Do twins share the same dress code? Quantifying relative genetic and environmental contributions to subjective perceptions of "the dress" in a classical twin study

Omar A. Mahroo*

Department of Ophthalmology and Department of Twin Research and Genetic Epidemiology, King's College London, St. Thomas' Hospital Campus, London, UK
Retinal Service, Moorfields Eye Hospital, London, UK
UCL Institute of Ophthalmology, London, UK
Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK

Katie M. Williams*

Department of Ophthalmology and Department of Twin Research and Genetic Epidemiology, King's College London, St. Thomas' Hospital Campus, London, UK

Ibtesham T. Hossain

Department of Ophthalmology, King's College London, St. Thomas' Hospital Campus, London, UK

Ekaterina Yonova-Doing

Department of Twin Research and Genetic Epidemiology, King's College London, St. Thomas' Hospital Campus, London, UK

Diana Kozareva

Department of Twin Research and Genetic Epidemiology, King's College London, St. Thomas' Hospital Campus, London, UK

Ammar Yusuf

Department of Ophthalmology, King's College London, St. Thomas' Hospital Campus, London, UK

Ibrahim Sheriff

Department of Ophthalmology, King's College London, St. Thomas' Hospital Campus, London, UK

Mohamed Oomerjee

Department of Ophthalmology, King's College London, St. Thomas' Hospital Campus, London, UK

Talha Soorma

Department of Ophthalmology, King's College London, St. Thomas' Hospital Campus, London, UK

Christopher J. Hammond

Department of Ophthalmology and Department of Twin Research and Genetic Epidemiology, King's College London, St. Thomas' Hospital Campus, London, UK

The phenomenon of contrasting color perceptions of "the dress" photograph has gained scientific interest. The mechanism underlying why individuals differ is yet to be fully explained. We use the powerful twin model design to ascertain the relative contribution of genetic and environmental factors on perception variation. A sample of 466 twins from the British TwinsUK registry were invited to report what color they saw in a standard image of the dress in standard illumination. The mean age of the participants was 49.5 ($SD = 17.8$) years, and 85% were female. When asked to choose between white and gold (WG) or blue and black (BB), 328 reported WG (70.4%) and 135 (29.0%) reported BB. Subjects choosing WG were significantly older ($p < 0.01$), but there was no significant difference in gender. Monozygotic (MZ) twins were more concordant in their responses than dizygotic (DZ) twins (0.46 vs. 0.36). Twin modeling revealed that genetic factors accounted for

34% (95% confidence interval, 5%–59%) of variation in the reported color of the dress when adjusted for age, whereas environmental factors contributed 66% (95% CI, 41%–95%). This study suggests environmental factors play a significant role in how an individual perceives the color of "the dress."

Introduction

Human trichromatic color vision is enabled by the presence, within cone photoreceptors, of photopigments with differing spectral sensitivities. However, subjective perceptions of color are the consequence of multiple layers of retinal and higher neuronal processing. Objects emitting or reflecting similar wavelengths may be perceived to have the same or different colors based on intrinsic and extrinsic properties such as type of illumination and prior experience.

The widely reported phenomenon of starkly differing, largely dichotomized, color perceptions between subjects viewing a particular photograph of a dress raises intriguing questions and avenues for exploring visual neurophysiology (Bach, 2015; Chetverikov & Ivanchei, in press; Hesslinger & Carbon, 2016; Lafer-Sousa, Hermann, & Conway, 2015; Moccia et al., 2016; Rabin, Houser, Talbert, & Patel, 2016; Schlaffke et al., 2015; Vemuri, Bisla, Mulpuru, & Varadharajan, 2016; Winkler, Spillmann, Werner, & Webster, 2015). Are the differences shaped by prior environmental experience in terms of how subjects have learned to name colors in different contexts, or is there genetic determination to how the dress is perceived, for example, relating to polymorphisms in opsin sequences?

