



HAL
open science

Qualia and spandrels: an engineering perspective

Jean-Louis Dessalles

► **To cite this version:**

Jean-Louis Dessalles. Qualia and spandrels: an engineering perspective. Technical Report ENST 2001-D-012, pp.350, 2001. hal-00614805

HAL Id: hal-00614805

<https://hal.science/hal-00614805v1>

Submitted on 22 Aug 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Qualia and Spandrels: an Engineering Perspective¹

Jean-Louis Dessalles

Ecole Nationale Supérieure des Télécommunications

46 rue Barrault - 75013 Paris - France

dessalles@enst.fr - <http://www.enst.fr/~jld>

If we consider that cognitive abilities evolved to perform sophisticated processing of sensorial inputs and allow for better adaptive behavioural responses, then it seems that there is no functional need for the qualitative aspects of subjective experience. Qualia have been shown to be dispensable when causal explanations of behaviour are requested. They look very much like epiphenomena. Though qualia are private experiences, the structure of qualitative spaces for various sensory modalities can be investigated. Surprisingly, what such studies reveal is inconsistent with epiphenomenality. Qualitative spaces offer a familiar look to the eye of engineers. There is some evidence that qualia are *optimally organised*, in each modality, for functional discrimination. It is thus reasonable to hypothesise that such a design was produced by natural selection. If so, we must admit that qualia play a causal role in cognitive processes, even if such a causality still lies beyond our understanding.

Keywords: qualia, phenomenal consciousness, evolution, epiphenomenon, qualitative spaces.

1. Introduction

1. Qualitative properties are the most immediate aspect of our conscious experience. Familiar examples are pain, pleasure, redness, the taste of red wine. Phenomenal consciousness² is a basic component of our perceptual and proprioceptual experience. As such, it can be considered as a biological property of our species. We may attribute to some non-human species a similar ability to experience smell, colour or pain, even if their experience may be qualitatively different from ours. An obvious question then arises: why did phenomenal consciousness emerge in the phylogenetic history of these species?

2. Such a question may seem alternatively pointless or hopeless, depending on how it is understood. It is pointless if phenomenal consciousness, as a scientific object, is on principle

¹ This research would not have existed if Tiziana Zalla had not introduced me to the problematics of consciousness. I thank her and Georges Rey for their fruitful remarks and critiques on a first draft of this paper. I also thank Rafael Núñez for his friendly and helpful comments. I thank Joëlle Proust, Jose Luis Bermudez and Thomas Polger for their numerous suggestions that helped me improve the quality of this paper.

² Following (Block 1995), we need to distinguish here *phenomenal consciousness* from other cognitive processes, from self-consciousness and from higher-order conscious states. The feeling of being a single entity, the fact that some recalled events look familiar, the feeling of “ownership” about our mental states, the first-person point of view, the ability to observe aspects of our cognitive functioning are other important features of what is called “consciousness”. Nevertheless, all of them might be different aspects of consciousness, each one might be related to different functions and may eventually call for different accounts (Zalla 1996).

reduced to structural or functional properties of the brain. We assume the traditional view that there is something like experiencing pain or redness, and that such qualitative experience is worth scientific explanation. We use the term *qualia* to refer to qualitative aspects of experience, being aware that the cognitive nature of qualia is highly problematic. The main claim of this paper should be easily transposed in any non-eliminativist theoretical framework about consciousness³.

3. Accounting for the very existence of qualia may seem rightfully hopeless, at least in a foreseeable future. It has often been stressed that we have no idea why qualia are what they are, why they are simply *felt* (Chalmers 1995; Harnad 2001). The problem addressed in the present paper is much less ambitious. We take phenomenal experience as granted, and acknowledge the fact that its very existence remains mysterious. Given that, we question the possibility that phenomenal experience be considered epiphenomenal.

2. Are qualia mere evolutionary epiphenomena ?

4. Can we deduce from the observation of phenomenal experience some of the reasons why it does exist as a biological endowment of our species? Or can we consistently accept the claim that phenomenal experience be a biological epiphenomenon? The usual way to understand why a given biological feature was selected in the phylogenetic past is to consider its function, and to examine how this function could be advantageous for those individuals endowed with it. Unfortunately, phenomenal consciousness seems to fulfill no obvious function. The fact that it plays a definite role in perceptive processes or in the production of behaviour, in itself, is questionable. Furthermore, what such a role would consist in remains mysterious, so that it seems pointless to speak of its adaptive value.

5. Several authors mention strong arguments supporting the idea that conscious experience may be an *epiphenomenon* (Jackson 1982, Flanagan, 1992, Flanagan & Polger 1995). It has often been claimed that our current knowledge of the laws of nature did not allocate any definite role to consciousness: “the view that qualia are epiphenomenal is a perfectly possible one” (Jackson 1982). Flanagan demonstrates that “conscious inessentialism” is consistent:

“[Conscious inessentialism is] the view that for any intelligent activity *i*, performed in any cognitive domain *d*, even if *we* do *i* with conscious accompaniments, *i* can in principle be done without these conscious accompaniments” (Flanagan, 1992)

6. Since there is no absolute way of telling the difference between sentient and non-sentient beings from the observation of behavioural performance, the epiphenomenality of qualitative experience cannot be ruled out, and we are left with an evolutionary puzzle:

“Telling a convincing story about the adaptive advantage of consciousness is very hard. The upshot is that there exist no good stories for why consciousness evolved in this actual world. There are as yet no credible stories about why subjects of experience emerged, why they might have won –or should have been expected to win – an evolutionary battle against very intelligent zombie-like information sensitive organisms.” (Flanagan & Polger 1995)

³ Some authors consider that presupposing the existence of qualia is already a faulty attitude (Dennett 1988). We will come back to this issue in the discussion, section 9.

7. If we can imagine that fully functional creatures may exist and behave as we do without having any conscious experiences at all, how can any serious claim be made about the biological role of phenomenal consciousness? Qualia, these qualitative aspects of experience, the ‘What it is like’ character of mental states (Nagel 1974), would appear as mere superfluous and fortuitous adjuncts to brain activity. Even if conscious inessentialism does not rule out a possible role for qualia, it makes epiphenomenalism appear as a perfectly sound doctrine.

8. Most of the complex processes going on in our body are indeed achieved without involving any conscious component. We are not conscious of our immune system, we do not feel each contraction of our stomach, we are not aware of maintaining our equilibrium at each moment. Many of our cerebral processes are performed without us being aware of them. The way we analyse a visual scene, the way we recognize words in a complex acoustic signal, the way we adapt our walk on an uneven ground are good examples of such processes. If we think that complex processes like shape and object recognition are performed unconsciously, we may wonder why perception involves consciousness at all. Why are we sentient beings, why are we not unconscious, robot-like creatures? What use is the ability we have to experience internal mental states or events in the outer world?

9. From a Darwinian perspective, the issue becomes even more enigmatic: if phenomenal consciousness plays no causal role in neural processes, how and why was it retained by natural selection?

10. The question of the very existence of qualia is one to the most difficult issues addressed by contemporary philosophy and cognitive sciences. It is a form of mind-body problem. Understanding the reason why we are feeling anything at all belongs to the so-called ‘hard problem’ (Chalmers 1995). Such questions have been discussed at length by many authoritative authors. What is proposed in this paper is to offer the external view of an engineer on one limited aspect of the problem, the epiphenomenality of qualia.

11. In what follows, I will first consider the possibility that qualitative aspects of experience are evolutionary epiphenomena. Then, in order to evaluate such a hypothesis, I will consider the way qualia are structured within a given sensory modality and their relation to sensory inputs. I will give evidence suggesting that this structure is, from an engineering perspective, locally ‘optimal’. The next step will be to claim that such local optimality is not here by chance, but must be ascribed to natural selection. The conclusion will be that qualia epiphenomenality should be considered unlikely: if there was a selective feed-back that enabled the emergence of locally optimal qualia, these qualia must be efficient in some way. The argument is rather indirect, though, and the questions of *how* mental qualities can be efficient and even *why* they exist at all will remain out of the scope of this development.

