Forthcoming in *Visual Neuroscience*

Reply to Philipona and O'Regan[1]

Kent Johnson and Wayne Wright
KJ: johnsonk@uci.edu
WW: wwright2@csulb.edu

Abstract: This paper responds to Philipona & O'Regan (2006), which attempts to account for certain color phenomena by appeal to singularities in the space of "accessible information" in the light striking the retina. Three points are discussed. First, it is unclear what the empirical significance/import is of the mathematical analysis of the data regarding the accessible information in the light. Second, the singularity index employed in the study is both mathematically and empirically faulty. Third, the connection drawn between their findings and some data from the World Color Survey is lacking in quantitative analysis in places where it is needed. The difficulties raised prevent Philipona & O'Regan's conclusions from being accepted.

Introduction

David Philipona and Kevin O'Regan (hereafter P&O; unnoted page citations are to their 2006) offer an account of certain color phenomena grounded entirely in what they term the "accessible information" in the light, to the exclusion of neural processing beyond the retina. For P&O (p.332), the accessible information "about a light [is] the restricted information about the spectral composition of that light which is accessible through an organism's photopigment set." They explicitly focus on the cone quantum catch, employing nothing dependent on subsequent neural coding or processing (p.332). Surprisingly, P&O report that, unlike most colors, focal examples of the four Hering primaries (red, green, blue, and yellow) are "singularities": i.e., one or two of their three coordinates in the particular bases P&O employ are approximately zero. They contend that this unexpected result accounts for data from psychophysical studies and the World Color Survey (WCS). That all this could be achieved without invoking specialized neural pathways and

[1] The authors' names are listed in alphabetical order. We would like to thank Kimberly Jameson, Louis Narens, A. Kimball Romney, and Jack Yellott for helpful discussion. We would also like to thank an anonymous referee of this journal for useful comments and suggestions.

cortical representations would be a major achievement and would suggest a distinctive new way of studying perception.

P&O's project has three parts: they (i) mathematically analyze the data, (ii) measure the results with a particular index of singularity, and (iii) compare these measurements with some empirical data (e.g., from the WCS). However, before P&O's conclusions can be accepted, several issues about each of these three parts must be addressed. We take them in turn.

Mathematical analysis

The first part of P&O's project involves representing the accessible information from illuminants, which may have been reflected off a surface, as 3×3 matrices. The rows of these matrices correspond to the information that is accessible to the L, M, and S photoreceptors, and the columns correspond to the three (hypothesized) components of natural light. P&O show that if the matrix associated with the unreflected light is invertible, then the amount of accessible information in a reflected illuminant that is accessible to a given type of photoreceptor is a linear function of the amount of accessible information in the unreflected illuminant available to the photoreceptors. P&O perform an eigenanalysis on the 3×3 matrix of coefficients of these linear combinations, using the resulting eigenvectors as a new basis for representing the 3×1 vectors of information accessible to the three photoreceptors. In this new basis, P&O report that best exemplars of the four Hering primaries are of less than full dimensionality.

Initially, it might appear that by switching to a basis of eigenvectors, P&O are employing some familiar technique such as principal component analysis (PCA), singular value decomposition (SVD), or the like (e.g. Basilevsky 1994). The latter techniques are often used in attempts to uncover the "true" dimensionality of a data set, by projecting the data into a subspace determined by the eigenvectors of matrices associated with, but not identical to, the original matrix of interest.[2] Moreover, the matrices from which the eigenvectors are extracted are symmetric (i.e., each such matrix is equal to its own transpose). The symmetry of these matrices is crucial to ensuring that their associated eigenvectors have the extremal properties that support many of their most important empirical interpretations and uses.[3]

In contrast, P&O's eigenanalysis is carried out directly on the matrices of interest, which will in general not be symmetric. In fact, P&O report real eigenvalues for only 88% of their matrices (p.334), so that 12% of them have complex eigenvalues. But it is a theorem that all the eigenvalues of a real symmetric matrix are themselves real. More generally, we should expect this lack of symmetry. After all, the *ij*th element of this matrix is $\sum_{h=1}^{3} c_h \int R_i(\lambda) S(\lambda) E_h(\lambda) d\lambda$, where $[c_1, c_2, c_3]^T$ is the *j*th column of the inverse of the matrix U^{-1}, where U has as its *kl*th entry the value $\int R_k(\lambda) E_l(\lambda) d\lambda$.[4] There is extremely little reason to think that we will often (or ever)

---

[2] A classical PCA might extract the eigenvectors from the covariance matrix of the variables defining the original data set A; SVD takes them from A^TA and AA^T.

