CO-054

SQUARING THE CIRCLE: BIVARIATE COLOR MAPS AND JACQUES BERTIN'S CONCEPT OF 'DISASSOCIATION'

REIMER A.
GFZ German Research Centre for Geosciences, POTSDAM, GERMANY

INTRODUCTION

The 'Semiology of Graphics' has made Jacques Bertin (1967/74/83) a household name in basically all disciplines involving two dimensional visualizations and Bertin's seminal work appears in the references sections of visualization and cartography papers in a nearly compulsory fashion. The concept of visual variables, probably his most famous concept, is mentioned and used on a wide scale. In spite of this, there is a growing amount of research that tacitly and implicitly does not adopt the full matrix of Bertin's visual variable typology. While some authors question the actual expressive capabilities of certain variables, such discussions are often only concerned whether a given variable is capable of expressing nominal, ordinal or numerical data. We argue that Bertin's twin concept of 'Associativity' and 'Disassociativity' of the respective visual variables has been mostly ignored or misinterpreted, especially in the english-speaking world.

Bertin has labeled the variables 'Value' and 'Size' as being 'Disassociative'. One major consequence of the postulated 'Disassociativity' of 'Value' is fundamentally at odds with all research efforts trying to construct so-called bivariate color maps: using 'Value' and 'Color' to construct a bivariate color matrix will destroy Bertinian 'Selectivity' i. e. the ability for the pre-attentive perception of areas with the same attributes.

We track the role of 'Disassociativity' (or lack thereof) in anglo-american textbooks on thematic cartography and the research papers from statistical graphics, visualization and cartography domains on bivariate color mapping starting with the U. S. Census efforts (Meyer et al. 1975) and that of Bruce Trumbo (1981) up to the present day. This is flanked by a comparison to continental European ideas of 'Color' and 'Value' usage in cartography. There is ample textual evidence that both traditions have structurally different interpretations of color as a general phenomenon and how Bertin's ideas have been received: The anglo-american tradition follows a development path that is more heavily influenced by the technical and computational representation of color in so-called color spaces whereas the continental European tradition is more influenced by heuristic, partially pre-scientific notions of best cartographic practices. We argue for a re-examination of both traditions in order to investigate to which degree bivariate color mapping is indeed possible at all. We show that the question of 'Disassociativity' lies at the heart of such a debate.

DISASSOCIATIVITY IN THE SEMIOLOGY OF GRAPHICS

Probably one of the most overlooked attributes of visual variables found in Jacques Bertin's Semiology of Graphics is the notion of 'Disassociativity' of 'Value' and 'Size'. Bertin's basic idea is that 'Associativity' of a variable is the norm. Associative visual variables allow the symbol to be differentiated by other visual variables, the variation of associative variables can be ignored. No matter which pattern, hue, orientation or form a symbol has, its other variables, if any, can be decoded without interfering selectivity. As such, Disassociativity is a negative attribute of a visual variable. An attribute, that negates the expressive powers of other variables, in other words, an attribute that is concerned with the interaction of visual variables with each other (see also Spiess, 1970). Bertin's main argument for explaining that behavior lies with the visibility of the symbols: 'Value' and 'Size' change the visibility of the symbols, so that a variation of visibility makes it impossible or very hard to ignore that variation. A variation in 'Value' as well as 'Size' dominates all other variations, and even destroys non-variations. In the Bertinian sense, 'Selectivity' can be understood as being equal in concept to preattentive perception capability. Selective perception in that sense denotes the ability to look at all symbols sharing a visual variable to be viewed at once, without effort. If a visual variable is not selective, or if selectivity is obstructed by 'disassociative' variables, then the viewer cannot look at all the symbols sharing another specific instance of a variable at once preattentively. The viewer has to compare each symbol with each other symbol to differentiate between them, i. e. search sequentially. While visual variables and their attributed expressive power have been investigated empirically (e. g. Garlandini and Fabrikant 2009), to our knowledge there has been no empirical study that was concerned with proving or disproving Bertin's notion of 'Disassociativity'. It is important to stress that Bertinian 'Associativity' is not about the ability to perceptually form a closed Gestalt, but strictly about the non-intervention with the 'Selectivity' of a given visual variable. A very simple example would be to imagine a large number of black symbols and two red symbols far away from each other. Bertinian selectivity of 'Hue' correctly predicts the preattentive ability to view all red symbols at the same time, not that they form a Gestalt group.

In his seminal work, “How Maps Work”, MacEachren (1995) mixes Bertinian selectivity with Gestalt-group formation, and comes to the conclusion that 'Associativity' and in turn 'Disassociativity' are false (“unlikely”, MacEachren 1995, p. 92) contentions, as they ignore distance between the symbols. The warrantors for his renunciation of Bertinian 'Associativity/Disassociativity', Pomerantz and Schwaitzberg (1975) were actually doing studies on symbols with the same 'Size' and 'Value' but different 'Form', and were indeed concerned with the formation of Gestalt-groups. While distance of symbols is important for Gestalt groups, it is irrelevant for 'Selectivity' and in turn 'Asocciativity'. Even if one follows MacEachren's own argument for a moment, citing Pomerantz and Schwaitzberg (1975) for highlighting the importance of distance of symbols, the choice seems unproductive. They say: “The fact that grouping breaks down over large distances, while reassuring, is not of great interest since it could hardly have turned out otherwise: certainly, if the elements were spaced so widely that they could not both be seen within a single fixation, no grouping could occur.” (p. 356). Apart from a misrepresentation of 'Disassociativity' itself, MacEachren cited work does not even prove his misrepresented point.