3. Epiphenomena and evolution

12. Evolutionary epiphenomena are known in biology. Features which were not selected for, but result from the selection of other characteristics, are evolutionary epiphenomena. A trivial example is the colour of our internal organs, which is a mere property of organ tissues,

but which was certainly not selected for its aesthetic value. Such epiphenomenal characteristics have been named ‘spandrels’⁴.

13. Any organ or behavioural disposition that is recognized as the direct product of natural selection can be shown to be locally ‘optimal’. For example, any minor change in eye architecture would result in a less efficient device for its typical owner. Such optimality is local. Major changes that would lead to better performance are not offered to natural selection. For instance, evolutionary biologists often mention the fact that the neural wiring of photosensitive cells would be better located behind the retina, as it is in mollusk eyes, rather than in front of it, as is the case in vertebrate eyes. But such a new design cannot be reached within a small mutational step, and intermediary mutants would have a very poor vision indeed. Gould and Lewontin rightfully denounce ‘panglossism’, the doctrine of Voltaire’s character Pangloss for whom every feature of the world is optimal (Gould & Lewontin 1979). Global optimality is not a sound concept of biology. Natural selection can only achieve local optimality. The vertebrate eye is not the best possible organ for vision, but it is locally optimal, and this local optimality can be directly assessed.

14. A common misunderstanding about evolution is that it is a slow process. As a consequence of this supposed slowness, all biological forms would be on the way of their improvement, and none of them would, even locally, be optimal. The fallacy comes from the fact that the judgement is made at the wrong time-scale. The mechanisms of evolution that make it rapid (as illustrated by bacterial adaptation to antibiotics) have been better understood since they are reproduced by engineers in an optimisation method called Genetic Algorithms (Holland 1975; Dessalles 1996). Genetic recombination and selection give rise to what has been called intrinsic (or implicit) parallelism (Holland 1975). Evolutionary change in presence of selection pressure operates in parallel in the population and rapidly reaches equilibrium, characterised by a local optimum (Dessalles 1996). Subsequent changes may be slow: they suppose a transition from the current local optimum to another, giving rise to a series of punctuated equilibriums (Gould & Eldredge 1977). We must retain that any sub-optimal biological set of features is rapidly improved until it reaches a local optimum.

15. Gould and Lewontin also denounce pan-selectionism, a doctrine claiming that all features of living beings have been selected as such. There are important exceptions: neutral characteristics and spandrels. For obvious reasons, neutral characteristics lack local optimality. Spandrels do have some local evolutionary necessity. For instance, the colour of our internal organs is, in some sense, locally necessary. Any change that would produce a new coloration would result from a change in the histological structure of organs. If, as we may suppose, this structure is the result of natural selection, such a modification will lead to sub-optimal organs. The local evolutionary necessity of a spandrel thus results from the local evolutionary necessity of other attributes, to which it owes its existence. In other words, the local evolutionary necessity of a spandrel is inherited from what *causes* it, whereas the local evolutionary necessity of a functional feature is due to its indirect *effects* on survival (or, ultimately, on gene propagation).

⁴ The term refers to a metaphor used by Gould and Lewontin (1979): ‘the spandrels of San Marco’. Spandrels are the triangular arch structures that support the upper domed roof of the cathedral of San Marco in Venice. These structures, which are integrated in the global decoration design as if they had been part of it in the first place, happen to unavoidably result from the presence of other architectural structures.

16. The crucial difference between spandrels and functional features is that the local evolutionary necessity of spandrels is only indirect. Since they do not have reliable effects on survival, their structure cannot show any optimality *per se*⁵. The green colour of bile is not locally necessary *per se*. We cannot account for it unless we consider the function of bile itself and the local optimality of its chemical composition.

17. Similarly, if we consider that consciousness is a mere epiphenomenal property accompanying some neural mechanisms, it must be an evolutionary spandrel, a fortuitous by-product of brain phylogenetic evolution. As Jackson puts it:

“The point is that all we can extract from Darwin’s theory is that we should expect any evolved characteristic to be *either* conducive to survival *or* a by-product of one that is so conducive. The epiphenomenalist holds that qualia fall into the latter category. They are a by-product of certain brain processes that are highly conducive to survival.” (Jackson 1982)

18. As such, qualia should not show any feature that could be directly identified as optimal. I will now suggest that this is not the case, and that qualia do exhibit an optimal arrangement. As Nichols and Grantham put it, “the phenomenal consciousness system exhibits a level of anatomical complexity that plausibly requires an adaptation explanation” (Nichols & Grantham 2000).

4. The structure of qualitative spaces

19. For each sensory modality, there is a qualitative space: qualitative experiences are more or less similar, depending on input stimuli. As I will suggest, this fact is highly unexpected from an epiphenomenalist perspective. Consider for example colour perception. The organization of colour experience for a given subject is well represented in a three-dimensional space (Munsell 1905; Clark 1993). This fact, in itself, is opaque to introspection⁶. It can be proved using psychological techniques. This finding is not deduced from physiological considerations, like the fact that we have three types of photosensitive cells in the retina⁷. The structure of qualitative colour space can be inferred from subjective judgments alone. From a set of similarity judgments made during controlled experiments, like ‘this colour is closer to this one than to that one’, a colour map can be deduced using

⁵ To show how the local evolutionary optimality of a functional feature and a spandrel differ, consider a characteristic E , for instance the aerodynamic efficiency of the wing of a bird, as a function of several relevant parameters a_i (e.g. the length of the wing, its curvature, etc.) An aerodynamicist could verify that actual birds show values $\{a_i^*\}$ that maximize E : $\partial E / \partial a_i(a_1^*, a_2^*, \dots) = 0$. Now, consider a spandrel s (e.g. the distance between the end of the wing and a red spot on the neck of the bird). We can write $s = S(a_i, b)$ where the parameter b determines the position of the spot. The value of s at equilibrium is constrained: $s^* = S(a_i^*, b^*)$, where b^* is determined by independent evolutionary constraints on b . In this sense, s has, indirectly, a local evolutionary necessity. However, s^* is not itself the solution of the aerodynamic problem. If the flight efficiency is rewritten as $E_f(s, a_1, a_2, \dots)$, then no definite value s^* corresponding to the maximum of E_f can be determined, neither by the aerodynamicist nor by natural selection. As b may vary across generations without altering flight efficiency, s may vary without affecting E_f . We thus have $\partial E_f / \partial s = 0$ for any s . The spandrel s is not a reliable parameter of flight efficiency.

⁶ When colour TV was introduced, people needed some explanation to believe that all the variety of colours could be generated using only three basic colour components.

⁷ There are actually four kinds of photosensitive cells, but the central part of retina contains mainly three types of cells, the so-called cones.

techniques like multidimensional scaling⁸. Such experiments give results that are compatible with a three-dimensional colour space (Shepard 1962; Clark 1993). The distribution of colours in this space can be represented as a solid shape. Intensity is one axis of this solid. For a given intensity value, we get a disk-like ‘slice’ of the solid. Colours corresponding to this intensity value are distributed over the disk: gray near to the center and saturated colours at the periphery, arranged in the same order as in the rainbow.

20. There is more to be said about this colour space. In each disk-shaped slice, which corresponds to a given intensity value, the blue-yellow axis and the green-red axis constitute two roughly orthogonal diameters⁹. Moreover, not all colours are equivalent. For each subject, there seems to be four ‘unique’ colours, located in the green, yellow, blue and red areas. All other hues are experienced as mixtures of two neighbouring unique hues. There is, however, a significant variance among subjects about unique colours that leads some authors to question the scientific value of the concept (Saunders & van Brakel 1997).

