Hearing Color: Radical Pluralistic Realism and SSDs

Zach Capalbo
Gordon College, Dept. of Computer Science
255 Grapevine Rd., Wenham, MA 01984
zach.capalbo@gordon.edu

Brian Glenney
Gordon College, Dept. of Philosophy
255 Grapevine Rd., Wenham, MA 01984
brian.glenney@gordon.edu

ABSTRACT

Pluralistic Realism argues that what makes color real is both its ecological significance for a particular species and the evolutionary adapted visual system which processes it. (Matthen 2005) We argue for a more radical position, that real color is dependent *only* on what a species determines as ecologically relevant, *regardless of visual system*. "Radical Pluralistic Realism" argues that real color content is processed by adaptive technologies such as Sensory Substitution Devices (SSDs)—devices that transform information accessible by one sensory modality , like vision, into another, like audition. We describe a color sonification SSD that translates colors into sounds and argue that subjects wearing this device actually hear color because they process real color content. even though they do not experience it as such,

Keywords: Philosophy of Color, Pluralistic Realism, Sensory Substitution Device (SSD), vOICe, Adaptive Technology, Color Perception, Enactive Perception

INTRODUCTION

African Grey Parrots (*Psittacus erithacus*) have a tetrachromat visual system with a fourth pigment sensitive to ultraviolet wavelengths (Bennett and Cuthill 1994). So, trichromat humans cannot be expected to share the same color categories as their feathered friends. For one, parrots perceive more colors. Two, this fourth band of wavelengths influences how all colors are perceived by parrots (Shepard 1997). For instance, it influences their categorization of colors at the opposite end of the spectrum, of red and orange colors, like the color of a ripe grapefruit (Pepperberg 1999). So, physiological differences suggest that color perception is better for parrots than humans and if not better, at least different. It is therefore prima facie unreasonable to think that the parrot and the human *both* see real colors.

It might also be argued that *ecological* differences—differences in how species respond to their environment with respect to different action potentials evolved out of their distinctive ecological niche—influence a species' color categories. (Matthen 2005) For instance, the color of a ripe grapefruit for a parrot might be the color of an unripe grapefruit for a human. So, the *real* color of a ripe grapefruit might be a point of irresolvable contention in a debate with parrot, for both physiological and ecological reasons. That is, unless *real* color is just a label for what's ecologically relevant to a species.

Mohan Matthen's (2005) "Pluralistic Realism" advocates this view. He claims that both humans and parrots have picked out the *real* color of a ripe grapefruit even if, in the end, both pick out a different color. "A general conception of *colour* and *colour experience* should allow us to treat the human and pigeon colours, and their systems, as instances of a general kind." (p. 164) In other words, for Matthen, the ecological significance of color and the species-specific processor underwrite the realness of color.

The central question we would like to raise in this paper is just how general the "color" kind is. Might the perception of real color include perception by *any* sensory system whose input was within the visible and near-visible spectrum? For instance, might color content be processed by an auditory system with color receptors? We argue yes, based on a device similar to that described above, a color sonification sensory substitution device (SSD) that processes color content using the auditory system. In other words, we argue that subjects wearing the color sonification SSD hear real color much like a parrot or a human sees real color.[1] We call this view "Radical Pluralistic Realism". What makes it radical is that, unlike Matthen's view, color content can be processed by a visual system that did not evolve out of processes relevant to an ecological niche. Adaptive technologies like SSDs are processors of real color for Radical Pluralistic Realists.

Radical Pluralistic Realism and Punctuated Equilibrium

Matthen argues that visual systems paired with a naturally evolved and thus ecologically embedded species are the only ones that process real color, a claim he calls, "Punctuated Equilibrium" (PE). "Punctuated Equilibrium implies that for most species now there is no mismatch between sensory capacities and innate sensory guided activities." (p. 206) PE forms a response to a problem for

AP-CAP 2009, October 1–2, 2009, Tokyo, Japan.