More recent studies, such as Garlandini and Fabrikant (2009) repeat the fateful wording “to group” (p. 196) to describe “Associativity”, potentially continuing the misrepresentation albeit technically staying true to Bertin's implications. As we will show below, most other cartographic literature does not even mention 'Disassociativity'.

CARTOGRAPHIC TEXTBOOKS AND CHOROPLETH MAPS UNTIL 1975

Before the idea of bivariate color maps became prominent, most textbooks on thematic cartography, no matter whether from a continental or anglo-american tradition, took similar stances on choropleth maps and their coloring. The anglo-american tradition being more influenced by statistics and the continental one being more encyclopedic, the former concentrated on choropleths, dot maps and chorochromatic maps and the statistical procedures used to obtain class divisions (e. g. Raisz 1962, Dickinson 1963, Truran 1975), the latter was most concerned with different taxonomies for a more extensive variety of thematic maps (e. g. Witt 1967, Imhof 1972, Arnberger 1977). Both traditions basically took a comparable stance regarding choropleth maps:

-use them mostly/only for densities

-be aware of their pitfalls in comparison to area-class/dasymetric maps

-select the colors or shadings to make differentiation of the classes obvious.

If more than one data layer is even considered in the respective textbook, the introduction of diagram-symbols (e. g. pie- or bar-charts) is suggested.

ANOTHER QUANTITATIVE REVOLUTION

This all started to change with the rise of the microcomputer in applied cartography and the U.S. Census maps of Meyer et al. (1975). A large body of maps depicting two variables (such as farm produce income in four classes crossed with size in acres in four classes, Figure 1) by color for administrative units was produced with the help of computers from heretofore unseen masses of compiled data.

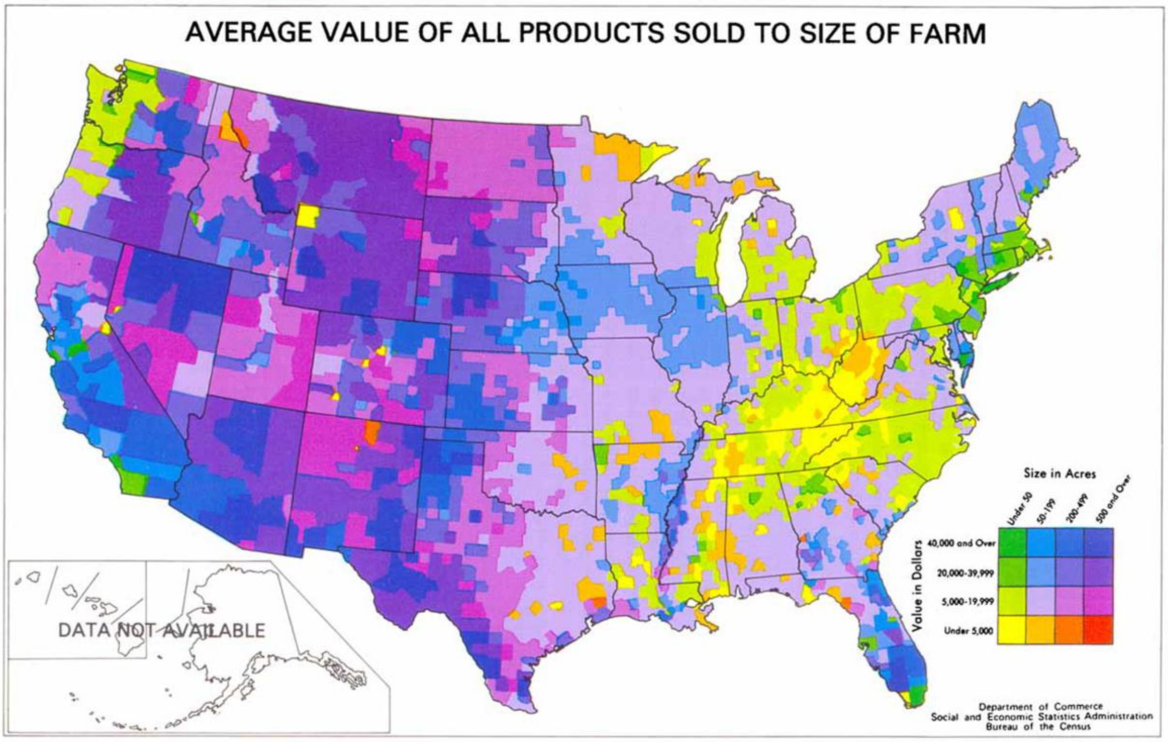

Figure 1: U.S. Census Map. Source: Olson 1981

While the maps spawned some controversy and empirical tests showed the inability of users to perceive overall distributions preattentively (Olson 1981), the idea itself was successful, especially in the US. Some skepticism remained, as can be seen from comments such as:

“Multi-subject mapping is warranted under only one other circumstance, namely when emphasis is on specific data – for each individual location or for a few locations close together-and general comprehension is not required.” (Fisher 1982).

