, 2007. Retrieved August 8, 2012, from www.sil.org/pacific/png/pubs/928474531217/Eng-Muyuw_lex.pdf.

Unknown. Arawak color words. NativeLanguages.org, 2008. Retrieved August 14, 2012, from http://www.native-languages.org/arawak_colors.htm.

Unknown. Ngata dictionary. Learning Media, 2010a. Retrieved on August 15, 2012, from http://www.learningmedia.co.nz/ngata/maori/kouru.

Unknown. Nepal Bhasa colors, 2010b. Retrieved July 13, 2012, from http://nepalbhasa.co.cc/learn/nb/0_colour.htm.

Unknown. Arabic colors with audio. MyLanguages.org, 2011. Retrieved July 14, 2012, from http://mylanguages.org/multimedia/arabic_audio_colors.php.

Unknown. Online English-Bhojpuri dictionary. Bhojpuria.com, 2012. Retrieved July 15, 2012, from http://bhojpuria.com/v2/dictionary.

Wheatbelt natural resource management. Nyungar Budjara Wangany, Nyungar Nrm word list & language collection booklet of the Avon Catchment region. Australian Government, n.d. Retrieved from http://www.wheatbeltnrm.org.au/resources/nyungar-dictionary.pdf.

Appendix B. Colour blindness incidence

References

Adam (1973). Colorblindness and gene flow in Alaskans. *American Journal of Human Genetics*, 25(5):564–566.

Adam, A., Doron, D., and Modan, R. (1967). Frequencies of protan and deutan alleles in some Israeli communities and a note on the selection-relaxation hypothesis. *American Journal of Physical Anthropology*, 26(3):297–305.

Aggarwal, S. and Sachdeva, M. (2004). A study of taste sensitivity of phenylthiocarbamide (PTC) and colour blindness among the Rajputs of Kasauli, Himachal Pradesh. *Anthropologist*, 6(4):289–290.

Agrawal, H. N. (1968). ABO blood groups, P.T.C. taste sensitivity, sickle cell trait, middle phalangeal hairs, and colour blindness in the coastal Nicobarese of Great Nicobar. *Acta Genetica Et Statistica Medica*, 18(2):147–154. PMID: 5301684.

Al-Aqtum, M. T. and Al-Qawasmeh, M. H. (2001). Prevalence of colour blindness in young Jordanians. *Ophthalmologica*, 215:39–42.

Appelmans, M. (1953). Colour defects among the natives of Congo. *Bull Soc Belg Ophthal*, (103):226–229.

Barnicot, N. and Woodburn, J. (1975). Colour-blindness and sensitivity to PTC in Hadza. *Annals of Human Biology*, 2:61–68.

Basu, A. (1964). The frequency of colour blindness in some population groups of Maharashtra (India). *Annals of Human Genetics*, 28(13):129–132.

Bhasin, M., Walter, H., Singh, I. P., Sing, S. M., and Singh, R. (1982). Genetic studies of Pangwalas, transhumant and settled Gaddis: 1. blood group polymorphisms and saliva secretor system. *Zeitschrift fr Morphologie und Anthropologie*, pages 77–96.

Bhasin, M. K. (1974). A genetic study on the Newars of Nepal Valley. *American Journal of Physical Anthropology*, 40(1):67–74.

Bhasin, M. K., Singh, I. P., Walter, H., Bhasin, V., Chahal, S. M. S., and Singh, R. (1986). Genetic studies of Pangwalas, transhumant and settled Gaddis: 4. colour blindness, mid-phalangeal hair, ear lobe attachment and behavioural traits. *Anthropologischer Anzeiger*, 44(1):45–53. ArticleType: research-article / Full publication date: Mrz 1986 / Copyright 1986 E. Schweizerbart'sche Verlagsbuchhandlung.

Bonn, B., Adam, A., Ashkenazi, I., and Bat-Miriam, M. (1965). A preliminary report on some genetical characteristics in the Samaritan population. *American Journal of Physical Anthropology*, 23(4):397–400.

Borges da Silva, G. and Gabaye Borges da Silva, G. (1983). The prevalence of color blindness in Senegal in a population of 4500 workers. *Bulletin De La Socit De Pathologie Exotique Et De Ses Filiales*, 76(5 Pt 2):841–845. PMID: 6608418.

Chan, E. and Mao, W. S. (1950). Colour-blindness among the Chinese. *The British Journal of Ophthalmology*, 34(12):744–745. PMID: 14821267 PMCID: 1323660.

Citirik, M., Acaroglu, G., Batman, C., and Zilelioglu, O. (2005). Congenital color blindness in young Turkish men. *Ophthalmic Epidemiology*, 12:133–137.

Clements, F. (1930). Racial differences in color-blindness. *American Journal of Physical Anthropology*, 14(3):417–432.

