



The topological space of subjective experience

Catherine Tallon-Baudry

► To cite this version:

Catherine Tallon-Baudry. The topological space of subjective experience. Trends in Cognitive Sciences, 2022, 26 (12), pp.1068-1069. 10.1016/j.tics.2022.09.002 . hal-03874056

HAL Id: hal-03874056

<https://hal.science/hal-03874056>

Submitted on 14 Dec 2022

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.

The topological space of subjective experience

Catherine Tallon-Baudry, Cognitive and Computational Neuroscience Lab, Inserm, ENS-PSL

Corresponding author: catherine.tallon-baudry@ens.psl.eu (C. Tallon-Baudry)

Abstract: Subjective experiences often feel rich, yet are most often quantified with simple metrics, such as a few levels on a predefined scale. What are the dimensions and topological organization of subjective experience? How do they relate to behavioral output, how do they map onto the classical cognitive domains?

Keywords: Consciousness, perception, emotion, decision, natural language processing

Our subjective experiences are determined by our perceptions of the external and internal world—our moods, thoughts, and bodily feelings. Understanding how subjective experience arises from biological activity is one of the most enduring questions in cognitive science, yet we know surprisingly little about the structure of subjective experience. Many standard behavioral measures do not capture subjective experience adequately – for instance, the ability to detect and discriminate visual stimuli does not necessarily permit inferences about subjective visual experiences, as blindsight patients can sometimes perform these tasks near perfectly despite not having visual subjective experiences [1]. When subjective reports are employed they are reduced to a few levels on a predetermined dimension (e.g., visibility, pleasantness, or confidence). Conversely, behavior can be influenced by factors that are not part of subjective experience, as in unconscious priming for instance. In other words, behavior and subjective experience cannot, and should not, be conflated. Compared to behavior, subjective experience has received much less attention in experimental sciences. One way to generate proper metrics of subjective experience is to establish a description of the topological organization of subjective experience.

Think about the last time you browsed through images –be it all TiCS covers since the journal’s first issue in 1997, or perhaps more commonly, vacation photos. For each image, you had a subjective experience, which was not only visual, but also tinged with an emotional component. Both perceptual and emotional components contributed to your decision to tag some items for later perusal. Visual perception, emotional feelings, and decision making correspond to three domains of investigation in cognitive science, with different questions asked about the same physical stimulus: for instance when presented with a painting, you can be asked to describe what you see (a sea-shore landscape, ...), what type of feelings it elicits (pleasant, calming, ...), and whether you want to buy it or not. Yet, it seems reasonable to assume that those three judgements are based on the same, or at least initially similar, subjective experience of the image. Beyond intuition, there are some experimental arguments in favor of this assumption. For instance, during a perceptual task (judging the age of paintings, houses, and faces), the human brain automatically rates the pleasantness of the image presented, even though the task is purely perceptual [2], and monkey primary visual cortex encodes reward value as early as 100ms [3].

What do the fields of emotion, vision, and decision making say about subjective experiences? Emotions are defined by the combination of subjective emotional feelings, specific behavior (e.g., freeze, fight, flight), and physiological reactivity. Emotional feelings share at least two dimensions, valence (positive/negative) and arousal (relaxing/exciting) [4]. Valence and arousal ratings are useful compact experimental measures, but projecting emotional feelings on two dimensions conflates distinct feelings, such as fear and anger. More complex semantic emotional spaces have been considered but these are often based on semantic tags preselected by the experimenters, which means some

dimensions may be omitted. For instance, in one recent study of semantic emotional space [5], none of the proposed dimensions was perceptual even though there are some systematic associations between perceptual dimensions and affective feelings (e.g., brightness and valence [6]).

In the visual domain, subjective reports are sometimes collected (for instance, in visual detection at threshold) but the dimensions and topological organization of the full-fledged experience of natural images are rarely studied *per se*. The field has mostly progressed by analyzing the links between behavioral and/or neural measures and perceptual dimensions pre-defined by the experimenter - although recently, perceptual and conceptual dimensions underlying object similarity judgements have been identified in a data-driven manner [7]. Finally, while it seems reasonable to assume that task-relevant dimensions are somewhat related to spontaneous visual experience, this has rarely been experimentally tested [8].

In the field of decision making, much attention has been paid to decision confidence, but other dimensions of subjective experience have largely been ignored. Decisions are based on value, often simplified to appear as a number on a common scale, the underlying space being a simple straight line with Euclidian distance [9]. While this concept has been very fruitful, it cannot easily explain the distinction between liking and wanting for instance, or how value is built from both internal [10] and external [11] information. Altogether, it appears that while we are beginning to get a glimpse of the dimensions of subjective experience, this view is only *partial* given that most dimensions are preselected and differ depending on the domain, and *indirect* given that these dimensions are revealed through specific tasks.