21. The existence of a subjective space for colour perception, characterized by a 3D-structure, is not an exception. Other modalities also give rise to qualitative spaces. Qualitative taste can be approximated by a tetrahedral space, with four basic tastes at the vertices (Clark 1993). Sound intensity and pitch are distributed along one-dimensional spaces, whereas judgements about timbre require four dimensions to be only approximated (Bonnet *et al.* 1989). In what follows, I suggest that this organization of qualia in definite spaces is not an anecdotal curiosity, but must result from a specific requirement.

5. Why are there qualitative spaces ? An engineering hypothesis

22. The existence of qualitative spaces, although their global structure is opaque to introspection, is not at odds with intuition. Colours, smells, sounds can be compared from one stimulus to the next. This statement is obvious, so obvious that we fail to notice that it could be otherwise. If qualia were genuine epiphenomena, it should be possible that they form no identifiable structure. It would even be possible that qualia in a given sensory modality cannot be compared. This would be a logical extension of what thought experiments about inverted spectra suggest. According to such ‘experiments’, qualitative experience may be distorted from one individual to the next without producing any behavioural difference:

“Though you and I have exactly the same functional organization, the sensation that you have when you look at red things is phenomenally the same as the sensation that I have when I look at green things. If this hypothesis is true, then there is a mental state of you that is functionally identical to a mental state of me, even though the two states are qualitatively or phenomenally different. So the functional characterizations of mental states fail to capture their ‘qualitative’ aspect” (N. Block. 1980)

23. What such experiments show is that there is no way to prove that, when looking at a cherry, a person A has the same qualitative experience as another person B. To go beyond the

⁸ For instance, by applying multidimensional scaling on a table giving inter-city distances, one can reconstruct a map with the location of all towns. This map may be rotated or inverted, as if seen in a mirror.

⁹ It should be stressed that the colour space of a given subject is only approximately structured this way. What we call ‘disk’ here refers to an approximately round shape. Some colour-blind individuals lack one or two dimensions in their perception of colour.

usual assumptions of such experiments, which merely postulate simple permutations or global inversions of qualia, why not suppose that colours experienced by A correspond to a complex permutation of B's subjective colours, or that A and B's experiences are radically different? Such a discrepancy in personal experiences would remain unnoticed, since both A and B would use the word 'red' to qualify cherries. If we believe that qualia can be manipulated this way without any predictable consequence, however, then the very existence of finely structured qualitative spaces becomes a wonder. Conversely, if we consider the structure of qualitative spaces, as it can be deduced from psychological experiments, from an engineering perspective, we must conclude that most qualia permutations, during thought experiments, are not admissible. Moreover, the structure of qualitative spaces makes sense if these spaces are part of a definite functional design. Several observations will lead us to these conclusions.

5.1 Continuous projection

24. The most obvious feature of qualitative spaces is that the stimulus-qualia projection is continuous. Two close visible wavelengths will be experienced as similar hues. In our perception of the rainbow¹⁰, there is a smooth transition between colours. We could think of another picture: the rainbow, which results from a continuously varying physical parameter, could be experienced as kind of jumble: similar colours would occur at different locations; neighbouring colours would be experienced as strongly dissimilar; there would be no order consistent with the topology of our qualitative space. In our actual experience of rainbows, though, yellow and red are separated by all distinguishable orange hues, the space between blue and green is like greenish blue and bluish green. By independent means, for instance technical devices, we are able to determine a stimulus space structure and to control the stimulus presented to subjects. When the stimulus is modified in a continuous way, as in the rainbow, subjects' experiences vary continuously.

25. This observation is a genuine psychological result. Qualitative spaces, in each modality, are structured in a definite way, as shown in great detail by Clark. Stimulus spaces also have their own structure. The fact that both structures correspond through a continuous relation is not a trivial phenomenon. In some cases perception, *e.g.* phonological perception, is consciously experienced as discontinuous (Harnad 1987), but the amplitude of discontinuities remains small¹¹.

26. The same seems to hold for all sensory modalities, even if it is not equally easy to demonstrate. We expect that molecules with similar chemical structures are likely to project onto neighbour points of taste qualitative space, as anyone can experience with various sugars (monosaccharides and disaccharides) or various types of salt. Continuous projections between stimuli and subjective experience are more apparent for sound intensity, sound pitch, warmth or pressure exerted on skin. For instance, sound intensity subjective perception can be approximated by a continuous function of acoustic pressure (Bonnet *et al.* 1989). The existence of continuous stimulus-qualia projections is a first fact that restricts what subjective experience can be like.

¹⁰ Actual rainbows offer less than seven discernible hues. We use the term rainbow to refer to a complete set of monochromatic stimuli at a given intensity.

¹¹ If p represents the stimulus-qualia projection, then p can be considered continuous. In the relaxed version, p must be such as $|p(s+h) - p(s)| < C$ for any stimulus s . The constant C measures the degree of continuity, and h measures the granularity at which qualia are considered

5.2 Directionality

27. The stimulus-qualia projection has another important property: it may be directional. This means that for some situations like the rainbow, qualia are not duplicated and appear to be totally ordered: we do not experience a U-turn when the stimulus increases. In other terms, the stimulus-qualia projection is monotonic¹². When we look at a rainbow, the colours seem well ordered and are not repeated. When the wavelength increases, we do not experience yellow twice. The same seems to hold for other modalities. Increased pressure or increased heat transfer on our skin generate new experiences¹³.

5.3 Regularity

28. The projection of stimulus space on qualitative space in a given sensory modality may have another important property. When we look at a rainbow, we experience several colours and a smooth transition between them. Moreover, these colours are gracefully distributed along the spectrum. This, also, is an unexpected phenomenon. We could easily imagine that one colour, some pure red for instance, occupies 90% of the visible spectrum, while all other hues are concentrated on a 10% of the available wavelength range. Colours seem, instead, to share the subjective colour spectrum in such a way that they are, more or less, evenly spread out over it¹⁴.

5.4 Dynamic range

29. Qualitative spaces offer a variety of vivid and rich experiences. Though it may be difficult to quantify precisely, the dynamic range can be defined intrinsically as the ratio between the greatest difference in the qualitative space and the smallest discernible difference between two qualia. In the case of colour space, we may try to estimate the dynamic range by counting the maximum number of discernible hues between two colours. Some 150 colours can be discerned in monochromatic waves; other dimensions of colour offer smaller dynamic range: one can distinguish some 25 saturation levels in the blue, and some 30 intensity levels between black and white (Sève 1996). Different qualitative spaces may differ significantly in their dynamic range. Sound intensity, for example, is perceived on a subjective scale with a dynamic range of 2000, while pitch is experienced with a maximum dynamic range of 3000

¹² For some one-dimensional stimulus sets (e.g. pure wavelengths at a given intensity), corresponding experiences are structured as an ordered, one-dimensional space as well. Moreover, for any two non-equal stimuli s_0 and s_1 , the sign of $(p(s_1) - p(s_0))/(s_1 - s_0)$ does not change. If the stimulus-qualia projection p can be considered as derivable, then directionality is captured by the condition that $p'(s)$ does not change sign, where s stands for the current value of the stimulus.

¹³ Directionality does not prevent the qualitative space to loop on itself. The last colour of the rainbow is close to the first one; some individuals have a brief experience of very high water as if it were cold (before pain is felt). For such situations, we may consider the relevant qualitative space as periodic, *i.e.* for any qualia q , $q+A = q$, where A stands for the total range of the relevant qualia, and consider directionality as a local requirement, namely that $(p(s+h) - p(s))/h$ does not change sign for some granularity h .

¹⁴ The requirement that qualia are evenly distributed over a given stimulus range can be captured, in the case of a one-dimension stimulus space, by a condition on the first derivative of the stimulus-qualia projection: $|p'(s)| < K$, where the constant K measures the regularity of the distribution. This condition prevents experience to change too rapidly when stimuli changes are small. If p cannot be considered as derivable, the condition must be replaced by $|(p(s+h) - p(s))/h| < K$, where h is the granularity at which qualia are considered.