Cuff and Mattson (1982) highlighted that only in the case of a handful of overlaps, were basically a new set of overlap classes would be created, should multivariate mapping be attempted. They were intrigued by the obvious and analytical nature of a bivariate color map legend, though, comparing it to the easily understandable scatterplot diagram. It were exactly these “scatterplot”-legends that became the focus of attention for the works following in the wake of statistician Bruce Trumbo (1981).

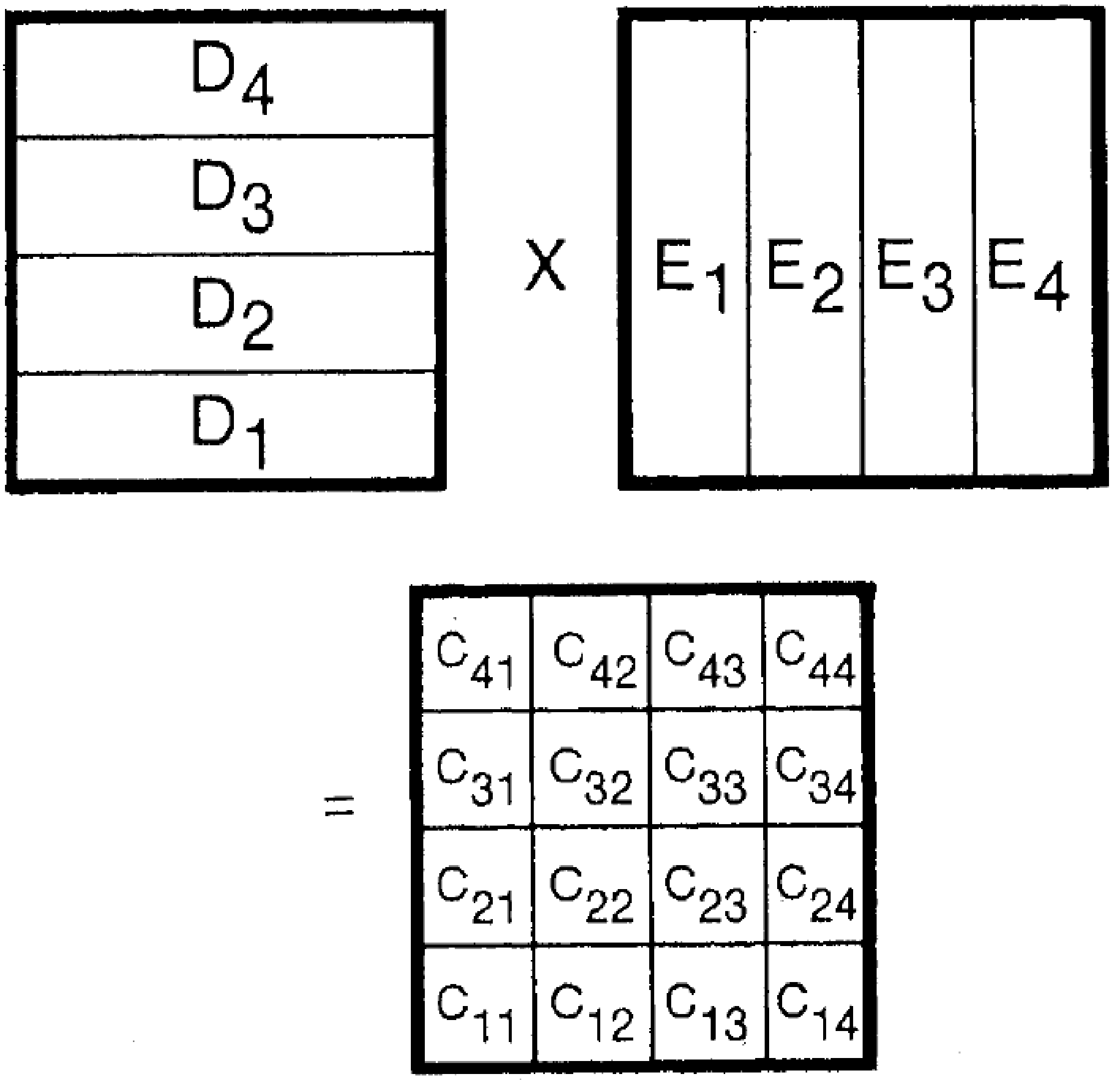

Figure 1. Crossing Two Retinal Variables To Form an Overlay Scheme

Figure 2: The basic idea of designing bivariate color schemes. Source: Trumbo 1981

The underlying idea of the numerous attempts (e. g. Eyton 1984, Robertson and O'Callaghan 1986, Rheingans 1999, Dunn 1989, Hagh-Shenas et al. 2007, Teuling et al. 2010) at the ideal “scatterplot”-legend for bivariate choropleth maps is that if color can be modelled by three numbers, as it has been since the rise of the microcomputer (RGB, LSH, CIE-lab etc.), these numbers can be used as new visual variables (Figure 2, Figure 3). This is a disconnect with cartographic tradition in general, Bertin in specific which only postulated the existence of 'Hue' and 'Value' in one form or the other. It continues to this day; as Kraak and Ormeling (2003, p. 101) point out, color is the main visual variable that has been extended upon by anglo-american cartography (e. g. MacEachren 1995, Tyner 2010 p. 137). Most prominent and de-facto standard for coloring choropleths is the approach by Cynthia Brewer (2005), which heartily includes strategies for bivariate color maps.