---

[1] We are not concerned with the experiential aspects of SSD use, but rather the kind of content being processed in the use of such devices. In having this specific concern, we are breaking away from an important theory of perception known as "intentionalism" (Byrne 2001), which claims that phenomenology supervenes on content.

Pluralistic Realism: that on such an account, visual systems are just different than others both between and within species: dogs and some humans are dichromatic and thus both possess different visual systems in comparison with trichromat visual systems of normal humans. But, most would identify differences of visual systems within the same species, like color blindness, as deficient, rather than different.

The deficiency claim gains strength as a problem for Radical Pluralism with the above-mentioned work of Shepard (1997) who demonstrated that the colors available to dichromats could not be the same as trichromats as they require comparisons with colors only available to trichromats. To avoid this within species deficit argument, PE grounds real colors in ecological contexts and so, provides Pluralistic Realism the conceptual space to claim that dichromatic humans *are* deficient, whereas dichromatic dogs are different.

Pluralistic Realism, however, bases the realness of color content on ecologically evolved action-potentials rather than evolved systems. This presents a kind of inconsistency if PE and Pluralistic Realism taken to be of a piece, particularly in cases where real color content is ecologically relevant but not processed by an evolutionarily-adapted visual system. For example Mirganka Sur (1990) rerouted ferret brain circuitry pathways that normally processed auditory information to process visual information. The auditory system was recruited to process visual content with little behavioral or processing change overall. This suggests that color content was in fact being processed by the ferret's auditory system in ecologically relevant ways without an innately given visual system.

Pluralistic Realism *sans* PE gains plausibility with the inverse claim, that if a visual system erred in obtaining ecologically relevant colors, then it would fail to acquire real colors, even if that system naturally evolved. So, whether or not a visual system evolved naturally is less relevant to whether it processes real color content, or at least makes much less a difference. It follows then, than a non-evolved visual system that processes ecologically relevant colors, like the color of a ripe grapefruit, processes real color.

Another less invasive icon of plasticity has been the development and use of sensory substitution devices. Like Sur's "Franken-ferrets", sonification SSDs take content that would normally be processed by the visual system and transforms it into sounds for processing by the auditory system. The question is, do these non-visual sensory systems process visual content? In other words, might one actually be able to hear color?

Radical Pluralistic Realism and SSDs

Since Radical Pluralistic Realism includes adaptive technologies into the space of visual processing systems that have real color content, it needs to account for how those using such devices might be said to possess real color content. We claim that users, whether blind, color blind or sighted but wearing blind goggles, do process real colors as long as the visual system does not revise the user's innately given action-potentials regarding these colors and thus a subject's color labeling. Whether subjects wearing color SSDs have the same color labeling is subject to empirical testing using a *color* sonification device. We include initial results using such a device below.

Additionally, another indication of SSDs processing real color content is if they function better than sonification SSDs that use mere luminance contrast as color SSDs will perform like an eye with cones, rather than an eye with just rods. For instance, the vOICe (Meijer 1992) uses rod-like luminance contrast rather than cone-like color as content and is by contrast 'less-visual' for trichromat humans than a device that does import color.

Though the vOICe is known to recruit parts of the visual cortex (Ptito 2005) and have some experiential effects (Auvray et. al. 2005), these results only vaguely suggest that its use is vision-like. For instance, though vOICe technology has long been cheap and accessible, there is no widespread use of it to aid visual deficits. One likely reason for this is that what SSDs provide is not sensory substution, but rather sensory augmentation (Auvray & Myin 2009). But another is that since color content is not being processed, it is not processing real human visual content, but rather something monochromat species would expect. One way to test whether real color content is being processed in a color sonification SSD is to show that it works better on trichromats than luminance contrast sonification SSDs like the vOICe.[2]

To show better functioning of color SSDs, our paper compares a color sonification SSD that produces a "sound locus" that correlates with a color locus as its output. A device that provides a sound locus has the potential of providing much of the spatial information made available by the vOICe, such as boundary detection, with the addition of providing color information.[3] Additionally, using color input makes discrimination possible in normal lit environments, as opposed to luminance contrast devices which function best in low lit and dark rooms with luminescent stimuli. In addition, such devices often confuse luminance changes within a color with object boundaries, something that a color SSD is not as susceptible to.