(Bonnet *et al.* 1989). Notice that contrary to colour spaces, these ranges are quite far from uniform, loud sounds or high tones being much less distinguishable. Our point is that all qualitative spaces possess a significant dynamic range.

30. These four properties: continuity, directionality, regularity and dynamic range are not, as I will suggest now, fortuitous features. In order to clarify this point, we must make a small digression.

5.5 Evidence from communication theory

31. Communication theory considers optimal projections. In digital communications, each elementary event is the transmission of a symbol taken from an alphabet. This alphabet may be as small as $\{0,1\}$, but it may also contain a significant number of possible symbols (more than 1000 for a typical telephonic modem). Symbols are abstract mathematical objects, but when it comes to actual transmission, they must be converted into physical signals. Let us assume that such signals belong to a two-dimensional space: the receiver will for instance measure two tensions v_1 and v_2 on two independent wires¹⁵. If we want to transmit with a two-symbol alphabet (like the Morse code or any binary code), then the corresponding signals should be chosen on two opposite corners of the available space: for instance a null signal for 0, a maximal value on both wires for 1. If we want to transmit messages based on a four-symbol alphabet, then the signals corresponding to the symbols should be placed on the four corners of the available signal space. If the number of symbols is large, the corresponding signals will be located at the vertices of a hexagonal lattice, each signal being surrounded by six neighbours (figure 1).

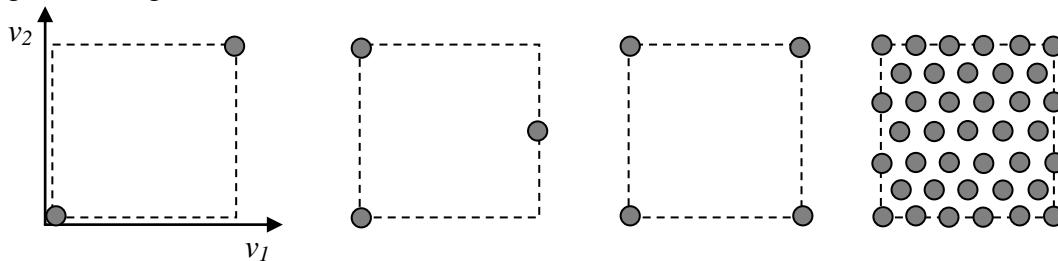


Figure 1: Examples of optimal arrangements of signals in the available signal space.

32. Why are such solutions optimal? The main problem in telecommunication is noise. Received signals are not identical to emitted signals. The problem is to avoid errors when recovering the abstract message from its physical realization. If signals representing symbols are maximally separated from each other, then the effect of noise on error rate is minimal. Hence the hexagonal pattern when many points have to be optimally spread out on a 2D-space. I will suggest that the structure of qualitative spaces is constrained by a similar requirement.

33. Qualitative spaces possess an intrinsic topology, however. They do not consist in a set of non-related points. This may also be the case in digital communication, when there is a topology in the set of symbols, *i.e.* when symbols are more or less similar to each other. The consequence of errors may be significantly different, depending on whether the erroneous

¹⁵ A more sensible example would be two parameters of a modulated carrier wave, as in quadratic amplitude modulation.

symbol is close to the correct one or far from it. If we consider the coding of music, for instance, a small shift in the signal space should not lead to distant values of the acoustic signal. If this constraint is respected, most actual errors will remain unnoticed. This has a consequence on the choice of the symbol distribution over the available signal space. An obvious solution is to place signals associated to neighboring symbols close to each other. As a consequence, the projection from signals to symbols is continuous. This is also a property found in the case of qualitative spaces.

34. In order to make the analogy with the structure of qualitative spaces more concrete, imagine we have to design a technical device to detect visible wavelengths, using a given number of sensors $\{S_1, S_2, \dots, S_n\}$ connected to a set of light-emitting diodes $\{d_1, d_2, \dots, d_n\}$. Let us assume that this latter set has an intrinsic topology: d_8 is closer to d_7 and d_9 than it is to d_1 or d_{10} . For instance, S_i may control the switching of the i th element of a row of light-emitting diodes (figure 2a). The best choice, when it comes to tuning the detecting elements S_i , is to divide the spectrum into n equivalent parts and have each d_i responding to the i th part of the spectrum. The solution results from a single requirement: *keep the average effect of errors as small as possible*. This requirement leads to four constraints: (1) continuous (topology preserving) projection; (2) connex receptor fields; (3) regular distribution; and (4) significant dynamic range. Thanks to (1), the topology of the diode row can be used to keep errors harmless. When, because of a small error, d_7 goes on instead of d_8 , the consequences may be almost ignored, considering that d_7 and d_8 are neighbours. If the projection were discontinuous (figure 2b), there would be no way to distinguish between small and large errors. Since the former are by far more frequent, a continuous projection, by reducing their effect, is much preferable¹⁶. The role of (2) is, very basically, to avoid indeterminacy. If d_4 can be triggered by two non-neighbour sensors (figure 2c), the performance of the system is affected by a fixed error rate. The main effect of (3) is to lower error probability. If S_7 responded to a major part of the spectral range while other detecting elements were crammed on the remaining part (figure 2d), then it becomes much more likely that noise affecting one of the crowded sensors will cause an error. The effect is rather dramatic: starting from an error probability of 10^{-3} when sensors are regularly spaced, we get an error probability as high as 0.38 when all but one of the sensors share 10% of the available range¹⁷. The effect of (4) is to lower the effect of quantification (figure 2e). A finite number of sensors on a continuum introduce a quantification noise, due to the imprecision of readings on the diodes. If the dynamic range is divided by 2, the quantification noise is multiplied by 4 or, equivalently, is increased by 12 dB¹⁸. Notice that the effect of (3) is also to avoid the unnecessary quantification noise due to the imprecision of S_7 (figure 2d).

¹⁶ For n sensors, the error amplitude average, for a small noise, is multiplied by $n/3$ when we go from a continuous projection to a random one.

¹⁷ In presence of a gaussian noise, the error probability is approximated by $2Q(A/(n\sigma))$, where A is the amplitude equally shared among sensors, σ is the noise standard deviation, and $Q(\alpha)$ is a rapidly decreasing function representing the probability for a gaussian variable to be higher than α .

¹⁸ The amplitude of the quantification noise, for uniformly distributed signals, amounts to $A^2/(4D^2)$, where A is the signal range and D is the dynamic range.