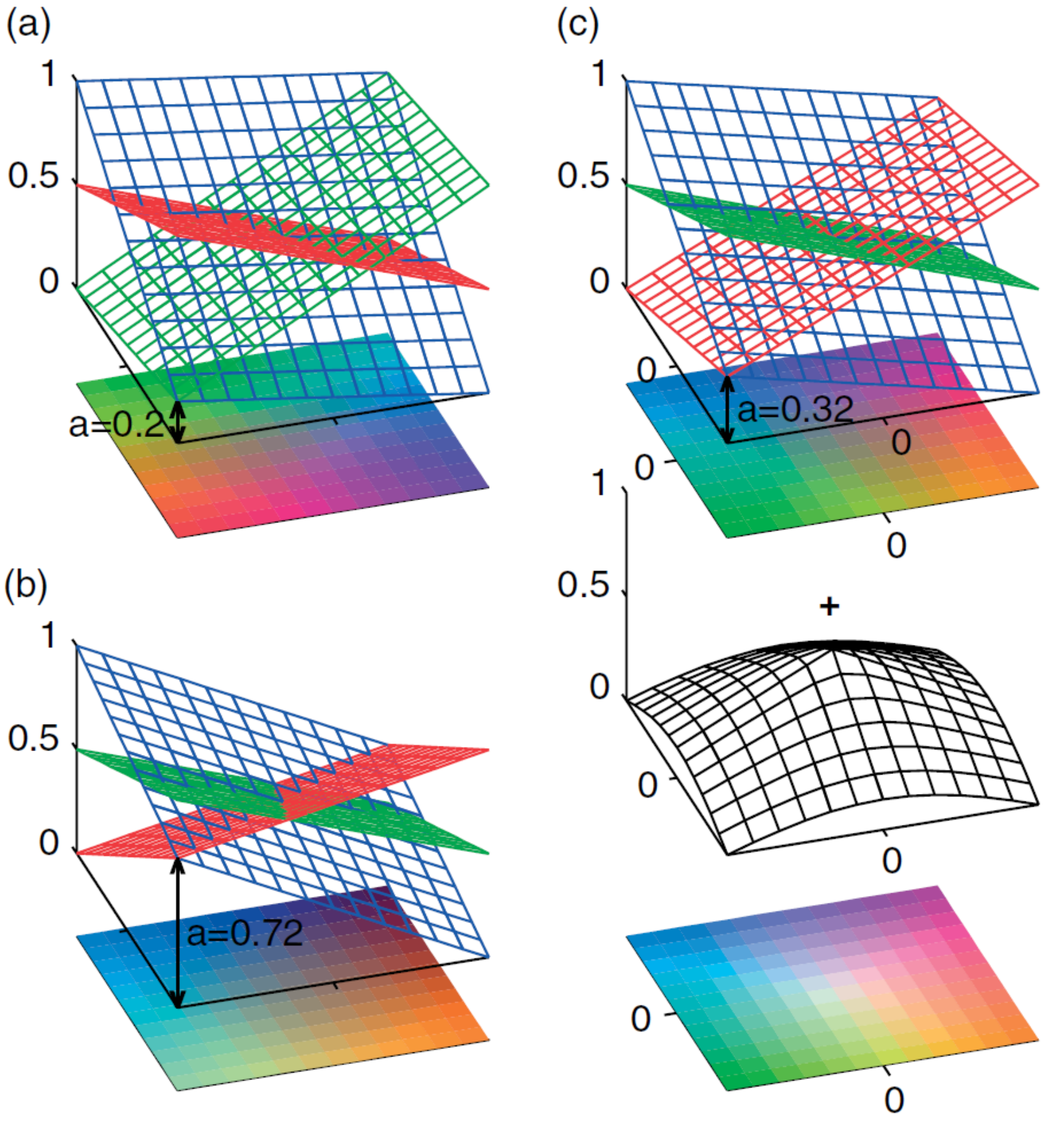

Figure 4. Construction of bivariate colour legends. (a and b) Examples for sequential colour legends. (b) is used in Figure 2. (c) Two-step construction of the diverging legend used in Figure 3. The parameter a controls the luminance gradient along the diagonal. Note that $a = 0.25$ results in equal luminance in all corners, but that at the same time $a \neq 0.5$ limits the effectiveness of the whitening kernel. This figure is available in colour online at www.interscience.wiley.com/ijoc

Figure 3: Analytical mapping of color space components to a bivariate color scheme. Source: Teuling et al. 2010.

All of these approaches can only hope to successfully work, if Bertin's concept of 'Disassociativity' of 'Value' is false, as nearly every one of them uses 'Value' or Lightness a visual variable that is crossed with some mixture of saturation and/or hue in the respective three dimensional color 'spaces'. This has been addressed not by a single one of those works, even those citing Bertin (1983) as reference and warrantor of sound strategy for visualization. In fact, the ongoing search for the ideal 'scatterplot'-legend for bivariate color maps, has produced contributions that question, if not disprove (empirically as in Olson 1981, Leonowicz 2006 or implicitly as in Ware 2009), the ability to construct such a color scheme while keeping preattentive selectivity for both variables.