We describe a set of preliminary experiments that compare the performance of the vOICe and the KromoPhone. The KromoPhone outperforms the vOICe in our search tasks, discrimination tasks, and wandering tasks with blindfolded normal subjects.

---

[2] A nice control for this experiment would be to test the inverse hypothesis: that monochromatic humans perform better at object identification tasks using the vOICe over a color SSD. No such experiment has been conducted as of yet.

[3] Another major difference of the vOICe is that it presents the entire visual field through a moving "soundscape".

Color Sonification with the KromoPhone

The KromoPhone translates color or properties of color intensity of a pixel into auditory frequencies. The input of our color SSD is a pixel source, localized either by a mouse cursor in images produced by a monitor or by the center of the image being processed from a head camera. The user is also fitted with a remote control, a “Wii remote”, which enables the user to remotely change the input from a single pixel to an average up to 80 square pixels around the cursor or center of the screen. There are three modes that allow the user to detect color, each of which can be toggled using the remote. The RGB mode provides the whole color wheel with three dominant sounds, the HSL mode correlates sounds with color properties rather than colors themselves. And the default RGBYW mode adds two further colors that are difficult to distinguish between in the RGB mode.

The RGB mode

The KromoPhone in RGB mode takes the intensity of a pixel into three main auditory frequencies. It scales the intensity of the red component of the pixel into a floating point value between 0 and 1. This value is then multiplied by a simple sinusoidal function of a frequency of 15.9 kHz to the red channel. This value is then multiplied by 1 on the right stereo channel and 0 on the left stereo channel. The same is done to the green channel, except the frequency is 8.4 kHz and 0.5 is multiplied by both channels, giving the sense of the sounds being heard in the middle. The same is done to blue, but with a frequency of 5.4 kHz, and 0 being sent to the right channel and 1 times the value being sent to the left channel. The sound is constant until the pixel intensity changes.

The HSL mode

The user also has the option of switching to HSL (Hue, Saturation, and Luminosity) mode, where hue is mapped to pitch, saturation is mapped to panning, and luminosity is mapped to volume. Only those with ‘perfect pitch’ perform well using the HSL mode as it requires a high grade of pitch discrimination to detect colors.

The RGBYW mode

An RGBYW mode is the default mode for the device. White and yellow are extracted from the RGB selection. Each color is translated into a distinct sinusoidal function of a distinct pitch, pan, and tone. The added tone axis helps the user to discriminate between the greater number of sound types. The intensity of each color in the pixel is mapped to the volume of the sound it produces. White is a high-pitched sound between left and right. Red is a high pitch in the right ear with a trumpet-like tone (Vergez 1997). Yellow is a high pitch in the left ear with a tone like a ukulele. Green is a medium pitch in the right ear with a violin-like tone. Blue also has a trumpet-like tone but is a low pitch in the left ear.

Figure 1: RGBYW Mode Mappingss

The figure indicates the volume of each different sound based on each color’s hue. The 5 colors fade in and out, as shown in the figure, based on the hue of the color.

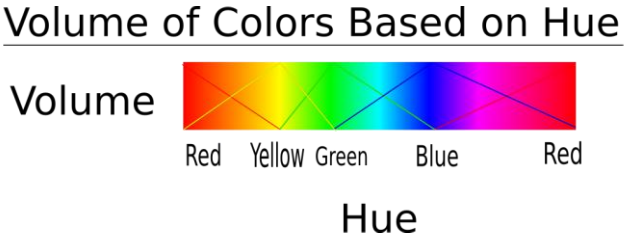
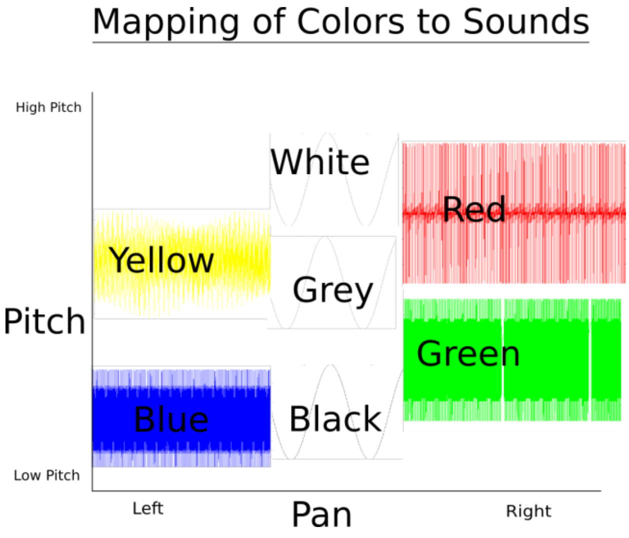