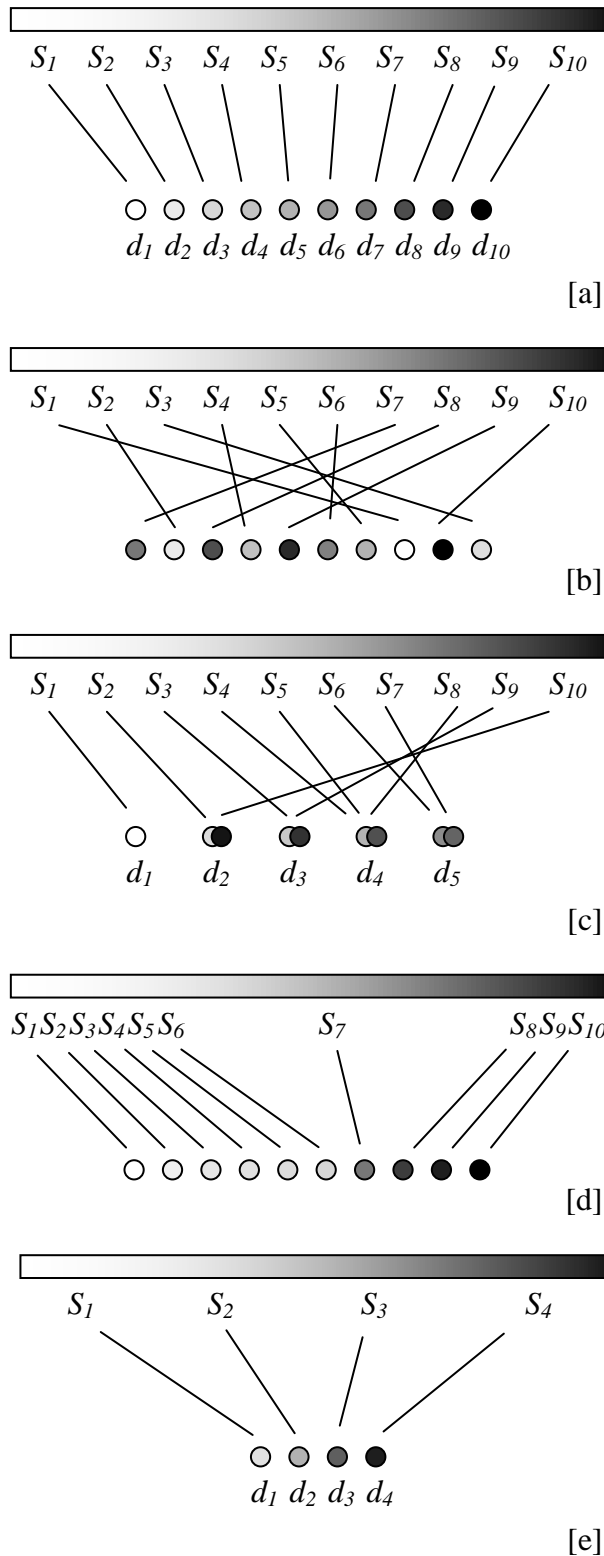


Figure 2: Illustration of [a] a continuous (topology preserving) and regular projection, [b] a discontinuous projection, [c] a continuous, though non connex projection, [d] an irregular projection and [e] a low dynamic range projection from a set of sensors S_i to a display consisting of a row of light-emitting diodes d_i .

35. Our analogy allows us to understand the role of the four properties of the stimulus-quale projection emphasized above: continuity, directionality, regularity and dynamic range. All four properties make perfect sense if they are involved in a discrimination process. The

role of continuity (*e.g.* the fact that colours follow each other quite smoothly on the rainbow) is to limit the effect of errors; the effect of directionality, in one dimension (*e.g.* the fact that colours occur only once on the rainbow), is to ensure the connexity of the projection (the fact that two disconnected stimulus ranges do not project on the same quale, introducing a systematic error source). The role of regularity (*e.g.* the fact that no single colour occupies 90% of the rainbow) is to minimise error probability and quantification noise. High dynamic range (*e.g.* the number of discernible colours in the rainbow) keeps the quantification noise low.

36. In the context of communication theory, such properties unavoidably suggest that *qualia are involved in a discrimination process*. Before drawing some consequences of this observation, we must make two remarks. First, we should not always expect strictly equal shares from the regularity of the stimulus-quale projection. If some parts of the stimulus space are more relevant than others, we expect more efficient discrimination in these zones to be correlated with a higher density of qualia¹⁹. If we follow the engineering guideline, what is important is that the practical consequences of errors are kept low. The correlation between stimulus relevance and quale variation should be, in principle, empirically testable.

37. The second remark concerns the inevitable distortion introduced in the stimulus-quale projection. The light spectrum emitted by a punctual source belongs to a space of infinite dimension. Our perception projects it onto a three-dimensional colour space. Such a projection cannot be made without distortion. Hence the fact, called metamerism, that many different stimuli are experienced identically. For instance, a superposition of two wavelengths corresponding to red and green will be experienced as an orange hue, as would be a single, intermediary wavelength. What we called a regular distribution of detecting fields can be represented as an optimal 'folding' of the target space when projected back into the stimulus space, sometimes called 'snake in the box' by theoreticians of error correcting codes.

6. Constraints on the inverted spectrum experiment

38. The fact that qualitative spaces may be triggered through a continuous, directional and regular projection, explain why their manipulation during inverted spectrum thought experiments is so delicate, sometimes believed to be practically impossible (Shoemaker 1975, Clark 1996a, Palmer 1999). It is indeed difficult to depict a scenario where experiences of red and green are inverted but all structural properties between colour experiences remain identical. For instance, you can't simply change a single quale. Suppose two persons A and B have the same colour experience, except for one quale, triggered when seeing a ripe cherry.

¹⁹ Suppose that a relevance pattern is available, *i.e.* the relevance (between 0 and 1) of each stimulus zone is given by $R(s)$. R can be seen as a distortion of the stimulus space: $dt = R(s) ds$. The new stimulus t varies more rapidly in relevant zones, when $R(s)$ is high. We have $t = T(s)$, where T is the cumulated relevance:

$$T(s) = \int_{s_0}^s R(u) du$$

The projection p is expected to be regular on the distorted stimulus t :

$$\left| \frac{dp}{dt} \right| = \left| \frac{dp}{ds} \times \frac{ds}{dt} \right| < K$$

We get the new constraint: $|p'(s)| < K R(s)$ for each $s = T^{-1}(t)$, which limits the variation of qualia $q = p(s)$ according to stimulus relevance (the lower the local relevance, the lower the quale variation).

Suppose that $red_A \neq red_B$, while $C_A = C_B$ for any other colour name C (colour names associated with suffixes stand for qualia). Then the stimulus-qualia projection cannot be continuous for both subjects, and at least one of them should notice it. For instance, when a single wavelength stimulus goes from, say, 625 nm to 675 nm (orange to red), one of the subjects will experience an abrupt transition that normal subjects do not experience. This effect of quale inversion limits the scope of the claim that qualia of different subjects may be changed at will in thought experiments with no observable consequence.

39. We should not yet conclude that all human beings must have identical qualitative experiences when presented with the same stimuli. There are indeed several ways to save part of the inverted spectrum claim. We could imagine that while A and B share part of their colour experience, red_B is a quale unknown to A, and that all intermediary colours, for instance between $yellow_B$ and red_B , are also completely different from what A has ever experienced. I do not see how such a claim can be refuted. However, it requires half of the qualitative space to be transformed into a brand new qualitative dimension. There is little ground to suppose that two roughly similar brains could generate completely different sets of qualia with similar structural properties.

40. A more plausible claim would be that A and B may have globally identical qualia sets, but that the stimulus-qualia projection, though continuous for both subjects, is nevertheless different. We must keep in mind that the two projections must respect continuity, directionality and regularity requirements, which can be experimentally checked and verified for each subject. As a consequence, the hypothesised mapping from A's qualitative space to B's cannot be a jumble. It must be continuous²⁰, directional²¹ and regular²². If we deal with the same set of qualia for A and B, only approximations of global shifts or global inversions

²⁰ Technically, if we note p_A the projection from the stimulus space on A's qualitative space, and M the hypothesised mapping between A's and B's qualitative spaces, then $p_B = M \circ p_A$. Since p_A and p_B are continuous, M must be continuous too, at least on all qualia in A's qualitative space reached by p_A .

²¹ Using the notations of preceding notes, the directionality of M results from the directionality of p_A and p_B . Let's consider two qualia q_0 and q_1 in A's qualitative space. We suppose $q_0 < q_1$. The corresponding qualia in B's qualitative space are consistently ordered (for the sake of simplicity, we ignore the case of periodic qualitative space and consider that both projections are derivable):

$$M(q_1) - M(q_0) = \int_{q_0}^{q_1} M'(q) dq = \int_{s_0}^{s_1} M' \circ p_A(s) p_A'(s) ds = \int_{s_0}^{s_1} p_B'(s) ds = p_B(s_1) - p_B(s_0)$$

Here, s_0 and s_1 are some stimuli that trigger experiences q_0 and q_1 respectively for A. Thanks to the directionality of p_A , the sign of $(s_1 - s_0)$ is the same as the (constant) sign of p_A' . If p_B' is positive, then $(p_B(s_1) - p_B(s_0))$ has the sign of $(s_1 - s_0)$, the opposite otherwise. We thus see that M is directional, and that its directionality equals the sign of $p_A' \cdot p_B'$.