[3] E.g., the symmetry (and positive semidefiniteness) of the relevant matrices in PCA and SVD ensure that the linear combination associated with the eigenvector with the *i*th ($1 \leq i < n$) largest eigenvalue maximizes the amount of remaining variance algebraically projectable onto the line it determines, and is orthogonal to the previous $i - 1$ combinations (e.g. Basilevsky 1994).

[4] The integrals are interpreted as follows. $E_h(\lambda)$ is the spectral power distribution of the *h*th component of the illuminant at wavelength λ, $S(\lambda)$ is the reflectance function of some given surface, and $R_i(\lambda)$ is the absorption rate of photoreceptor *i* at wavelength λ.

have matrices where $a_{ij} = a_{ji}$ always holds.[5] In short, P&O's analysis must receive its interpretation and support without the help of such techniques.[6]

We wish to raise two issues regarding P&O's transformations to the new bases of eigenvectors. The first issue is that when switching from a basis whose members have some clear empirical interpretation to a new one (e.g., one based on the eigenstructure of a matrix), there is no guarantee that the new basis will have any empirical interpretation. But the empirical interpretability of these vectors is important. To see this, consider a baker's claim to have developed a cake made with just one ingredient. What is this ingredient? It is 48 parts flour, 16 parts butter, 32 parts sugar, 16 parts egg, 10 parts milk, and 1 part baking powder. That one could reorganize one's kitchen to contain linear combinations of ingredients lends little credence to the claim that there "really is" a one-ingredient cake mix (i.e., a mix of "reduced dimensionality"). Similarly, the fact that the *original* basis vectors could each be legitimately viewed as empirically unidimensional does not ensure that various linear combinations of them, such as eigenvectors, will themselves be empirically meaningful. In the general case they are not, and cases of reduced dimensionality have no real interpretation.[7]

---

[5] As P&O themselves note (p.333), there could be more than three relevant illuminant spectra. This would mean that the matrices would not be square, and hence not symmetric.

[6] An anonymous reviewer observed that there are other, potentially more relevant, options than those involving the eigendecomposition of the matrices. Independent components analysis, for instance, attempts to represent the data as a combination of sources that are not merely linearly independent, but fully statistically independent from one another (e.g., Bell and Sejnowski 1995, Comon 1994). Such techniques involve methods very different from those used by P&O.

[7] This is especially so because eigenvectors are typically drastically different from the original basis, even when all their elements are real numbers.

Another way to see this last point is that we need some reason to focus on some particular types of bases, or else it becomes all too easy to find cases of reduced dimensionality. After all, *any* vector has reduced dimensionality in *some* bases. (*Proof:* the zero vector in P^n is singular in any basis. For any other non-zero vector $\mathbf{x}$, we can expand it into a basis $\{\mathbf{x}, \mathbf{y}_1, \ldots, \mathbf{y}_{n-1}\}$. In this new basis, $\mathbf{x}$ is singular, having the coordinates [1, 0, 0, 0,…, 0].) So if there is no reason why a basis of eigenvectors is somehow privileged, the significance of the results becomes unclear.