Collins, J. (1988). Prevalence of selected chronic conditions, United States, 1983-85. *Advance Data From Vital and Health Statistics*, (155):1–16.

Crooks, K. B. and Institute, H. (1934). On the incidence of color-blindness among negroes. *Science (New York, N.Y.)*, 80(2073):269. PMID: 17741747.

Cruz-Coke, R. and Barrera, R. (1969). Color blindness among Aymara in Chile. *American Journal of Physical Anthropology*, 31(2):229–230.

Das, S. R., Kumar, N., Bhatacharjee, P. N., and Sastry, D. B. (1961). Blood groups (ABO, m-n and rh), ABH secretion, sickle-cell, P.T.C. taste, and colour blindness in the Mahar of Nagpur. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, 91(2):345–355. ArticleType: research-article / Full publication date: Jul. - Dec., 1961 / Copyright 1961 Royal Anthropological Institute of Great Britain and Ireland.

Davies, L., Macdermid, C., Corbett, G., McGurk, H., Jerrett, D., Jerrett, T., and Sownden, P. (1992). Colour terms in Setswana: a lingusitic and perceptual approach. *Linguistics*, 30:1065–1103.

Dobson, T., Jackson, S., and Metcalfe, J. (1967). Red/ green colourblindness on the island of Ibiza. *Human Heredity*, 17:460–464.

Farokhfar, A. (2002). Prevalence of color blindness in the primary school students in Sari township 1999. *Journal of Mazandaran University of Medical Sciences*, 11(31):57–62.

Fernandes, J. L., Junqueira, P. C., Kalmus, H., Ottensooser, F., Pasqualin, R., and Wishart, P. (1957). P.T.C. thresholds, colour vision and blood factors of Brazilian indians. *Annals of Human Genetics*, 22(1):16–21.

Fischbach, L. A., Lee, D. A., Englehardt, R. F., and Wheeler, N. (1993). The prevalence of ocular disorders among hispanic and caucasian children screened by the UCLA mobile eye clinic. *Journal of Community Health*, 18(4):201–211.

Floris, G., Murgia, E., and Sanciu, M. G. (1992). Frequency of color blindness in Sardinia (Italy). *Bulletins et Mmoires de la Socit d'anthropologie de Paris*, 4(1):105–110.

Freire-Maia, A., Freire-Maia, N., and Quelce-Salgado, A. (1960). Genetic analysis in Russian immigrants PTC sensitivity, finger prints, color vision, hand clasping, and arm folding. *American Journal of Physical Anthropology*, 18(3):235–240.

Garth, T. (1936). Color blindness in Turkey. *Science*, 84(2169):85.

Garth, T. R. (1933). The incidence of color blindness among races. *Science*, 77:333–334.

Giles, E., Hansen, A. T., McCullough, J. M., Metzger, D. G., and Wolpoff, M. H. (1968). Hydrogen cyanide and phenylthiocarbamide sensitivity, mid-phalangeal hair and color blindness in Yucatín, Mexico. *American Journal of Physical Anthropology*, 28(2):203–212.

Grassivaro Gallo, P., Panza, M., Lantieri, P., Risso, D., Conforti, G., Lagonia, P., Piro, A., Tagarelli, G., and Tagarelli, A. (2003a). Some psychological aspects of colour blindness at school: A field study in Calabria and Basilicata (Southern italy). *Color Research & Application*, 28(3):216–220.

Grassivaro Gallo, P., Romana, L., Mangogna, M., and Viviani, F. (2003b). Origin and distribution of Daltonism in Italy. *American Journal of Human Biology*, 15(4):566–572.

Halberstein, R. A. and Crawford, M. H. (1974). Anomalous color vision in three Mexican populations. *American Journal of Physical Anthropology*, 41(1):91–94.

Haloi, A. (2011). Genetic characterization of the Biates of Saipung village of Jaintia Hills District, Meghalaya. *Anthropologist*, 13(3):239–240.

Heapost, L. (2000). A population genetic characterization of Estonians. *Anthropologischer Anzeiger; Bericht ber Die Biologisch-Anthropologische Literatur*, 58(2):137–154. PMID: 10962711.

Junqueira, P. C., Kalmus, H., and Wishart, P. (1957). P.T.C. thresholds, colour vision and blood factors of Brazilian Indians. II. Carajas. *Annals of Human Genetics*, 22(1):22–25. PMID: 13488181.

Kajanoja, P. (1972). A contribution to the physical anthropology of the Finns. variations of the ABO, rhesus, MN, P and Lewis blood group frequencies, PTC taste ability and colour blindness. *Annales Academiae Scientiarum Fennicae. Ser. A.5, Medica*, 153:1–12. PMID: 4629863.

Kalmus, H. (1957). Defective colour vision, P.T.C. tasting and drepanocytosis in samples from fifteen Brazilian populations. *Annals of Human Genetics*, 21(4):313–317. PMID: 13435643.