SQUARING THE CIRCLE?

As we have noted, according to Bertin's contention of the 'Disassociativity' of 'Value', it is impossible to combine any variation of color ('Hue') with the variation of 'Value' and still have 'Selective' (preattentive) perception. While this contention has not been proven it also has neither been falsified. Generally speaking, it just has been ignored by those who attempt to build legible bivariate color maps. This mirrors a grander tendency to downplay or only superficially touch the visual variables and their expressiveness for scale levels of data in seminal works such as Dent (1993), MacEachren (1995), Robinson et al. (1995), Slocum et al. (2010) or Tyner (2010).

While we cannot prove Bertin's contention, we can point to some results from cognitive neuropsychology that provide strong pointers as to an explanation why especially 'Value' could be considered 'Dissacociative'. Dale Purves (see Purves et al. 2001) is the most prominent cognitive neuroscientist concerned with vision that postulates our visual perception operates in an empirical manner. This theory at its simplest, postulates that what we see is not mainly a function of the physical information encoded in the lightwaves that reach our retina. Rather, the image we see is already an interpreted rendition, a guess of our brain. Why would our brain have to guess what it sees? Again, the most simple argument is the well known problem of inverse evidence, for our case, the inverse optics problem (Wojtach 2009). Any given retinal image is a two dimensional projection that can have multiple explanations in the three dimensional world. A small object on our retina could be a) large and far away, it could be b) close and small, c) seen at an angle or d) any combination in between. It is the same with color vision, as a given received wavelength could originate from a colored light source, the texture of an object or any combination of the two as well as ambient lighting and shadows (Lotto et al. 2002a, Figure 4). Purves theory states that from the earliest childhood, our eye-brain system is trained by perception of the outside world and interaction with it. The brain collects data on encountered configurations of object size, texture color, shadows, lighting conditions and so forth. These datasets are 'stored' and if proven successful for the interaction and interpretation of the outside world, form the empirical basis for the guesses that our eye-brain system takes. As Purves' multiple studies indicate, his theory can explain many of the well-known optical illusions as being situations in which our eye-brain system takes a wrong guess (for our discussion the Cornsweet-effect being the most important). His theory can also be used to predict and construct new optical illusions (Lotto et al. 2002b). As the interaction with the three dimensional world are the formative experiences for our eye-brain system, a compelling argument can be made for 'Disassociativity' for cartographic applications: Human beings have experiences with 'Value' primarily as complex of shadows crossed with color, texture and ambient light (Figure 5), and with 'Size' as a complex visual cue to distance, view angle and object size. Both being the retinal variables that are most prone to the inverse optics problem and thus the ones in greatest need of empirical guesswork for our brain to make sense of. Some of the maps produced by Roth et al. (2010) give a good impression of lights and shadows that our brain tries to guess the right way, while distorting the message of the map-maker (Figure 6). In light of this theory, the failure to produce preattentively selectable bivariate color maps is the result of the constant attempt of our eye-brain system to assign three dimensional sense to a “senseless” situation on a 2d screen or map.