Figure 2: RGBYW Mode Sounds

The figure indicates the distinctive pitch, pan, and tone matched to each color.



The Scan Setting

Lastly, a soundlocus “scan” can be toggled on the remote that emulates the side-to-side head movement which user’s naturally develop when using the device for movement around obstacles. On the scan mode, subjects hear the soundlocus crossing the screen back and forth at one to eight scans per second, a speed remotely set by the user.

Preliminary Experimentation

One question is whether the Kromophone can inform subjects of the correct color labels of objects that would be expected by normal sighted humans. A second question is whether it outperforms a sonification SSD without color information. So, after a very minimal training period involving about five minutes to hear the sounds as they correlate with the color wheel, subjects were asked to perform two tasks: to pick certain fruits and vegetables known to correlate with certain colors and to perform a search task for luminescent disks in a small area.

Preliminary Experiment 1: Fruit and Vegetable Picture Discrimination Task

Using the RGB mode of the Kromofone, subjects were able to distinguish between different kinds of fruits and vegetables pictured on a masked screen. Furthermore, musically trained subjects performed better at the fruit recognition task than non-musically trained subjects suggesting that errors were due to the matching the sounds to colors correctly. The best musically trained subject correctly identified 5 out of 6 fruits or vegetables, whereas our best non-musically trained subject identified only 3 out of 6 fruits or vegetables. In comparison, none of the three subjects trained using the vOICe could identify any of the fruit, either by the shape contours or luminance differences.

A total of 6 subjects attempted the fruit and vegetable identification task using the RGB setting. Together they made a total of 40 attempts at determining the fruits and vegetables. In total, 18 were identified correctly. 3 of the subjects identified the fruit with 50% accuracy or better. The best result was 5 out of 6 fruits correctly identified, while the worst result was 0 out of 5 fruits correctly identified. When asked for the an explanation of the poor results, subjects indicated a misidentification of the sounds with the colors rather than misidentifying the fruits when the colors were known.

Preliminary Experiment 2: Localization and Luminance Task

Three naive subjects, completely untrained in SSd devices, wore the KromoPhone and vOICe in a search task. (Reynolds and Glenney 2009) Subjects were asked to find three light sources placed randomly placed in a circle of 15 chairs approximately 5 meters in diameter. When using the KromoPhone on the RBG mode, users had equal times as when using the vOICe in a dark room, averaging three minutes per trial. However, as the ambient overhead lighting increased to a low lit 100 Lux (so low reading a book is near impossible), vOICe users were unable to find the light sources whereas KromoPhone users remained at approximately three minutes. Finding light sources using the KromoPhone in a fully lit room was much more difficult, but still possible for one subject, while impossible for anyone when using the vOICe.

Preliminary Experiment 3: Use of the KromoPhone in an urban setting.

One subject submitted to approx. four hours of training on the KromoPhone in RBGWY mode in the natural setting of a college campus. For instance, they were trained through trial, error, and encouragement on how to follow a path and to avoid obstacles like garbage cans and trees. We then allowed the subject to wander free for two ten-minute sessions.

In the first session the subject was able to avoid all objects in a 300 sq. m. courtyard, including trees, a concession stand, a garbage can, and a bike rack.