The twin study design permits a relative quantification of the importance of genetic and environmental factors by examining concordance within monozygotic (MZ) and dizygotic (DZ) twin pairs. Twin pairs share a common environment to some degree, but in addition, MZ pairs share 100% of their genome whereas DZ pairs share 50%. Traits that show significantly higher concordance in MZ compared with DZ pairs are likely to be more genetically determined. Calculation of heritability allows a formal calculation of the proportion of variance in a trait that is attributable to genetic factors. In this study, participants from the TwinsUK cohort (Moayyeri, Hammond, Hart, & Spector, 2013) were shown a standard image of "the dress" to explore heritability of the color perception in this interesting case.

Materials and methods

As part of a wider twin study, participants were recruited from the TwinsUK cohort over a one-year period (October 2015 to September 2016). TwinsUK is a nationwide registry in the United Kingdom that comprises several thousand adult twins who have volunteered to participate in research studies (Moayyeri et al., 2013). The cohort is mostly female and of European ancestry. The study had local ethics committee approval and conformed to the tenets of the Declaration of Helsinki.

Participants were shown a standard image of the dress, displayed on a flat screen computer monitor (luminance at viewing distance c.100 photopic cd m^{-2}) in standard indoor incandescent illumination (mean luminance c.40 photopic cd m^{-2}). They were asked to name the colors they saw. If their response differed from the two common alternatives "white and gold" (WG) or "blue and black" (BB), they were then asked to make a forced choice between these two alternatives. They were also asked whether or not they had come across the image before.

Proportions of participants choosing WG or BB were calculated. Casewise concordance was calculated for MZ and DZ pairs as $2C/(2C + D)$, where C is the number of pairs concordant and D the number of pairs discordant for seeing the dress as BB (which was the less common choice in our cohort).

In addition, we used a liability threshold twin model to estimate the relative genetic and environmental contributions to twin resemblance, as described in previous studies of binary traits (Tariq et al., 2014). A liability model assumes that the tendency to have a particular trait is continuous and normally distributed within the population and that those who exceed a certain threshold will express the trait of interest, in this example, seeing the dress as BB. Trait variation is decomposed as arising from three factors: the additive genetic component (A), the shared environment (C) or the nonadditive genetic component (D), and the unique environment (E), the latter also comprising measurement error. Given that the proportion of participants seeing the dress as WG or BB varies with age, we adjusted the liability thresholds according to age of the participant at the time of testing using the definition variable approach. The definition variable approach allows use of all the information on age, rather than separation or use of a narrow range of ages, by using the observed age to define the covariance structure for that trait (Neale, Boker, Xie, & Maes, 1999).

Maximum likelihood structural equation models, to compare variance and covariance in twin pairs, were constructed using the OpenMx (http://openmx.psyc.virginia.edu) package in R. A model with all parameters is specified (ACE/ADE), and then parameters are removed in a stepwise manner to create three parsimonious models with more degrees of freedom (AE, CE, and E). Twice the difference in log likelihoods between the full and submodels follows a

x^2 distribution, with the degrees of freedom equal to the difference in degrees of freedom between the full and subsequent models (likelihood ratio test). Each model is compared with the full model, and the best-fitting model is ascertained by identification of the model with the lowest Akaike information criterion (AIC). AIC is a measure of the relative quality of statistical models for a given set of data; the estimate is dependent on the number of parameters and the maximum value of the likelihood function of the model. The AIC gives no indication of heritability itself but enables appropriate selection of a model from which heritability can then be extracted. Heritability will then be estimated from the best-fitting model from the total contribution of genetic factors to trait variance, equivalent to A or A + D.

Results

Demographics and responses from all participants

In total, 466 twins participated in the study. The mean (*SD*) age was 49.5 (17.8) years, and 85% were female. When asked the colors of the dress, the majority (298 subjects, 63.9%) said the dress was WG, 126 subjects (27.0%) perceived BB, and 42 (9.0%) chose other colors (listed in Table 1). Of the 445 subjects (95.4%) from whom a response was recorded as to whether they had seen the dress image before, 312 participants (70.1%) had seen it before and 133 (29.9%) had not.