In fact, P&O offer a "straightforward interpretation" of the basis of eigenvectors. They write that "the three eigenvectors of A^S will provide a basis for the accessible information space such that elements of this basis do not mutually interact with each other when they are reflected by the surface $S(\lambda)$, and are simply individually scaled by the associated eigenvalues" (p.333). However, the appeal of this interpretation depends in part on the degree to which there is an empirically (as opposed to merely mathematically) significant interpretation of the new basis of eigenvectors. If one or more of the basis eigenvectors were such a strange or twisted combination of the original elements – i.e., the nine quantities $\int R_i(\lambda)S(\lambda)E_h(\lambda)d\lambda$ – so as to lack any plausible interpretation, it is hard to see the empirical import of their findings. (Indeed, even if one is apriori committed to an analysis along the lines of the last term of P&O's proposition (4), uninterpretable basis vectors would suggest a reexamination of other aspects of P&O's methods; e.g., Does the basis become more interpretable if the natural light is analyzed differently? What is the effect of deriving a distinct basis for each surface?) Moreover, from the perspective of the original basis, none of the original "interactivity" is eliminated; rather, it is encoded into the coordinates of the eigenvectors. Also, these eigenvectors are individually determined by the global structure of the matrix.[8] Empirically, this means that, e.g., the element of the vector associated with information from the S photoreceptor will be influenced by information from the M and L receptors. While this outcome is mathematically innocuous, it does increase the difficulty of empirical interpretation.[9]

We turn now to the second issue concerning P&O's mathematical transformation of the data. As described above, the elements of the matrix that produces the basis of eigenvectors are partly determined by a surface reflection function S, which will be different for different surfaces. Hence, for different surfaces, the resulting matrix and the basis of eigenvectors it generates will be different. Thus each of the singularities P&O find comes from a distinct basis. This explains why the coordinates for green and blue are, up to P&O's own acceptable rounding error, identical (p.335). Such an outcome is unproblematic, since they are coordinates from different bases.

---

[8] This can be seen by considering the coefficients of the matrix's characteristic polynomial, which determines the eigenvalues, which in turn determine all the possible eigenvectors.

[9] Two additional comments about P&O's use of eigenvectors are worth noting. (i) Since the matrix from which they were extracted is not symmetric, the eigenvectors themselves will not in general be orthogonal to one another. Thus, in one sense, the eigenvectors do interact with one another: they will not lie at right angles to one another and the (Pearson) correlation between pairs of them is nonzero. (ii) P&O's method does not involve the simple change of basis technique familiar from introductory algebra. As their construction of the matrix A^S and their discussion in equation (4) show, only the vector u(E) is expressed in terms of a basis of eigenvalues. A^S, however is not modified; it is left as a transformation in the original coordinate system. Thus, from the perspective of the basis of eigenvectors, the elements of A^S cannot be interpreted as linear combinations of integrals that give information about the photoreception of certain types of reflected illuminants. Of course, in the original basis, this is precisely the information A^S contains. With respect to the basis of eigenvectors, though, it is doubtful that A^S has any empirical interpretation at all.

In any one 3-dimensional basis, there are 6 different types of (non-zero) singular coordinates: a vector's coordinates could have exactly one zero coefficient in three ways, and there are another three ways it could have exactly two zero coefficients. It is not clear why P&O claim that their methods predict exactly 4 types of singular coordinates (p.335). In any case, as long as each basis allows for at least one type of singularity, P&O's analysis makes no predictions about how many chips will have singular coordinates, or which ones will do so. After all, their coordinates are each from different bases. Thus, in principle, every surface could have singular coordinates with respect to its matrix, or none could. In order to better understand the phenomenon P&O report, it would be useful to know whether there were any other near singularities in the full analyzed data set and, if so, which ones. Since P&O round values as large as .08 down to 0, in fairness, this standard should be observed in the consideration of the other color chips' new coordinates.

Singularity index

Let us leave aside the concerns regarding P&O's method of coordinate transformation, and assume that it is appropriate for their purposes. The next stage of P&O's project is to measure the "singularity" present in the new coordinates. P&O describe two types of singularity, but the type of singularity which they claim is possessed by (and only by) focal examples of the four Hering primaries is that which occurs when one or more coordinate in the new basis is (nearly) zero. The most natural way to measure this type of singularity would be to take the minimum of the three coordinates. Surprisingly, however, this is not what P&O do. Instead, when (a, b, c) are the new coordinates of a surface (ordered according to size: $a \geq b \geq c$), the singularity of these coordinates is measured as the maximum of a/b and b/c (where each of these are scaled by a constant to always yield a value between 0 and 1). Importantly, this measure is unreliable: sometimes it delivers the normatively correct ordering, sometimes not. Consider the vector coordinates A = (.4, .22, .1), B = (.6, .4, .2), and C = (1, 1, .4). In terms of approaching reduced dimensionality, the vectors should be ranked ABC. P&O's singularity index ranks them CAB. This singularity index suffers from other anomalies, as well. It is invariant over multiplication of the vectors by a nonzero constant: e.g., multiplying A, B, and C by 100, 10, and .5 will not change their singularity indices or relative rankings. Similarly, (.09, .09, .09) would be considered less singular than (.9, .8, .7). (In fact, P&O's singularity index appears better suited to measure their other notion of singularity, which concerns the degree to which the three coordinate take the same value.[10])