In the second session, the subject was able to negotiate his way through a portion of Harvard square, a place teeming with people and obstacles, including steps, fire hydrants, and various bulkheads. In one case, the subject was surprised to discover that a fire hydrant was blue rather than red. The subject was also able to differentiate the sounds of vehicles on a road nearby from the sounds of the device, and thereby avoided walking onto the road.

Discussion

A primary function of color perception in humans is object recognition, enabling quick identification of natural objects such as flowers and fruit. Non-natural objects, such as signs and stoplights, are also encoded with colors for ease and speed of identification. While shape discrimination is of more import to mammal eyes (Land 2003) color discrimination has great significance. Supplanting color discrimination for blinds and color blinds is thus a significant achievement. Additionally, most sonification SSDs base their input on luminance variations, making stimuli with little luminance variation difficult (Auvray 2007). Hence, using color as input provides a means for compensating for low luminance contrast settings that other devices cannot handle. Lastly, the sounds correlated with colors are distinguishable by pitch and thus provide subjects with a good ear near immediate use of the device for color discrimination tasks.

The preliminary studies above indicate that a color sonification SSD will outperform a luminance contrast sonification SSD. Since color is a uniquely visual content, this suggests that real color content is being transformed and processed by the KromoPhone. Furthermore, KromoPhone use provided subjects the ability to label various objects with color and identify them on this basis. Such identification suggests that there were no changes to what subjects took to be ecologically relevant. For instance, it was still expected by our trained subject that a fire hydrant would be the color red, as they were surprised by its blue sound appearance. In sum, preliminary experiments using a color sonification SSD suggest that it processes real color content and thus its users hear *real* color.

CONCLUSION

One of Matthen's arguments in support of Pluralistic Realism, what might be dubbed the "clinical" argument, is that only Pluralistic Realism can provide real grounds for experimentation with color perception in other species. "*Colour* is not 'our' term: we can make it so only at the exorbitant cost of abandoning the comparative (i.e. cross-species) study of colour vision." (p. 163) If parrot colors are not real, for instance, then there is no object of study for those researching color vision in parrots. In the same manner, Radical Pluralistic Realism extends color research into the sector of those working on color using adaptive technologies.

Barry Mound has argued against the clinical argument, claiming that there are "pragmatic" alternatives to our color understanding that need not be called "realist", and yet provide grounds for a clinical understanding of color. "All that Matthen has provided a case for, at best, is that we should have an additional concept to the anthropocentric one." (p. 32) In response, we suggest that Matthen's account of color *is* a pragmatic alternative that doubles as a metaphysical account of the nature of color and thereby offers those experimenting on color perception a practical and robust research program, a robustness that can be extended into research on adaptive technologies if a Radical Pluralistic Realism is accepted.

Color SSDs and Realism's Problem of the Criterion

Little has been said so far in support of color realism. We end with a problem for color realism, the problem of the criterion for what would make color real, and a suggested empirical response using our color sonification SSD.

To be a realist does not entail that the conditions of correct color perception are fulfilled by human visual systems. As Hardin argued (1998), if to be a color is to be a reflectance property ordered by wavelength comparisons, then only perception which orders colors by an objectively given color scale counts as color perception. But because humans and other species like parrots see violet as more similar to red than green when in fact green is closer to violet on the color scale, humans and other species do not see violet, or any color, correctly.

Hardin's inadequacy claim can be made more radical (Hardin 2003)—no color perception is correct *in principle* because there is no 'in principle' standard of correctness available; there are any number of facts about color that might count as a criterion for correctness, but no way to determine *which* criterion gets color right. So, any realist theory of color need not only wield a correctness criterion, but a criterion for what makes that criterion the correct one. The challenge for the realist is to find a criterion for correctness that falls out of the nature of color perception and thereby providing a self-evident criterion.

We conclude this paper by suggesting an experiment using the KromoPhone to provide grounds for color realism. The experiment is grounded on the following scenario: if a subject were put into an environment of novel color perception, whatever criterion that made that environment one involving color perception would be in this literal sense a self-evident criterion. This self-evident criterion would then serve as a normative basis of correctness for color perception.