²² We can check the regularity of M by computing M' on any quale $q = p_A(s)$ experienced by A when exposed to stimulus s . Since $p_B = M \circ p_A$, we have $p_B'(s) = M' \circ p_A(s) p_A'(s)$. This relation puts a constraint on M' :

$$|M' \circ p_A(s)| \times |p_A'(s)| < K_B$$

where K_B measures the regularity of p_B . The only possibility for M' to reach high values is when $p_A'(s)$ is small or zero. At such locations, B would experience (rapid) qualitative change, while A would notice little change or no change at all. This discrepancy would contradict the basic assumption of the thought experiment that A and B would fail to communicate any difference in their experience. As a consequence, we must conclude that the mapping M between A's and B's qualitative spaces must be regular.

are admissible. This constitutes already a significant constraint on colour spectrum manipulation²³.

41. The preceding discussion does not rule out the theoretical possibility of spectrum discrepancies between two human beings. However, it highlights the importance of the continuous and evenly distributed stimulus-qualia projection. Because of it, no change of the spectrum can be envisaged except global shifts or global inversions. The fact that actual qualia offer resistance to manipulation during thought experiments indicates that they are not as fortuitous as would be expected from epiphenomena. Qualitative spaces are organised entities, and their structure shows characteristics which appear optimally designed for discrimination. Now we may try to relate this observation to the fact that qualia played a role in our evolutionary past.

7. Qualia and evolution: an analogy

42. Phenomenal consciousness is a biological characteristic of our species, and possibly of many others²⁴. From the preceding development, we conclude that qualitative spaces possess certain optimal properties. Is it sufficient to conclude that qualia play a causal role in cognition? It can still be argued that qualia, though optimally organized, are a mere accessory feature accompanying neural processes, that they are an effect without being the cause of anything. I propose here an indirect argument in support of the opposite option. In order to do so, I will use an analogy.

43. When ethologists study bird songs or calls, they are able to describe certain features as locally optimal. For instance, the acoustic spectrums of some song show features like large bandwidth and frequency modulation that seem correlated with an easy location of the singing animal and species-specific recognition (Hauser 1996; Naguib & Wiley 2001; Mathevon & Aubin 2001). Songs performing similar functions in different species tend to show the same type of characteristics. For instance, the maximum frequency of the song and intervals between notes are correlated with habitat density across species (van Buskirk 1997). From the ethologist's point of view, such acoustic characteristics are well designed for their observed function. They are not here by chance. Their local optimality only makes sense if they were produced by natural selection, since natural selection is regarded as the only source of locally optimal design in biology²⁵ (Nichols & Grantham 2000).

44. If we consider now the bird pharynx (syrinx) from a physiologist's perspective, it may be possible to show that some of its characteristics are locally optimal to emit a specific song. For instance, the complex syrinx of passerines, especially oscines, can be correlated in some detail with the complexity of their song. However, if we ignore the song, those characteristics

²³ Some further constraints come from the existence of four 'unique' colours for every subject. Any shift or inversion of the spectrum must globally preserve the set of unique colours.

²⁴ This does not preclude the philosophical possibility that artificial devices could have conscious experience.

²⁵ This argument is taken as granted by most evolutionary biologists. As Pinker and Bloom say in their efforts to bring language back into the scope of evolutionary science, "there are clear criteria for when selectionist and nonselectionist accounts should be invoked to explain some biological structure: complex design" (Pinker & Bloom 1990).

of the syrinx will remain incomprehensible. The local optimality of the syrinx structure can only be explained by the optimality of the song's acoustic properties.

45. Ancestors of song birds were selected for their ability to sing. Should we consider that singing itself was selected, or rather that the syrinx was selected in order to allow territory signalling? Perhaps we should look at the neural circuits involved in singing and say they were also selected for territorial signalling purposes. And what did selection retain after all, if not the genetic changes that make the difference between song birds and their non-singing ancestors? From genes to neural circuits, syrinx and song, there is a long chain of embryological and functional events. Each of these events is explained by its role in allowing the song. When ethologists study those birds, they do not exclude the song, they do not consider that the sole syrinx was shaped by evolution to perform territory signalling.

46. In this analogy, the song represents qualia, and the syrinx may stand for their neural correlates. What seems obvious in the case of avian ethology should hold for some aspects of the study of consciousness. Because they exhibit an optimal constitution, qualitative spaces, like bird song, should not be excluded from evolutionary accounts, but should rather be seen as a genuine link in the causal chain linking genes to biological function.

47. We can go one step further and suggest that *qualia should be taken into account* if we want to find some evolutionary justification for the corresponding brain states. Consider again the causal chain going from genes to biological function. Genes are not directly selected. They produce phenotypes, which undergo competition and selection. From genes to phenotypic characteristics, a long causal chain of structures is involved. The efficiency of the phenotype for gene transmission is what confers a Darwinian justification to the whole chain. Each link is explained in reference to the next link in the chain. At one end of the chain, genes are mere DNA sequences. They have no intrinsic feature that may justify their existence. To justify their presence in a given species, one must ultimately locate some advantageous phenotypic characteristic that they contribute to cause.

48. Phenotype is properly defined as the set of individual characteristics upon which natural selection directly acts (Dawkins 1982; Dessalles 1996). Qualitative spaces, like bird song, as far as they exhibit a locally optimal constitution, are either phenotypic themselves or part of the genotype-to-phenotype chain. Ignoring them would result in a broken chain with no evolutionary account. There would be no point in trying to explain why the syrinx has its actual characteristics if it is studied as if it belonged to a mute bird species. Likewise, neuroscientists cannot avoid considering qualia if they want to find some evolutionary justification for the presence of their neural correlates. The reason for this, once again, is that qualitative spaces show complex design that an engineer will inevitably ascribe to a discrimination function.

49. The comparison with ethology is no more than an analogy. It is of course thinkable, though extremely unlikely, that qualitative spaces evolved their orderly structure by pure chance, without having been selected for that. It is also thinkable that this structure, which seems so nicely designed for discrimination, fulfils another function. But those who support one of these two options are now in trouble, either being suspected to believe in miracles, or being requested to show how the structure of qualitative spaces relates in a locally optimal way to their supposed function.

8. Discussion: qualia and ghost spandrels

50. The conclusion we arrived at, that qualitative spaces are part of the genotype-to-phenotype chain, has further consequences. Like any biological characteristic showing functional design, qualitative spaces must have been shaped by natural selection. This can only be understood if qualia make a difference, if they helped our ancestors to be our ancestors. This conclusion is indirect evidence in support of a functional role of qualia. Even if such a functional role is difficult to conceive, it would be even harder to believe that qualia could be subject to natural selection without having any effect on cognitive processing.

51. An alternative account that may restore the epiphenomenality of qualia comes from the idea that the structure of qualitative space would be a mere consequence of physiology. This idea is captured by Chalmers' principle of structural coherence (in this quotation, awareness stands for cognitive processes):

“It is this isomorphism between the structures of consciousness and awareness that constitutes the principle of structural coherence. This principle reflects the central fact that even though cognitive processes do not conceptually entail facts about conscious experience, consciousness and cognition do not float free of one another but cohere in an intimate way.” (Chalmers 1995)

52. The principle presupposes an isomorphism. We rejected the possibility that qualia be mere by-product of neural activity, because as such, they would be evolutionary spandrels. Yet spandrels, contrary to qualia, show no intrinsic optimality. Chalmers could insist that qualia constitute a special type of spandrel, some sort of ghost spandrel. If qualia were a kind of echo of neural activity, qualitative experience could be epiphenomenal and yet appear as optimally structured. It would inherit its structure from its neural correlates.