When asked to choose between WG and BB, 328 chose WG (70.4%) and 135 (29.0%) chose BB; this response was not recorded from three subjects (0.6%). The mean (*SD*) ages were 51.5 (17.0) years and 44.3 (18.9) years for WG and BB, respectively. The mean age was significantly older for participants choosing WG compared with BB ($p < 0.0001$, unpaired *t* test). The ages of the two groups are shown as boxplots in Figure 1. The proportions of men in both groups were similar: 14.6% and 17.0% for WG and BB, respectively ($p = 0.514$). There was a trend toward a greater proportion of older subjects seeing the dress as WG, for both men and women (Figure 2).

Responses and concordances in MZ and DZ pairs

After exclusion of unpaired twins or pairs whose zygosity was unknown, responses were included from 368 twins (117 MZ and 67 DZ pairs). Of the 234 MZ twins, 69 chose BB, and of the 134 DZ twins, 39 chose BB. These proportions were similar: 29.5% and 29.1%, respectively ($p = 0.938$). The concordance (concordant

Response	Number	%
White and gold	298	63.95
Blue and black	126	27.04
Blue and gold	3	0.64
Light blue and gold	1	0.21
Pale blue and gold	4	0.86
Pale blue and gray	1	0.21
Light blue and coffee	1	0.21
Blue and green	3	0.64
Blue and brown	1	0.21
Blue and gray	3	0.64
Blue and white	1	0.21
Blue and tan	1	0.21
Sky blue and taupe	1	0.21
Brown	1	0.21
Brown and cream	1	0.21
Green/pale blue and yellow	1	0.21
Lilac and beige	2	0.43
Lilac and black	1	0.21
Lilac and brown	5	1.07
Lilac and gold	2	0.43
Lime green and white	1	0.21
Off blue and green	1	0.21
Pale gold and green	1	0.21
White and mustard	1	0.21
Silver gray and brown	1	0.21
Purple and blue	2	0.43
Purple and black	1	0.21
Purple and gold	1	0.21

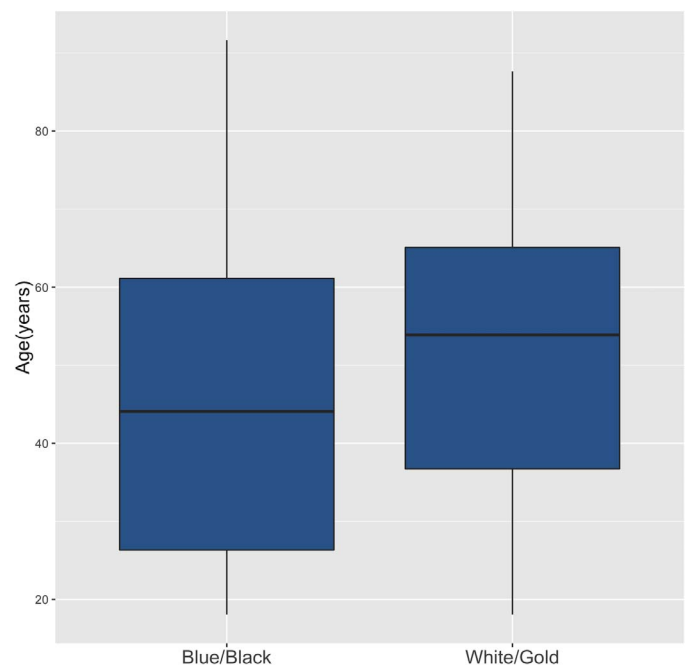Table 1. Initial responses from subjects asked to name the colors they saw in the dress.