Data from the WCS

We now set apart the preceding concerns to consider the final stage of P&O's project: their use of the singularity indices taken from the transformed coordinates. Of particular interest is their comparison of a graph of singularity measures to a plot of data from the WCS (p.335, Fig.3; see also Fig.2 of Regier et al 2005).

P&O appear to be comparing their singularity index plot to that of a focal choice task, where for each of their language's basic color terms,[11] WCS participants were to select from the stimulus array a particular Munsell chip as the best exemplar of that term. About the relationship between WCS focal choice data and singularity index scores for the WCS stimuli, they write:

---

[10] Although, for reasons already discussed, it is unlikely that identical coordinates in the new basis will typically correspond to achromaticity.

[11] A language's basic color terms form the smallest set of simple terms that could be used to name any color. For more on how basic color terms were determined, see Cook et al (2005).

> The chips most often given a name by widely separated human cultures … and which we call ‘red’, ‘yellow’, ‘green’, and ‘blue’ in English, can be seen to be within one chip of those having maximally singular reflecting properties …. It could thus be argued that the reason the colors ‘red’, ‘yellow’, ‘green’, and ‘blue’ are so often singled out among all other colors as being worth giving a name, is that surfaces of these colors have the particularity that they alter incoming light in a simpler way than other surfaces. (pp.335 – 336)

However, of the 110 languages surveyed in the WCS, only 38 had basic color terms for all four Hering primaries and some of those languages had more than one basic term corresponding to the same Hering primary; see Kuehni (2005).[12] Several WCS languages that lacked a basic term for one of the Hering primaries had basic terms for colors considered non-fundamental from the standpoint of the Hering primaries. Furthermore, while many languages have separate terms for green and blue, many instead have a basic color term that covers both; i.e., “grue” terms. In sum: 65% of WCS languages find at least one of the Hering primaries to not be “worth giving a name” and many languages have basic color terms that do not have a simple correspondence with them. Thus, clarification is needed of just what color naming phenomenon P&O aim to account for when they highlight the special status of certain surfaces according to their singularity index.

Inspection of the full WCS focal choice data presented in Cook et al (2005) and Regier et al (2005) raises further questions about just what P&O can claim to have accomplished. Approximately 2640 total speakers took part in the WCS. The peaks of the focal choice distribution, along with the number of participants who selected a particular Munsell chip as a

[12] Jameson (forthcoming) emphasizes Kuehni’s findings in arguing against the idea that pan-human features of color experience are the largest factor in explaining cross-cultural color naming patterns.

best exemplar of one of his or her language's basic color terms are: A0 (2048), J0 (1988), G1 (668), C9 (752), F17 (351), F29 (253). A0 and J0 are at the extreme ends of the range of achromatic Munsell chips used in the WCS and correspond to best exemplars of English "white" and "black", respectively. The remaining chips (G1, C9, F17, and F29) align well with averaged focal points for English "red", "yellow", "green" and "blue" respectively; see Cook et al (2005), Sturges & Whitfield (1995). As P&O (p.335) mention, the WCS focal choice peaks line up nicely with the chips having the highest singularity scores: G2, C9, F16, and H31. (Note, however, that H31 is two chips and four coordinate changes away from F29, not just one, as P&O claim.)