53. Such a claim is hard to refute, since Chalmers' isomorphism must exist *ex post*: any difference between conscious experiences must be ascribed to differences between their neural correlates. The question is to know whether qualia were the target of natural selection or not. My argument in favour of qualia is that the structure shown by qualitative spaces is *simple*. This structure is, yet, optimally related to stimuli to allow discrimination. The set of corresponding brain states must have a much richer structure. No neuroscientist would accept the possibility that projecting the brain states corresponding to colour vision onto a three dimensional space can exhaust what can be said of these brain states. Why would Chalmers' isomorphism just pick out the few properties of brain states that are relevant to optimal discrimination²⁶? A proficient neurologist could perhaps trace acoustic properties of bird songs back to the corresponding brain states. But acoustic frequency and modulation are manifest in sound waves, not in brain states which encode them. Not only do qualitative spaces manifest optimal structure for discrimination, but they show no other structure. This constitutes a strong argument against the fact that qualia could be a kind of ghost spandrel, a faithful but superfluous echo of neural activity.

54. The alternative hypothesis, which consists in considering qualitative experience as a key link in the genotype-to-phenotype chain, explains why the structure of qualitative spaces is both poor and optimal for discrimination. Such a view, however, requires that qualitative experience play a definite biological role.

²⁶ If, as is claimed here, qualitative space are much poorer than corresponding brain activities, then Chalmers' term *isomorphism* is inappropriate, and should be replaced by *homomorphism*.

9. A biological role for qualia

55. The preceding development about the locally optimal structure of qualitative spaces and its putative connection with discrimination functions should make no sense from an eliminativist perspective. Some authors stress the fact that the notion of quale is very difficult to ground epistemologically. It is claimed that the notion of quale is vague and ill-defined, and that any phenomenon accounted for by a pure change in qualitative experience can also, in principle, be explained by physiological modifications (Dennett 1988). As any scientific concept, the notion of quale is as temporary as was the notion of light before it was reduced to electromagnetic phenomena. However, in the present state of knowledge, it seems difficult to describe the properties of experience that were highlighted here (continuity, directionality, regularity, dynamic range) without considering qualitative spaces. In the case of temperature, for instance, every scientist can check for herself those properties: higher heat transfers are experienced as warmer. Such simple facts are not predicted in eliminativist frameworks. They nevertheless constitute a genuine scientific problem that demands scientific explanation. One may object that qualia cannot be recorded and that verbal reports do not constitute an ultimate evidence of their existence. This is not sufficient to rule out qualia from the scientific scene. In conventional science, scientific observers are never sure to watch exactly the same phenomenon. They rely on the fact that their perceptions (possibly mediated by some artificial device) are correlated to conclude that they are dealing with a genuine scientific fact²⁷. It is not absurd to rely on internal checking when it comes to phenomena like continuity, directionality, regularity or dynamic range of qualitative experience. The fact that we can correlate our observations through language and mathematical formalism is sufficient to proceed with standard scientific procedures.

56. The central notion needed to establish the main claim of this paper is the notion of qualitative space. We do not need to consider other aspects of consciousness and their relation to phenomenal consciousness (Block 1995, Zalla 1996), though some authors may find it impossible to isolate phenomenal consciousness from higher aspects of cognition. For instance, crucial issues about qualia are the nature of their intentionality and of their mode of presentation, which may depend on language and/or a theory of mind (Carruthers 2000) or not (Siewert 1998). Strictly speaking, we do not have to decide about such matters. Our discussion is about the structure of qualitative spaces, not their relation to other aspects of cognition (like memory, theory of mind, concepts or attention). We should note, however, that our claim about the evolutionary role of qualitative spaces is hardly compatible with a position that would restrict phenomenal consciousness to humankind. Moreover, relevant properties of qualitative spaces (continuity, directionality, regularity, dynamic range) are not predicted by accounts that make them result from high-level cognitive abilities like language or theory of mind.

57. Qualitative spaces like colour space, for which there is some evidence of continuous, directional and regular stimulus-quale projections, seem to be optimally designed for discrimination. If, as we must presume, they result from natural selection, the next natural step

²⁷ Two individuals never see exactly the same rainbow in the sky at the same time, since the observer has to stay on the rainbow axis, at a definite distance, so as to see the rainbow with a 42.52° aperture. Similarly, two different persons, or the same person at two different moments, cannot be sure to deal with the same phenomenon when considering falling objects. Only correlations and mathematical laws like the law of gravity make them confident that their perceptions are alike.

in the reasoning is to suppose that qualia *must play a causal role* for this discrimination function. There is some evidence in support of this causal role. In colour vision, for instance, many different physical stimuli are experienced as identical (see above, end of section 5). Nevertheless, such stimuli cannot induce different behavioural responses. We are aware of every colour shade that we can behaviourally discriminate. As noticed by Chalmers, “every color distinction corresponds to some reportable information” (Chalmers 1995). This would be hard to explain if phenomenal consciousness was not involved in the functional process of discrimination.

58. If we follow this line of argument, we should conclude that *we perform discriminations on the basis of phenomenal qualities*. From this perspective, phenomenal consciousness is what led evolution in our lineage towards increasing discriminatory capacities. One may wonder how we can define *what* is being discriminated thanks to phenomenal experience. From the engineering perspective adopted here, we don’t need to consider this issue: qualia *ought to be* involved in a discrimination process. For the sake of the evolutionary argument, suffice it to say that improved discrimination is always better (at equal cost) and is selected, no matter what is discriminated which correlates with food, danger, mating opportunity, etc. We may add that, since many animals share neural organisation with us and perform the same kind of perceptual discrimination, we can hardly suppose that phenomenal qualities are a human privilege, based on the use of language, conceptual thought or theory of mind, as is sometimes claimed²⁸ (Carruthers 2000).

59. The hypothesis about the functional role of qualitative experience predicts our actual phenomenal variety. If phenomenal qualities were epiphenomenal spandrels, even ghost spandrels, there would be no reason why our perceptions would give rise to such a variety of phenomenal states. How could we explain the structure, the richness and the vividness of the qualitative repertoire if it was a fortuitous echo of neural activity? Conversely, if qualia are causally involved in the process of discrimination, then we understand not only their structural properties, but also their variety and their intensity, what we called the dynamic range of qualitative spaces.

60. Conscious species were selected according to the discrimination abilities that consciousness, for some reason beyond our reach, made possible. As a consequence, a rich gamut of phenomenal qualities was selected in each modality. The role played by conscious experience in discrimination is consistent with the fact that qualitative experience has a modular structure that systematically mirrors the functional organization of perceptual systems. It explains why phenomenal consciousness meets the fundamental constraints of signal discrimination efficiency by keeping relevant qualitative properties scrupulously apart. If, as claimed here, conscious experience is a powerful means for brains to perform efficient discrimination, then we must admit that it was selected for its own sake as an essential component of the perceptual process.

61. Phenomenal consciousness should thus be considered as a proper biological character, on which natural selection could act. Any increase in qualitative variety was likely to induce a

²⁸ Part of Carruthers’ claim, namely that phenomenal qualities are systematically available to higher cognitive processes (Carruthers 1996), remains perfectly sound in our context. What is no longer acceptable from the perspective put forward in the present paper is that phenomenal consciousness could emerge as an evolutionary by-product of the ability to read one’s own mind (Carruthers 2000), even if it might be the case for other aspects of consciousness.

better discrimination ability and, within certain limits²⁹, a more probable survival of individuals. Even if it remains far from obvious, we must consider that qualitative experience should be seen as a driving element in the evolutionary process which produced enhanced discrimination in perception.

10. Conclusion

62. The purpose of this paper was not to give a formal proof for the existence of qualia. There was no attempt either to explain why perception is accompanied by experience. By emphasizing the fact that qualitative spaces are not randomly arranged, we pursued only one goal: it should be much harder, now, to think that qualitative experience can be epiphenomenal.