Figure 1. Age distribution of subjects choosing blue and black or white and gold.
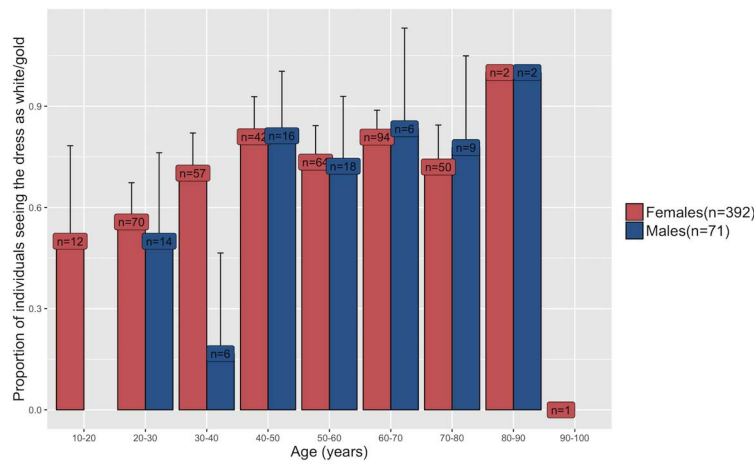
Figure 2. Proportion of subjects seeing the dress as white and gold by age and gender (proportional fraction and upper 95% confidence limit; number of subjects tested by age and gender = n).

twin pairs were pairs in which both twins saw the dress as BB) was 0.46 for MZ twins and 0.36 for DZ twins. Proportions and concordances for BB are shown in Figure 3.

We repeated concordance calculations with subjects who did not say WG or BB initially (see Table 1) removed, and found similarly higher concordance in MZ compared with DZ twins; when these subjects were removed (9%), concordances were 0.49 for MZ twins and 0.33 for DZ twins.
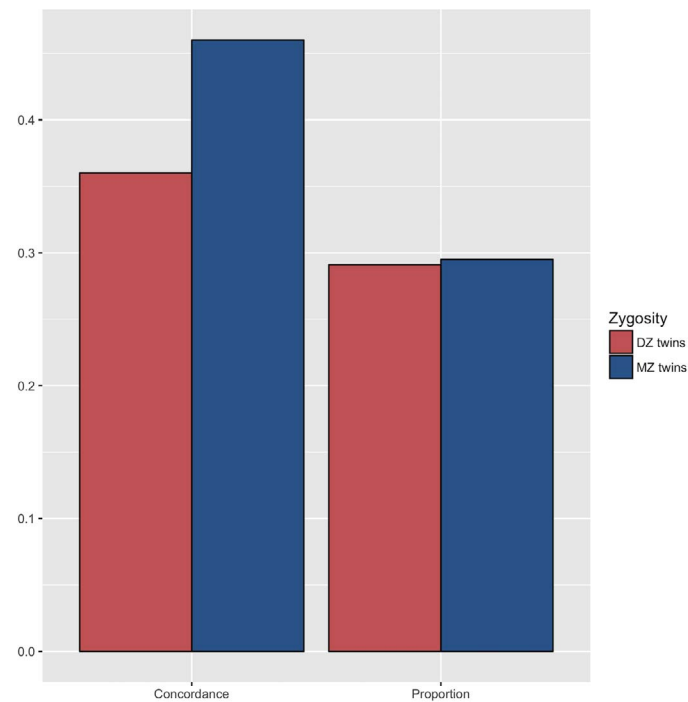

Figure 3. Proportions of subjects choosing blue and black in monozygotic and dizygotic twins and concordance for this choice within twin pairs.

The mean (*SD*) age was 46.5 (18.0) years and 53.9 (17.0) years for MZ and DZ twins, respectively. MZ twins were significantly younger ($p = 0.0001$, unpaired *t* test). As response differs by age, analysis was repeated on an age-matched sample of 224 twins (56 MZ and 56 DZ pairs), with the same number of twin pairs per decade. Mean (*SD*) ages were 52.6 (17.8) and 52.0 (18.0) for MZ and DZ pairs, respectively (*p* value for difference in ages 0.822). Concordances were 0.48 and 0.39 for MZ and DZ pairs, respectively.