Kalmus, H., Degaray, A. L., Rodarte, U., and Cobo, L. (1964). The frequency of PTC tasting, hard ear wax, colour blindness and other genetical characters in urban and rural Mexican populations. *Human Biology*, 36:134–145. PMID: 14167588.

Kaur, N., Kumar, A., Kaur, G., Kaur Dhillon, J., and Singh, K. (2011). Study of colour blindness in Tibetan population. *Delhi Journal of Ophthalmology*, 21(3):45–47.

Khrumian, R. and Pickford, R. (1959). *Hrdit et frquence des anomalies congnitales du sens chromatique (dyschromatopsies)*. Vigot Freres, Paris.

Kilborn, L. G. and Beh, Y. T. (1934). The incidence of color-blindness among the Chinese. *Science (New York, N.Y.)*, 79(2037):34. PMID: 17813441.

Kim, B., Lee, S., Choe, J., Lee, J. H., and Ahn, B. H. (1989). The incidence of congenital color deficiency among Koreans. *Journal of korean Medical Science*, 4(3):117–120.

Koliopoulos, J., Iordanides, P., Palmeris, G., and Chimonidou, E. (1976). Data concerning colour vision deficiencies amongst 29,985 young Greeks. *Modern Problems in Ophthalmology*, 17:161–164. PMID: 1085863.

Kumar, N. and Sastry, D. B. (1961). A genetic survey among the Riang: A Mongoloid tribe of Tripura (North east india). *Zeitschrift fr Morphologie und Anthropologie*, 51(3):346–355. Article-Type: research-article / Full publication date: September 1961 / Copyright 1961 E. Schweizerbart'sche Verlagsbuchhandlung.

Kumar, S. and Kapoor, A. (2003). Taste sensitivity to PTC and incidence of colour blindness among the schedule castes of district Mandi, Himachal Pradesh. *Anthropologist*, 5(1):57–59.

Larrick, J. W., Yost, J. A., Kaplan, J., King, G., and Mayhall, J. (1979). Part one: Patterns of health and disease among the Waorani Indians of Eastern Ecuador. *Medical Anthropology*, 3:147–189.

Luxmi, Y. and Kapoor, A. (2011). A study of taste sensitivity of phenylthiocarbamide (PTC) and colour blindness among the Rajputs of Dadra and Nagar Haveli. *Anthropologist*, 13(2):163–165.

Mann, I. and Turner, C. (1956). Color vision in native races in Australasia. *American Journal of Ophthalmology*, 41(5):797–800. PMID: 13313671.

Maurya, N., Sachdeva, M., and Kalla, A. (2004). Incidence of colour blindness among the Nicobarese of Car Nicobar. *Anthropologist*, 6(4):299.

Mehra, K. S. (1963). Incidence of colour blindness in Indians. *The British Journal of Ophthalmology*, 47(8):485–487. PMID: 14189716 PMCID: 505834.

Mitchell, R. (1977). Red-green colour blindness in the Isle of Man and Cumbria. *Annals of Human Biology*, 4:577–579.

Modarres, M., Mirsamadi, M., and Peyman, G. (1997). Prevalence of congenital color deficiencies in secondary-school students in Tehran. *International Ophthalmology*, 20.

Mol, d. V.-d. and Went, L. (1978). Frequencies of different types of colour vision defects in the Netherlands. *Human Heredity*, 28:301–316.

Mueller, W. H. and Weiss, K. M. (1979). Colour-blindness in Colombia. *Annals of Human Biology*, 6:137–145.

Neel, J. V., Salzano, F. M., Junqueira, P. C., Keiter, F., and Maybury-Lewis, D. (1964). Studies on the Xavante Indians of the Brazilian Mato Grosso. *American Journal of Human Genetics*, 16(1):52–140. PMID: 14131874 PMCID: 1932461.

Niroula, D. and Saha, C. (2010). The incidence of color blindness among some school children of Pokhara, Western Nepal. *Nepal Med Coll J*, 12(1):48–50.

Norn, M. (1997). Prevalence of congenital colour blindness among inuit in east greenland. *Acta Ophthalmologica Scandinavica*, 75(2):206–209.

Oppolzer, A. and Winkler, E. M. (1980). Incidence of color blindess in East African negroes. *Anthropologischer Anzeiger; Bericht ber Die Biologisch-Anthropologische Literatur*, 38(2):117–120. PMID: 6968538.

Osuobeni, E. P. (1996). Prevalence of congenital red-green color vision defects in Arab boys from Riyadh, Saudi Arabia. *Ophthalmic Epidemiology*, 3(3):167–170. PMID: 8956321.

Post, R. H. (1962). Population differences in red and green color vision deficiency: a review, and a query on selection relaxation. *Eugenics Quarterly*, 9:131–146. PMID: 13985679.