How can we specify the multi-dimensional space of subjective experience? First, we have to accept subjective experience at face value to be able to ultimately explain it. The most obvious proxy for subjective experience is its verbal description in humans. Verbal descriptions have pitfalls and caveats, notably biases related to cultural or social differences which would have to be carefully taken into account, but they are extremely rich. So far, their very richness has hampered their use for lack of appropriate methods of analysis. However, the novel tools developed in the field of Natural Language Processing (NLP) can extract meaningful information from human verbal descriptions. They are already employed – with caution – in psychiatry, and have for instance revealed structured patterns from descriptions of hallucinogenic experiences [12]. Briefly, NLP tools can generate multi-dimensional spaces where semantically related terms cluster together – in a purely data-driven manner. Coupled with graph analysis, this seems a promising starting point to probe the topological organization of subjective experience and derive both novel metrics and experimentally testable predictions. How many dimensions do we need to account for subjective experience, and how do those dimensions differ between individuals? What are the properties of the space defined by those dimensions, are they homogeneously distributed or are there some regions with specific properties? How is the subjective space altered by task instructions, and do all task-related variables map onto subjective space?

Why does exploring the topological structure of subjective experience matter? Perhaps most obviously, understanding how we go from neuronal to subjective space is an important question in cognitive neuroscience – and my personal interest would be to probe the role played by interoceptive signals in this transformation. But this research program would also offer a formal framework to study perception in a more ecological manner – for instance, using visual scenes it might be possible to construct a model relating physical image properties, neuronal activity, task variables, and subjective experience - with a huge potential to improve the translation of results obtained in controlled laboratory settings to everyday life. Such a description would also offer a unifying framework to integrate conscious perception, emotional feelings, and choices, offering a glimpse of how the unity of consciousness could be achieved and helping us better understand the origin of some decisional biases. The nature of the dimensions is also likely to be informative – does it map onto classical domains as

currently in use in cognitive science? For instance, is there a “past-present-future” axis and how does it relate to the domains of memory, planning, or learning? Recently, it has been proposed to revise the definition and classification of cognitive processes inherited from psychology or in use in psychiatry based on the evolution of both anatomical data and behavior [13], or based on task-related neuronal data [14]. How do those novel classifications maps onto the dimensions of reported subjective experience? Are there discrepancies, which could be related to the mismatch between behavior and subjective experience? Finally, while the topological structure of subjective experience would initially stem from verbal descriptions, non-verbal markers might be derived to probe the existence and quality of subjective experience in humans unable to communicate (post-comatose patients, pre-verbal infants but also maybe fetuses) as well as in animals or artificial systems, with important societal implications.

References

1. Sanders, M.D. et al. (1974) "Blindsight": Vision in a field defect. *Lancet* 1 (7860), 707-8.
2. Lebreton, M. et al. (2009) An automatic valuation system in the human brain: evidence from functional neuroimaging. *Neuron* 64 (3), 431-9.
3. Stanisor, L. et al. (2013) A unified selection signal for attention and reward in primary visual cortex. *Proc Natl Acad Sci U S A* 110 (22), 9136-41.
4. Russell, J.A. and Barrett, L.F. (1999) Core affect, prototypical emotional episodes, and other things called Emotion: Dissecting the elephant. *Journal of Personality and Social Psychology* 76 (5), 805-819.
5. Cowen, A.S. and Keltner, D. (2017) Self-report captures 27 distinct categories of emotion bridged by continuous gradients. *Proc Natl Acad Sci U S A* 114 (38), E7900-E7909.
6. Cavanaugh, L.A. et al. (2016) Perceptual dimensions differentiate emotions. *Cogn Emot* 30 (8), 1430-1445.
7. Hebart, M.N. et al. (2020) Revealing the multidimensional mental representations of natural objects underlying human similarity judgements. *Nat Hum Behav* 4 (11), 1173-1185.
8. Campana, F. et al. (2016) Conscious Vision Proceeds from Global to Local Content in Goal-Directed Tasks and Spontaneous Vision. *J Neurosci* 36 (19), 5200-13.
9. Levy, D.J. and Glimcher, P.W. (2012) The root of all value: a neural common currency for choice. *Curr Opin Neurobiol* 22 (6), 1027-38.
10. Keramati, M. and Gutkin, B. (2014) Homeostatic reinforcement learning for integrating reward collection and physiological stability. *Elife* 3.
11. Pelletier, G. et al. (2021) A Preferential Role for Ventromedial Prefrontal Cortex in Assessing "the Value of the Whole" in Multiattribute Object Evaluation. *J Neurosci* 41 (23), 5056-5068.
12. Ballentine, G. et al. (2022) Trips and neurotransmitters: Discovering principled patterns across 6850 hallucinogenic experiences. *Sci Adv* 8 (11), eabl6989.
13. Cisek, P. (2019) Resynthesizing behavior through phylogenetic refinement. *Atten Percept Psychophys* 81 (7), 2265-2287.
14. Beam, E. et al. (2021) A data-driven framework for mapping domains of human neurobiology. *Nat Neurosci* 24 (12), 1733-1744.