Conventionally, concordance is calculated for the less common trait (akin to the "disease"), and hence, this was performed for the BB choice. If the more frequent trait is used, because this is more common anyway, concordances become deceptively high. Concordance rates for the WG combination would be 0.78 and 0.74 for MZ and DZ, respectively, but these are high largely because the majority of subjects chose this combination overall.

Estimation of heritability

From the full sample of complete twin pairs ($n = 184$ twin pairs), model selection statistics suggested that the AE model provided the most parsimonious fit to the observed data (Table 2, best fit model highlighted in bold). The AIC values between the models are comparable, indicating that there is little difference in how well the models best explain the variance in the observed data. However, the AIC value for the AE model is lowest (AIC for AE model is 2 lower than for ACE), showing that the addition of an extra parameter (C) does not improve the fit and is therefore not justified; thus, we chose the more parsimonious AE model. In this model, additive genetic factors (A) were estimated to contribute 38.3% to the total variance in

Variable	Model tested	Comparison model	minus2LL	df	AIC
Dress viewed as blue and black	ACE	—	546.992	453	−359.008
	ADE	ACE	546.973	453	−359.027
	AE	**ACE**	**546.992**	**454**	**−361.008**
	CE	ACE	547.783	454	−360.217
	E	ACE	554.024	455	−355.976
Dress viewed as blue and black (adjusted for age)	ACE	—	429.292	365	−300.708
	ADE	ACE	429.278	365	−300.722
	AE	**ACE**	**429.292**	**366**	**−302.708**
	CE	ACE	429.885	366	−302.115
	E	ACE	434.431	367	−299.569
Dress viewed as blue and black (age matched)	ACE	—	264.475	221	−177.525
	ADE	ACE	264.352	221	−177.648
	AE	**ACE**	**264.475**	**222**	**−179.525**
	CE	ACE	265.609	222	−178.392
	E	ACE	269.836	223	−176.165

Table 2. Results of the twin model fitting. The Akaike information criterion (AIC) is used to evaluate the most parsimonious model, which is highlighted in **bold**. For each model, the minus 2 log-likelihood (minus2ll) and degrees of freedom (df) for model comparison are given. A = additive genetic effects; D = dominant genetic effects; C = common environmental effects; E = unique environmental effects (and measurement error).

response (95% confidence interval [CI], 10.3%–62.1%), whereas environmental factors (E) contributed the remainder of variance (61.7%; 95% CI, 37.9%–89.7%).

Likewise, in the same sample adjusted for age at testing, AE was the best-fit model, and additive genetic factors (A) accounted for 33.9% of total variance (95% CI, 4.7%–59.0%) and environmental factors (E) 66.1% (95% CI, 41.0%–95.3%).

Using the smaller age-matched sample ($n = 112$ twin pairs), the AE model was again the most parsimonious fit, and heritability (A) was estimated at 47.2% (95% CI, 7.5%–76.9%). This latter estimate entails large CIs, reflective of the smaller sample size used for the calculations.

Discussion

This study aimed to explore, using the classic twin study approach, the relative contributions of genetic and environmental factors to subjective perceptions of color in the particular case of the dress image. When considering the choice as a binary outcome, concordance for choosing BB (the less common choice in our cohort) was higher in MZ than in DZ pairs, suggesting genetic factors play a role. Heritability was estimated at 34% when adjusted for the known influence of age. This would suggest that a significant proportion of the variance in response is determined by environmental factors.