This further experiment using the KromoPhone situates subjects in a novel color perception experience as suggested above. Like Auvray's (2005) experiment, subjects are only told to wear a set of headphones and a backpack while blindfolded and that the device will produce a unique array of sounds. Subjects would be 'visually' presented with an assortment of color patterns as stimuli, but without knowledge that these are stimuli. In addition, a control group will be given the same device but with preset sounds that do not correlate with the colors presented to the camera.

We expect, following Auvray's (2005) results, that subjects will infer a correlation between the sounds and their environment and further that some subjects will specifically infer that the sounds are correlated with colors. If these results hold, subject questionnaire reports may then be evaluated to determine the basis for their inferences of color perception. We hypothesize, given our initial tests, that color permanence will be what suggested to the subjects this correlation, making color permanence a normative criterion for color realism, much like object permanence is for object realism. (Piaget 1969)

ACKNOWLEDGMENTS

We would like to thank Gordon College for an internal grant to support our research. We also thank Zachary Reynolds, David Botticello, and Sara Hendron for help with our experiments and Irene Pepperberg for helpful suggestions on parrot color perception.

REFERENCES

Aurvray, M., Myin, E. *Perception with compensatory devices: From sensory substitution to sensorimotor extention*, Cognitive Science, 2009: forthcoming

Auvray, M., Hanneton, S., & O'Regan, J. K. *Learning to perceive with a visuo-auditory substitution system: Localization and object recognition with The Voice*. Perception, 2007 36, 416-430.

Auvray, M. Hanneton, S., O'Regan, J. K. *There is something out there: Distal Attribution in Sensory Substitution, Twenty Years Later*: Journal of Integrative Neuroscience 2005 4 4: 505-521

Bennett, A.T.D., Cuthill, I.C. *Ultraviolet vision in birds: What is its function*? Vision Research, 1994 34: 1471-1478

Bynre, A. *Intentionalism Defended*; The Philosophical Review, 110 no. 2 April 2001: 199-240

Hardin C.L., *A Spectral Reflectance Doth Not a Color Make*: Journal of Philosophy 100 2003: 191-200

Hardin C.L., Color for Philosophers: Unweaving the Rainbow; Hacket: Indianapolis, 1988

Land, M.F., Nilsson, D-E Animal Eyes; Oxford University Press: New York, 2002

Matthen, M Seeing, Doing, and Knowing; Oxford University Press: New York, 2005

Meijer PBL, An experimental system for auditory image representations, *IEEE T Bio-med Eng* **39**:112–121, 1992.

Maund, B. *The Philosophy of Color* Stanford Encyclopedia of Philosophy; 2006 edition, URL = http://plato.stanford.edu/entries/color/

Pepperberg, I.M. Alex Studies: Cognitive and Communicative Abilities of Grey Parrots, Harvard University Press: Cambridge MA, 1999

Piaget, J., Inhelder, B The Psychology of the Child: Basic Books: New York; 1969

Ptito, M., Moesgaard, S.M., Gjedde, A., Kupers, R. *Cross-modal plasticity revealed by electrotactile stimulation of the tongue in the congenitally blind*, Brian 2005; 128 (pt 3): 606-14

Reynolds. Z., Glenney, B.R. *Enactive Training for Sensory Substitution Devices* AP-CAP 2009

Shepard, R.N. *The Perceptual Organization of Colors: An adaptation to Regularities of the Terrestrial World?"* Readings on Color Vol. 2: The Science of Color, MIT, Cambridge, MA 1997: 311-56

Sur, M., Pallas, S.L, Roe, A.W. *Cross-modal plasticity in cortical development: differentiation and specification of sensory neocortex* Trends in Neurosciences, 13 6,1990: 227- 233

Vergez, C., Rodet, X. *Model of the Trumpet Functioning: Real Time Simulation and Experiments with an Artificial Mouth*: Proceedings of the 1997 International Symposium of Musical Acoustics. St. Albans: Institute of Acoustics, pp. 425-432