63. The structure of qualitative spaces, which makes sense if qualia are involved in a discriminative process, gives us only an indirect evidence of the functional role of phenomenal consciousness. We concluded that qualia are not spandrels, that they must have some functional role, but this does tell neither why they are there, nor why they should be functionally necessary. The very nature of phenomenal consciousness is not addressed by the arguments we developed here. The claim about the optimality of the stimulus-qualia projection concerns only the *structure* of qualitative spaces. It does not tell much about the *intrinsic* properties of qualia. The way we experience heat or taste, for instance, and the fact that it differs from colour experience, lie out of the scope of the argument, which only concerns the relationships between qualia within a given modality.

64. What we showed, however, is that qualia cannot be excluded from a naturalistic account of our cognition, or even of our brain. Their structural properties, optimally designed for discrimination, strongly suggest that they are genuine links in the Darwinian selection loop going from genes to phenotype and back to genes.

65. At this point, we are in a situation in which qualia, because they are not spandrels, must be given a role in the functional organization of the mind. The structure of qualitative spaces reveals that they do have such a role, and that this role may be related to perceptual discrimination. However, we still lack the conceptual frame that could help us understand how this role is carried out. It looks as mysterious as Newton's forces through which corpses can act on each other at a distance. If we ever understand how qualia help us discriminate, we will also understand why evolution, at least in our phylum, had to produce phenomenal consciousness in order to design beings endowed with a powerful discriminative ability.

References

Block, N. (1980). "Are absent qualia impossible?". *Philosophical Review*, 89, 257-274.

Block, N. (1995). "On a confusion about a function of consciousness". *Behavioral and Brain Sciences*, 18, 227-287.

<http://www.bbsonline.org/documents/a/00/00/04/31/index.html>

²⁹ Clark rightfully observes that the emergence of a new dimension in the visual qualitative repertoire goes with a loss of spatial resolution (Clark 1996b).

- Bonnet, C., Ghiglione, R. & Richard, J-F. (1989). *Traité de psychologie cognitive*. Paris : Dunod.
- Carruthers, P. (1996). *Language, Thought and Consciousness*. Cambridge, MA : Cambridge University Press.
- Carruthers, P. (2000). "The evolution of consciousness". In P. Carruthers & A. Chamberlain (Eds), *Evolution and the human mind: modularity, language and meta-cognition*. Cambridge : Cambridge University Press, 254-275.
<http://cogprints.ecs.soton.ac.uk/archive/00001205/>
- Chalmers, D. J. (1995). "Facing up the problem of consciousness". *Journal of Consciousness Studies*, 2(3), 200-219. <http://www.imprint.co.uk/chalmers.html>
- Clark, A. (1993). *Sensory qualities*. Oxford : Clarendon Press.
- Clark, A. (1996). *Inversions spectral and bright - Comments on Melinda Campbell*. San Francisco : Society for Philosophy and Psychology, 31 May 1996.
<http://www.ucc.uconn.edu/~wwwphil/camptalk.html>
- Clark, A. (1996). "True theories, false colors". *Philosophy of science, PSA Suppl. Issue*, 143-150. <http://www.ucc.uconn.edu/~wwwphil/falseco.html>
- Dawkins, R. (1982). *The Extended Phenotype - The Gene as the Unit of Selection*. Oxford : W.H. Freeman & Co.
- Dennett, D. C. (1988). "Quining qualia". In A. J. Marcel & E. Bisiach (Eds), *Consciousness in contemporary science*. Oxford : Clarendon Press, 42-77.
<http://ase.tufts.edu/cogstud/papers/quinqual.htm>
- Dessalles, J-L. (1996). *L'ordinateur génétique*. Paris : Hermès.
<http://www.infres.enst.fr/~jld/papiers/pap.evol/96061301.html>
- Flanagan, O. (1992). *Consciousness Reconsidered*. Cambridge : MIT Press.
- Flanagan, O. & Polger, T. (1995). "Zombies and the function of consciousness". *Journal of Consciousness Studies*, 2(4), 313-321.
<http://oz.uc.edu/~polgertw/Polger-ZombiesJCS.pdf>
- Gould, S. J. & Eldredge, N. (1977). "Punctuated equilibria : the tempo and mode of evolution reconsidered". *Paleobiology*, 3, 115-151.
- Gould, S. J. & Lewontin, R. C. (1979). "The spandrels of San Marco and the Panglossian program: a critique of the adaptationist program.". *Proceedings of the Royal Society of London*, 205, 281-288.
- Harnad, S. (1987). "Psychophysical and cognitive aspects of categorical perception: A critical overview". In S. Harnad (Ed), *Categorical Perception: The Groundwork of Cognition*. Cambridge : Cambridge University Press, 1-52.
<http://cogprints.ecs.soton.ac.uk/archive/00001571/>
- Harnad, S. (2001). "No Easy Way Out". *The sciences*, 41(2), 36-42.
<http://cogprints.soton.ac.uk/documents/disk0/00/00/16/24/index.html>
- Hauser, M. D. (1996). *The evolution of communication*. Cambridge : The MIT Press.
- Holland, J. H. (1975). *Adaptation in natural and artificial systems*. Ann Arbor : The University of Michigan Press.

- Jackson, F. (1982). "Epiphenomenam qualia". *Philosophical Quarterly*, 32, 127-136.
<http://members.aol.com/NeoNoetics/Mary.html>
- Mathevon, N. & Aubin, T. (2001). "Sound-based species-specific recognition in the blackcap sylvia atricapilla shows high tolerance to signal modifications". *Behaviour*, 138, 511-524. <http://h0.web.u-psud.fr/cb/BlackcapBehaviourMathevon.pdf>
- Munsell, A. H. (1905). *A color notation*. Baltimore : Munsell Color Company.
- Nagel, T. (1974). "What is it like to be a bat ?". *Philosophical Review*, 83, 435-450.
- Naguib, M. & Wiley, R. H. (2001). "Estimating the distance to a source of sound: mechanisms and adaptations for long-range communication". *Animal behaviour*, 62, 825-837. <http://www.unc.edu/home/rhwiley/pdfs/NaguibWiley2002.pdf>
- Nichols, S. & Grantham, T. (2000). "Adaptive complexity and phenomenal consciousness". *Philosophy of science*, 67, 648-670.
[http://www.cofc.edu/~nichols/evolcons\(final\).htm](http://www.cofc.edu/~nichols/evolcons(final).htm)
- Palmer, S. E. (1999). "Color, Consciousness, and the Isomorphism Constraint". *Behavioral and Brain Sciences*, 22(6), 923-990.
<http://www.bbsonline.org/documents/a/00/00/05/04/index.html>
- Pinker, S. & Bloom, P. (1990). "Natural language and natural selection". *Behavioral and Brain Sciences*, 13(4), 707-784.
<http://www.bbsonline.org/documents/a/00/00/04/99/index.html>
- Saunders, B. A. C. & van Brakel, J. (1997). "Are there non-trivial constraints on colour categorization?". *Behavioral and Brain Sciences*, 20(2), 167-228.
<http://www.bbsonline.org/documents/a/00/00/04/87/index.html>
- Sève, R. (1996). *Physique de la couleur*. Paris : Masson.
- Shepard, R. N. (1962). "The analysis of proximities: Multidimensional scaling with an unknown distance function". *Psychometrika*, 27, 219-245.
- Shoemaker, S. (1975). "Functionalism and qualia". *Philosophical Studies*, 27, 291-315.
- Siewert, C. P. (1998). *The significance of consciousness*. Princeton : Princeton University Press.
- Van Buskirk, J. (1997). "Independent evolution of song structure and note structure in American wood warblers". *Proceedings of the Royal Society of London*, B(264), 755-761.
<http://www.zool.unizh.ch/ecology/jvanbuskirk/ProcRoySocLond1997.pdf>
- Zalla, T. (1996). *Unité et multiplicité de la conscience : une étude critique des théories contemporaines à la lumière d'une hypothèse modulariste*. Paris : PhD Dissertation, Ecole Polytechnique.

Jean-Louis Dessalles is Associate Professor at the Ecole Nationale Supérieure des Télécommunications, Paris, where he studies artificial evolution (genetic algorithms) applied to the evolution of cognition and communication.