Previous studies have suggested that perceptions of the dress may relate to subjects' assumptions regarding the blue light component of the illuminant (Lafer-

Sousa et al., 2015; Winkler et al., 2015) as well as its position (Chetverikov & Ivanchei, in press). Other studies have reported an association with natural pupil diameter (Vemuri et al., 2016) and with the density of the macular luteal pigment (Hesslinger & Carbon, 2016). Blue light detection begins with differential activation of cone opsins in the retina after passing through the anterior structures of the eye, in particular the lens, which filters out increasing proportions of blue light with age as cataract develops. Cone opsin sequences are genetically determined. Also, moderate to high heritability has been shown for cataract (Hammond, Snieder, Spector, & Gilbert, 2000) pupil diameter (He et al., 2009), and macular pigment optical density (Liew et al., 2005; and distribution, Tariq et al., 2014). Taken together, these findings might predict a substantial genetic contribution to subjects' response to the dress. Our study supports a significant genetic contribution but suggests that nongenetic, environmental factors may play an equal or even greater role. These may include prior lifetime experience to different spectral and luminance environments, shaping retinal and higher neuronal processing, and the development and evolution of naming of object colors in various contexts. Our findings suggest that color perception of the dress is largely influenced by environmental factors; these could be as important, or even quantitatively more important, than genetic factors. However, as our study relates directly to perceptions of the dress, it is difficult to estimate the reliability of any extrapolation beyond this special case to color perception more generally.

The CIs of our estimates of the contribution of genetic and environmental effects are relatively large.

This is in part due to the limitations of a small sample size but also reflective of similar twin studies. For example, in a twin study of the cognitive ability of face perception, similarly wide CIs were obtained for face recognition ability (heritability = 39%; 95% CI, 20%–54%; Zhu et al., 2010). It should also be noted that measurement error contributes to "E" (unique environmental effects) in twin studies. We did not formally measure test-retest repeatability in our study. It is known that some subjects can switch between perceiving the dress as WG or BB. In a previous study, three of 28 subjects who were retested in particular laboratory conditions reported a switch (Lafer-Sousa et al., 2015), which would give an estimate of 11% (95% confidence interval 0%–22%). The estimated proportion of variance in color perception attributable to environmental factors in this study was 62% (95% CI, 38%–90%), and so one might estimate that about one sixth of this could be due to test-retest variability acting as measurement error. However, it is also possible that the test-retest variability varies according to laboratory conditions, making this estimate somewhat tentative.

The present findings support the notion that observers who chose WG have a higher mean age than those choosing BB as reported previously (Lafer-Sousa et al., 2015; Moccia et al., 2016). Interestingly, we found WG to be the more frequent choice, whereas one previous study found BB to be more frequent (Lafer-Sousa et al., 2015) and another (Vemuri et al., 2016) found equal numbers making each of the two common choices. We suspect that the different age (and possibly gender) distributions in the different samples are contributory, with additional variation also potentially attributable to differences in viewing conditions (such as luminance of image and surroundings).

Some limitations of the study deserve mention. The specific demographics of the cohort (mostly northern European descent and mostly female) may limit the generalizability of the findings to other populations. Also, the confidence intervals for the age-adjusted heritability (likely to be more accurate as age is associated with the color interpretation of the dress) are relatively wide: The confidence interval for genetic influences is 5%–59%, whereas environmental influences fall within the range of 41%–95%. These results are reflective of the small sample size. Despite this, there are clear trends for a more environmental component to variation in responses across the three models we assessed. Given that many traits have a genetic component that is frequently in excess of 50%, the finding that variance in this trait appears to be more environmentally determined is interesting. Additional twin pairs, particularly DZ, would be helpful in increasing the power of the study. Third, as the twins were taking part in a number of eye tests, some of which required removal of spectacle correction, not all

twins may have viewed the image with optimal refraction, although we would not expect this to have differed substantially within twin pairs or between MZ and DZ pairs. Finally, for color vision studies in women, account should be taken of the fact that the existence of two X chromosomes (with different chromosomes inactivated in different cells) potentially with different genetic sequences for L and M opsins may affect color perception. Our cohort is mostly female, and limiting the analysis to men would have reduced the power considerably.

Our twin study has provided a preliminary estimate of the contribution of genetic factors to the variance in perception of the colors of the dress with a point estimate of heritability of 34%.