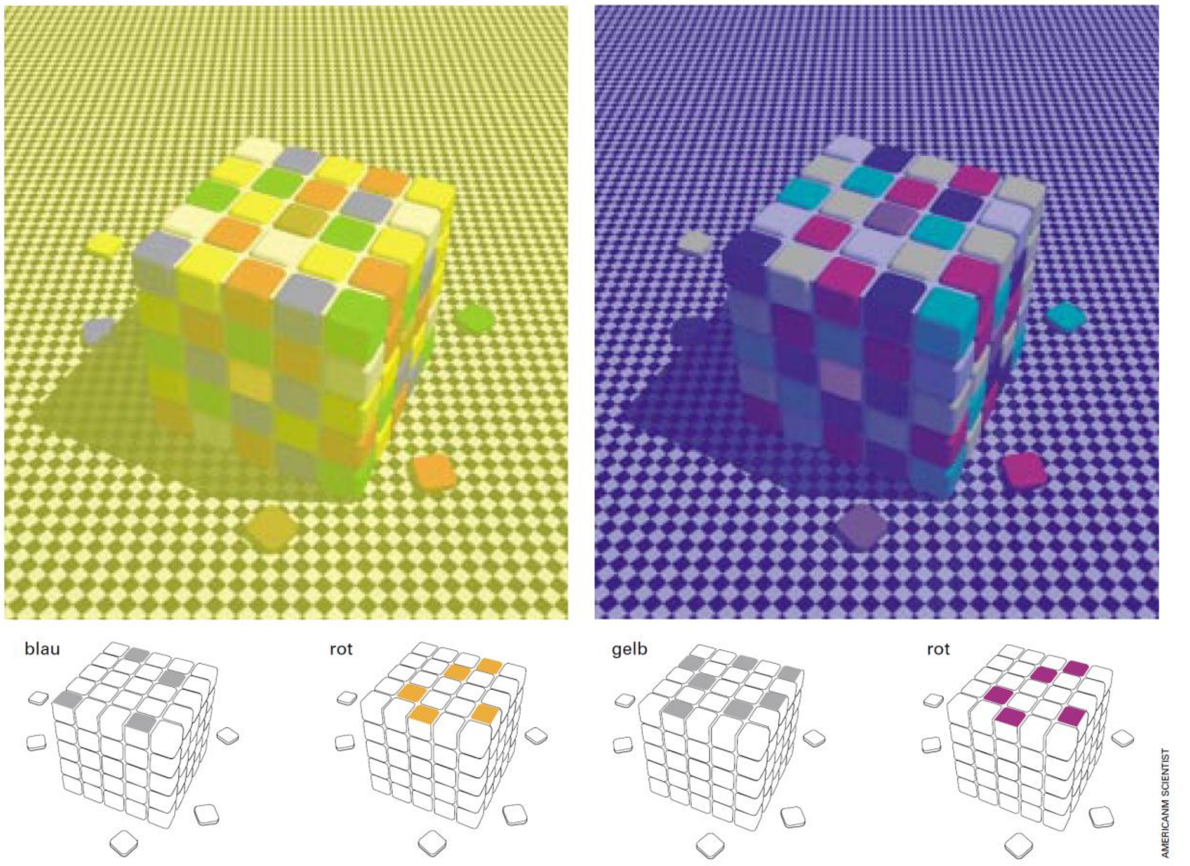

Figure 4: Color illusions; although we see both cubes as being similiarly covered, they are not. What is blue on the left cube is actually grey, the bright yellow fields on the right cube are grey, too. Source: Lotto et al. 2002b, image taken from the German 'Spektrum der Wissenschaft'-version of the same paper.

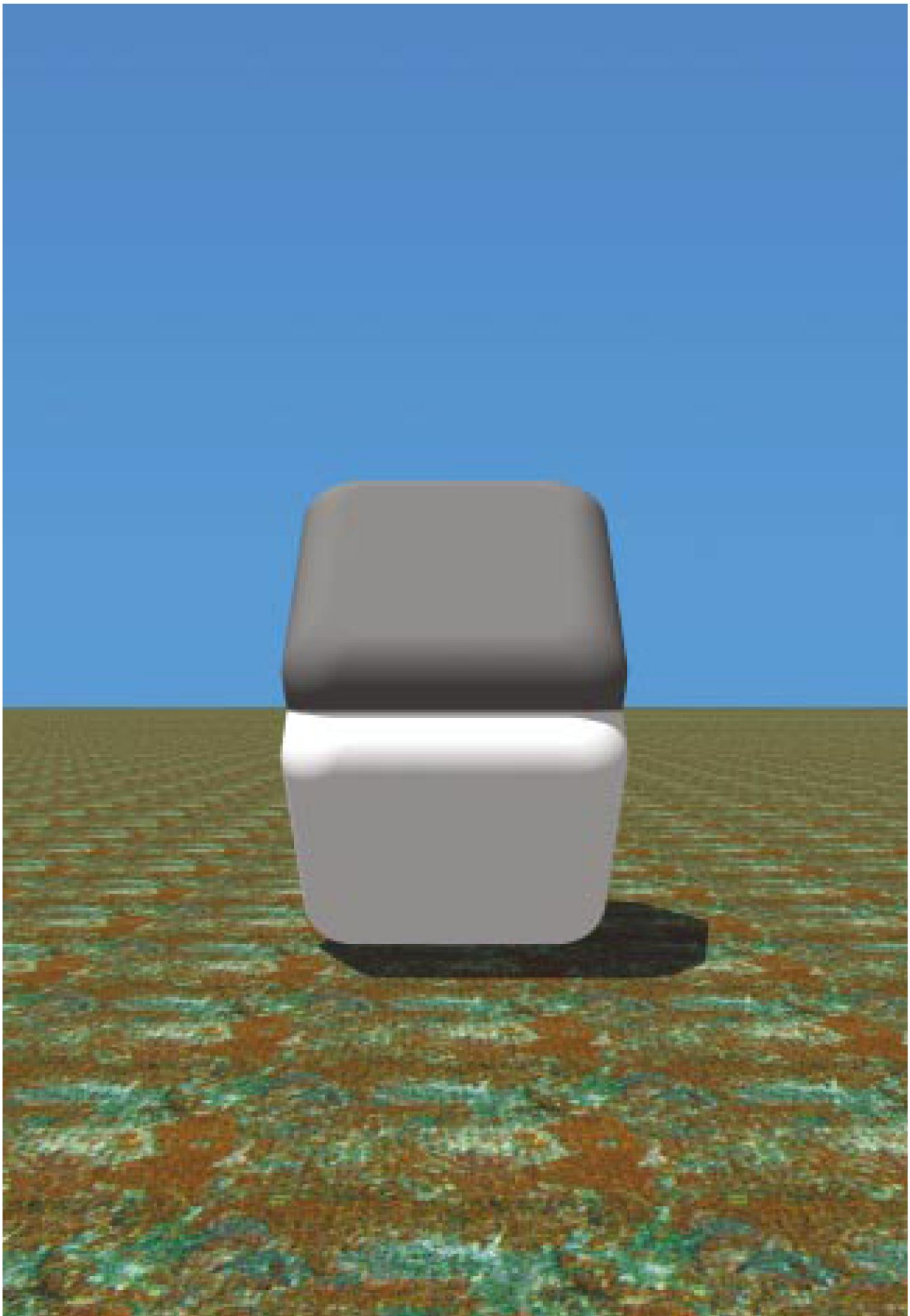
AMERICAN SCIENTIST

Figure 5: The Cornsweet-effect, amplified. the upper and the lower half of the object are actually the same shade of grey. Source: Lotto et al. 2002b