Pramanik, T., Sherpa, M., and Shrestha, R. (2010). Color vison deficieny among medical students: an unnoticed problem. *Nepal Med Coll J*, 12(2):81–83.

Pullin, E. W. and Sunderland, E. (1963). A survey of phenylthiocarbamide (P.T.C.)-Tasting and colourblindness in Pembrokeshire, Wales. *Man*, 63:52–55. ArticleType: research-article / Full publication date: Apr., 1963 / Copyright 1963 Royal Anthropological Institute of Great Britain and Ireland.

Rahman, S. A., Singh, P. N., and Nanda, P. K. (1998). Comparison of the incidence of colour blindness between sections of Libyan and Indian populations. *Indian Journal of Physiology and Pharmacology*, 42(2):271–275. PMID: 10225056.

Rajasekhara, R. K. (2006). Incidence of colour blindness among the scheduled caste Madigas of Andhra Pradesh, India. *Man in India*, 86(1-2):105–110.

Rebato, E. and Caldern, R. (1990). Incidence of red-green color blindness in the basque population. *Anthropologischer Anzeiger; Bericht Über Die Biologisch-Anthropologische Literatur*, 48(2):145–148. PMID: 2378506.

Salzano, F. (1972). Visual acuity and color blindness among Brazilian Cayapo Indians. *Human Heredity*, 22:72–79.

Salzano, F. M. (1961). Rare genetic conditions among the Gaingang Indians. *Annals of Human Genetics*, 25(2):123–130.

Salzano, F. M. (1964). Color blindness among indians from Santa Catarina, Brazil. *Human Heredity*, 14:212–219.

Salzano, F. M. (1980). New studies on the color vision of Brazilian Indians. *Rev. bras. gent*, 3(3):317–327.

Salzano, F. M. and Neel, J. V. (1976). New data on the vision of South American Indians. *Bulletin of the Pan American Health Organization (PAHO)*, 10(1):1–8.

Sanghvi, L. D. and Khanolkar, V. (1949). Data relating to seven genetical characters in six endogamous groups in Bombay. *Annals of Human Genetics*, 15(1):52–76.

Serebrovskaia, R. (1930). Genetique du Daltonisme. *J. medicobiologique*, 4(5):355.

Seth, P. K. and Seth, S. (1973). Genetical study of Angami Nagas (Nagaland, india): A1A2BO, MN, Rh blood groups, ABO(H) secretion, PTC taste sensitivity and colour-blindness. *Human Biology*, 45(3):457–468. PMID: 4201492.

Shuey, A. M. (1936). The incidence of color-blindness among Jewish males. *Science*, 84(2175):228.

Siddiqui, Q. A., Shaikh, S. A., Qureshi, T. Z., and Subhan, M. M. (2010). A comparison of red-green color vision deficiency between medical and non-medical students in Pakistan. *Saudi Medical Journal*, 30(8):895–899. PMID: 20714688.

Simon, K. (1951). Colour vision of Buganda Africans. *East African Medical Journal*, 28(2):75–79. PMID: 14840361.

Singh, S. and Singh, H. (2007). A genetic study on the Purum (Chothe) tribe of Manipur. *Anthropologist*, 9(2):161–162.

Sunderland, E. and Ryman, R. (1968). P.T.C. thresholds, blood factors, colour vision and fingerprints of Jivaro Indians in Eastern Ecuador. *American Journal of Physical Anthropology*, 28(3):339–344.

Tagarelli, A., Piro, A., and Zinno, F. (2000). Color-blindness in Calabria (Southern italy): A north-south decreasing trend. *American Journal of Human Biology*, 12:17–24.

Vernon, P. E. and Straker, A. (1943). Distribution of colour-blind men in Great Britain. *Nature*, 152:690–690.

Vyas, G. N., Bhatia, H. M., Banker, D. D., and Purandare, N. M. (1958). Study of blood groups and other genetical characters in six Gujarati endogamous groups in Western India. *Annals of Human Genetics*, 22(3):185–199.

Vyas, G. N., Bhatia, H. M., Sukumaran, P. K., Balkrishnan, V., and Sanghvi, L. D. (1962). Study of blood groups, abnormal hemoglobins and other genetical characters in some tribes of Gujarat. *American Journal of Physical Anthropology*, 20(3):255–265.

Weinstein, E. D., Neel, J. V., and Salzano, F. M. (1967). Further studies on the xavante indians. VI. the physical status of the Xavantes of Simões Lopes. *American Journal of Human Genetics*, 19(4):532–542. PMID: 6036273 PMCID: 1706314.

Zein, Z. A. (1990). Gene frequency and type of colour blindness in Ethiopians. *Ethiopian Medical Journal*, 28(2):73–75. PMID: 2364935.