Keywords: twin study, genetics, dress

Acknowledgments

*These authors contributed equally to this work.
Commercial relationships: none.
Corresponding author: Christopher J. Hammond.
Email: chris.hammond@kcl.ac.uk.
Address: Departments of Ophthalmology & Twin Research, King's College London, St. Thomas' Hospital, London, UK.

References

Bach, M. (2015). Excitement over the color of a dress: Blue-black or white-gold or else? *Ophthalmologe*, *112*, 512–516.

Chetverikov, A., & Ivanchei, I. (in press). Seeing "the

dress" in the right light: Perceived colors and inferred light sources. *Perception.*

Hammond, C. J., Snieder, H., Spector, T. D., & Gilbert, C.E. (2000). Genetic and environmental factors in age-related nuclear cataracts in monozygotic and dizygotic twins. *New England Journal of Medicine, 342,* 1786–1790.

He, M., Wang, D., Console, J. W., Zhang, J., Zheng, Y., & Huang, W. (2009). Distribution and heritability of iris thickness and pupil size in Chinese: the Guangzhou Twin Eye Study. *Investigative Ophthalmology & Visual Science, 50,* 1593–1597. [PubMed] [Article]

Hesslinger, V. M., & Carbon, C. C. (2016). #TheDress: The role of illumination information and individual differences in the psychophysics of perceiving white-blue ambiguities. *Iperception, 7*(2), 2041669516645592.

Lafer-Sousa, R., Hermann, K. L., & Conway, B. R. (2015). Striking individual differences in color perception uncovered by 'the dress' photograph. *Current Biology, 25,* R545–R546.

Liew, S. H., Gilbert, C. E., Spector, T. D., Mellerio, J., Marshall, J., van Kuijk, F. J., . . . Hammond, C. J. (2005). Heritability of macular pigment: A twin study. *Investigative Ophthalmology & Visual Science, 46,* 4430–4436. [PubMed] [Article]

Moayyeri, A., Hammond, C. J., Hart, D. J., & Spector, T. D. (2013). The UK Adult Twin Registry (TwinsUK Resource). *Twin Research and Human Genetics, 16,* 144–149.

Moccia, M., Lavorgna, L., Lanzillo, R., Brescia Morra, V., Tedeschi, G., & Bonavita, S. (2016). The Dress: Transforming a web viral event into a scientific survey. *Multiple Sclerosis and Related Disorders, 7,* 41–46.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (1999). *Mx: Statistical modeling* (5th ed.). Richmond: Department of Psychiatry, Medical College of Virginia.

Rabin, J., Houser, B., Talbert, C., & Patel, R. (2016). Blue-black or white-gold? Early stage processing and the color of 'the dress.' *PLoS One, 11,* e0161090.

Schlaffke, L., Golisch, A., Haag, L. M., Lenz, M., Heba, S., Lissek, S., . . . Tegenthoff, M. (2015). The brain's dress code: How The Dress allows to decode the neuronal pathway of an optical illusion. *Cortex, 73,* 271–275.

Tariq, A., Mahroo, O. A., Williams, K. M., Liew, S. H., Beatty, S., Gilbert, C. E., . . . Hammond, C. J. (2014). The heritability of the ring-like distribution of macular pigment assessed in a twin study. *Investigative Ophthalmology & Visual Science, 55,* 2214–2219. [PubMed] [Article]

Vemuri, K., Bisla, K., Mulpuru, S., & Varadharajan, S. (2016). Do normal pupil diameter differences in the population underlie the color selection of #thedress? *Journal of the Optical Society of America. A, Optics, Image Science, and Vision, 33,* A137–A142.

Winkler, A. D., Spillmann, L., Werner, J. S., & Webster, M. A. (2015). Asymmetries in blue-yellow color perception and in the color of 'the dress.' *Current Biology, 25,* R547–R548.

Zhu, Q., Song, Y., Hu, S., Li, X., Tain, M., Zhen, Z., . . . Liu, J. Heritability of the specific cognitive ability of face perception. *Current Biology, 20,* 137–142.