The numbers associated with the chromatic chips reveal substantially less consistency than is found for the achromatic ones. If a chip gets a very high singularity score, fewer than 30% of the participants will pick it as a best exemplar. Using the focal choice data for F17 and F29 to evaluate chips F16 and H31 that P&O have identified as special, fewer than 15% of participants select as best exemplars chips with singularity values in the middle of the scale.

P&O do not quantitatively analyze the similarities and differences between their singularity index values and the WCS focal choice data. Hence, it is difficult to know what to make of the patterns just noted, to the (unknown) extent they exist. For example, simply looking at reflectance coefficients and singularity values does not reveal why a high score on a measure of chromatic singularity should lead to only ca. 10-29% of participants picking a chip as a best exemplar. Moreover, there appear to be marked differences in the graphs (Fig. 3) comparing the WCS with their singularity index. E.g., the spread, skew, and kurtosis about the corresponding peaks look rather distinct. Also, a cluster in the focal choice plot in the vicinity of hue line 30 has two peaks while its singularity index counterpart has only one. Without a more detailed analysis of the geometry of these two plots, it is virtually impossible to draw any conclusions whatsoever about their apparent similarities.

Finally, P&O's strategy does not suit their aim to explain the existence of "four special surface colors" by taking into account only "constraints satisfied by natural illuminants and surfaces" (p.331). For that purpose, it would be appropriate for P&O to plot the Munsell chips from the WCS stimulus set according to a metric defined by the bases and reflectance coordinates they derive through their eigendecomposition.[13] Ideally, chips with high singularity values would appear at the edges of such a plot and chips would cluster in the space in a way that suggests four distinct categories. That is, P&O should be able to (in some sense) reconstruct the WCS stimulus array from their biological account of reflectance and thus also derive any similarities between singularity values and focal choice results. However, P&O cannot do this, since, as we have seen, their eigendecomposition is done chip-by-chip, rather than across the entire set of chips. They are then left to use singularity values alone, plotted on the original stimulus array, to make their case. That approach, however, relies heavily on the conceptual organization of the Munsell chip set used in the WCS to establish the four special color categories; the chips are arranged in 40 equal hue steps at eight levels of lightness.

Conclusion

P&O's project is clearly "high risk – high yield." The difficulties we have raised regarding interpretation and evidential strength prevent their conclusions from being accepted. We stress, though, that we do not claim P&O's conclusions are false, only that they are in need of further justification. Upon further analysis, it may be that P&O have found a real surprise in need of

[13] We thank an anonymous referee for suggesting the basic idea behind this point.

explanation. In that happy circumstance, though, more work will be needed to interpret the nature of the surprise, and its plausible explanation(s). It will be a fascinating finding if color research can productively proceed as P&O envisage, by eschewing consideration of neural pathways and cortical representations.

References

Basilevsky, A. 1994. Statistical Factor Analysis and Related Methods. New York: Wiley-Interscience.

Bell, A., and T. Sejnowski. 1995. "An Information-Maximization Approach to Blind Separation and Blind Deconvolution." Neural Computation, 7, 1129 – 1159.

Comon, P. 1994. "Independent Component Analysis, A New Concept?" Signal Processing, 36, 287 – 314.

Cook, R., P. Kay, & T. Regier. 2005. "The World Color Survey Database: History and Use." In H. Cohen and C. Lefebvre (eds.) Handbook of Categorisation in the Cognitive Sciences. Amsterdam and London: Elsevier.

Jameson, K. forthcoming. "Where in the World Color Survey is the support for the Hering Primaries as the basis for Color Categorization?" In J. Cohen & M. Matthen (eds) Color Ontology and Color Science. Cambridge, MA: MIT Press.

Kuehni, R. 2005. "Focal Color Variability and Unique Hue Stimulus Variability." Journal of Cognition and Culture, 5, 409-426.

Philipona, D. & J.K. O'Regan. 2006. "Color naming, unique hues, and hue cancellation predicted from singularities in reflection properties." Visual Neuroscience, 23, 331-339.

Regier, T., P. Kay, & R. Cook. 2005. "Focal colors are universal after all." Proceedings of the National Academy of Science, 102, 8386–8391.

Sturges, J. & T. Whitfield. 1995. "Locating basic colours in the Munsell space." Color Research and Application, 20, 364-376.