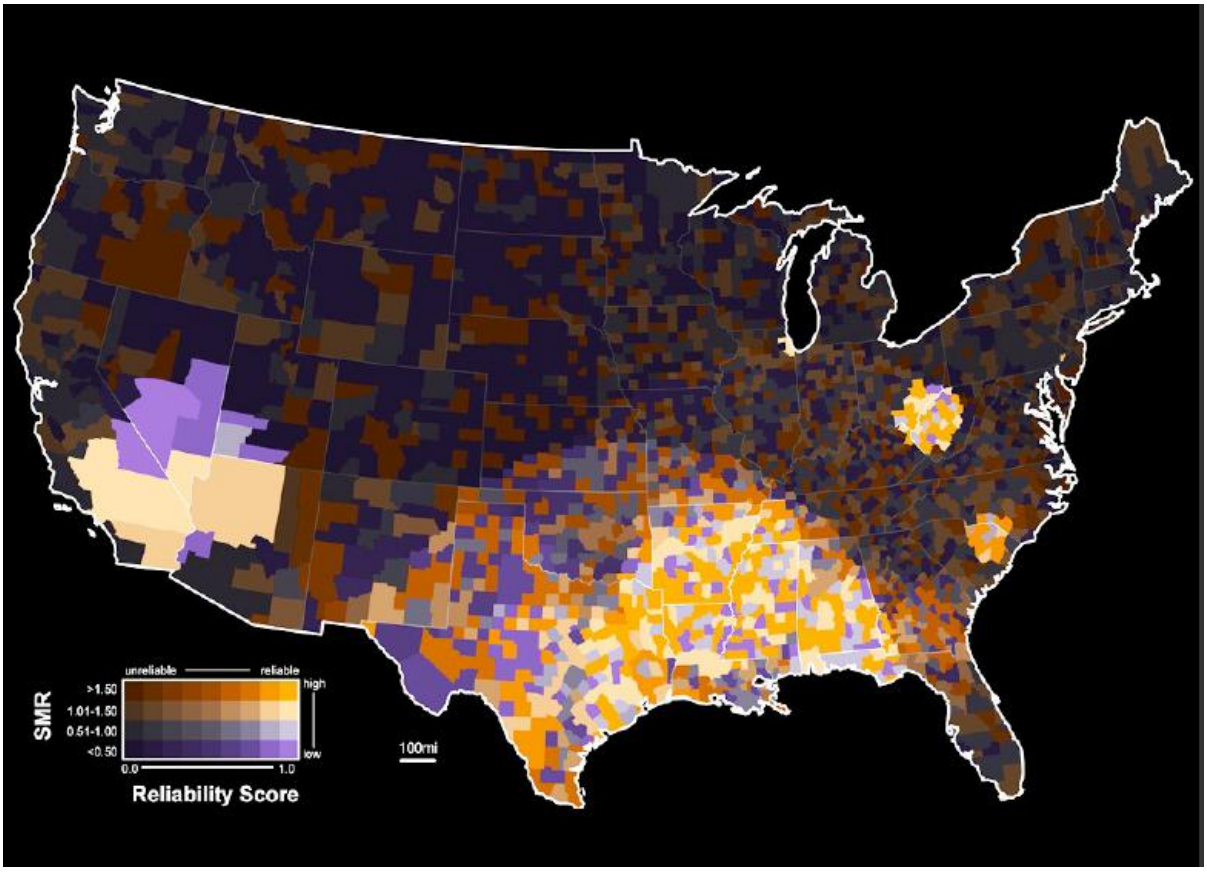

Figure 6: A quasi bivariate-color map. Note the difference in selectivity in the area that can be interpreted as being continually tinted by a spotlight and the non-continous light areas in comparison to the rest. Source: http://www.personal.psu.edu/rer198/presentations/WoodruffEtAl_2009_NACIS.pdf (NACIS 2009 presentation of the work documented in Roth et al. 2010)

CONCLUSION

It is beyond this contribution to give a definitive answer to the questions raised. It was our goal to highlight the fact that ignoring Bertinian 'Disassociativity' or postulating visual expressiveness for the dimensions of technical color spaces is just as contentious as Bertin's concept itself. In the light of Purves' theory of vision, the debate should move from assumptions toward proof or falsification of the results that are obtained when visual variables are mixed. At the very least, we hope to have brought forth some intriguing questions to consider before embarking on further quests for bivariate color schemes.

REFERENCES

Arnberger, E. (1977): Thematische Kartographie. Braunschweig.

Bertin, J. (1967): Sémiologie graphique. Paris.

Bertin, J. (1974): Graphische Semiologie. Diagramme Netze Karten. Berlin.

Bertin, J. (1983): Semiology of Graphics: Diagrams, Networks, Maps. Madison.

Brewer, C. (2005): Designing Better Maps: A Guide for GIS Users. ESRI Press.

Cuff, D. J.; Mattson, M. T. (1982): Thematic maps: their design and production. London.

Dent, B. D. (1993): Cartography. ThematicMapDesign, Third Edition. Dubuque.

Dickinson, G.C. (1963): Statistical Mapping and the Presentation of Statistics. Leeds.

Dunn, R. (1989): A Dynamic Approach to Two-Variable Color Mapping. In: The American Statistician, Vol. 43, No. 4.

Eyton, J. R. (1984): Complementary Color, Two-Variable Maps. In: Annals of the Association of American Geographers, 74 (3), pp. 477-490.

Fisher, H. T. (1982): Mapping Information: The Graphic Display of Quantitative Information. Abt. Associates.

Garlandini, S.; Fabrikant, S. I. (2009): Evaluating the Effectiveness and Efficiency of Visual Variables for Geographic Information Visualization. In: Hornsby et al. (Eds.): COSIT 2009, pp. 195-211.

Hagh-Shenas, H.; Kim, S.; Interrante, V.; Healey, C. (2007): Weaving Versus Blending: a quantitative assessment of the information carrying capacities of two alternative methods for conveying multivariate data with color. In: IEEE Transactions on Visualization and Computer Graphics, 13(6):1270–1277.

Imhof, E. (1972): Thematische Kartographie = Lehrbuch der Allgemeinen Geographie Vol. 10. Berlin.

Kraak, M.-J.; Ormeling, F. (2003): Cartography: Visualization of Geospatial Data, 2nd Edition. Essex.

Leonowicz, A. (2006): Two-variable choropleth maps as a useful tool for visualization of geographical relationship. In: Geografija 42, 1, pp. 33-37.

Lotto, R. B.; Purves, D. (2002a): A rationale for the structure of color space. In: Trends in Neurosciences, Vol. 25, No. 2.

Lotto, R. B.; Purves, D.; Nundy, S. (2002b): Why We See What We Do. A probabilistic strategy based on past experience explains the remarkable difference between what we see and physical reality. In: American Scientist, Vol. 90, No. 3, p. 236.

MacEachren, A. M. (1995): How Maps Work. Representation, Visualization and Design. New York.

Meyer, M. A.; Broome, F. R.; Schweitzer, R. H. (1975): Color Statistical Mapping by the U.S. Bureau of the Census. In: The American Cartographer, Vol. 2, No. 2, pp. 100-117.

Olson, J. M. (1981): Spectrally encoded two-variable maps. In: Annals of the Association of American Geographers, Vol. 71, No. 2.

Pomerantz, J. R.; Schwaitzberg, S. D. (1975): Grouping by proximity: Selective attention measures. In: Perception & Psychophysics, Vol. 18 (5). pp. 355-361.

Purves, D.; Lotto, R. B.; Williams, S. M.; Nundy, S.; Yang, Z. (2001): Why we see things the way we do: evidence for a wholly empirical strategy for vision. In: Philosophical Transaction of the Royal Society of London, 356, pp. 285-297.

Raisz, E. (1962): Principles of Cartography. New York.

Rheingans, P. (1999): Task-based Color Scale Design. In: Proc. Applied Image and Pattern Recognition, Oct. 1999.

Robertson, P. K.; O'Callaghan, J. F. (1986): The Generation of Color Sequences for Univariate and Bivariate Mapping. In: IEEE Computer Graphics and Applications, Vol. 6, No. 2, pp. 24-32.

Robinson, A. H.; Morrison, J. L.; Muehrcke, P. C.; Kimerling, A. J.; Guptill, S. C. (1995): Elements of Cartography, Sixth Edition. New York.

Roth, R. E.; Woodruff, A. W.; Johnson, Z. F. (2010): Value-by-alpha maps: An alternative technique to the cartogram. In: The Cartographic Journal, 47(2).

Slocum, T. A.; McMaster, R. B.; Kessler, F. C.; Howard, H. H. (2010): Thematic Cartography and Geovisualization. Third Edition. Upper Saddle River.

Spiess, E. (1970): Eigenschaften von Kombinationen graphischer Variablen. In: Grundsatzfragen der Kartographie, pp. 280-293, Wien.

Teuling, A. J.; Stöckli, R.; Seneviratne, S. I. (2010): Bivariate colour maps for visualizing climate data. In: International Journal of Climatology. DOI: 10.1002/joc.2153

Trumbo, B. E. (1981): A Theory for Coloring Bivariate Statistical Maps. In: The American Statistician, Vol. 35, No. 4.

Truran, H.C. (1975) A practical guide to statistical maps and diagrams. Portsmouth.

Tyner, J. A. (2010): Principles of Map Design. New York.

Ware, C. (2009): Quantitative Texton Sequences for Legible Bivariate Maps. In: IEEE Transactions on Visualization and Computer Graphics, Vol. 15, No. 6.

Witt, W. (1967): Thematische Kartographie – Methoden und Probleme, Tendenzen und Aufgaben. Hannover.

Wojtach, W. T. (2009): Reconsidering Perceptual Content. In: Philosophy of Science, 76, pp. 22-43.