

# Qualia: The Hard Problem

Todd W. Griffith\* and Michael D. Byrne†

\*College of Computing  
Georgia Institute of Technology  
Atlanta, Georgia 30332-0280  
griffith@cc.gatech.edu

†School of Psychology  
Georgia Institute of Technology  
Atlanta, GA 30332-0170  
byrne@cc.gatech.edu

## Abstract

One issue that has been raised time and again in philosophy of mind and more recently in cognitive science is the question of qualia, or “raw feels.” What are qualia and how do they fit into the cognitive science conception of mind? We consider some of the classic qualia thought experiments and two proposed solutions to the qualia problem, eliminativism and content-dependence. While neither of these solutions are actually able to dismiss or explain qualia as claimed, the content-based solution does clarify the relation between cognitive science and qualia. Because qualia are precisely the part of our experiences that are not related to informational content (and therefore intersubjective), and cognitive science is primarily based on information content, qualia are not within the domain of cognitive science.

## Introduction

The nature of qualia or “raw feels” has always been a philosophical issue at least implicitly, and recently cognitive science has tried its hand at the problem. In this paper we argue that the real issue of qualia is not an issue for cognitive science as it presently exists, but remains a strictly philosophical issue. Qualia is the term which is applied to indivisible primary feelings such as the feeling of seeing red in an apple or the feeling of pain in one’s foot. Traditionally qualia are considered to be purely first-person or subjective in nature. Today, however, with the aid of neuroscience some researchers in cognitive science have claimed to have solved the qualia problem in an objective sense. We disagree.

The problem of qualia is a problem of where. Where do qualia reside? It is the central question of the dualist/materialist debate and a core issue in the philosophy of mind. This question also reduces the question of consciousness to its primitives, e.g. where is my sensation of pain, of red, etc. The problem is best illustrated through an example: Suppose that I take a pin and push it into my finger (OK, so I’m not too bright). This action is objectively observable by normal means. As the pin enters my finger chemicals in the damaged tissue around my finger are released, which cause nerves to fire impulses to my brain. This too is an objectively observable phenomenon, albeit not by normal means. Now let us suppose that we trace these events full circle. We can, in theory at least, objectively observe all the neurons firing

throughout my brain in a causal series, eventually ending in a signal from my brain to my hand causing it to move. Break the phenomenon down as you prefer; perhaps to chemical reactions or quantum phenomena. The question is still: where’s the pain?

Is the pain reducible to my behavior? Clearly not, given that I can conceive of a situation where I would behave similarly without experiencing the pain (e.g. reflex actions work even on anesthetized patients, and simple robots react to negative stimuli). The problem is that qualia cannot be reduced to a third-person perspective. Searle (1992) attacks materialism by arguing that the first person point of view cannot be ignored, e.g. my pain is clearly mine, and all pain is clearly someone’s. Nagel’s (1974) idea of “what it is like to be” something also comes to mind. Searle points out that the insistence on objective criteria for truth is misplaced when one is talking about mental activity. Mental activity has a character that is outside the scope of objective observation yet is still clearly true. It is the “terror” of the subjective that has led to many of the current materialist positions in the philosophy of mind.

Two now classic thought experiments have been presented which illustrate some of the difficult issues regarding the nature of qualia: Thomas Nagel’s (1974) “What is it like to be a bat?” and the inverted spectrum. Nagel argues that consciousness is an issue of “what it is like to be” something. He presents the argument that given a complete neurophysiological account of the workings of a bat’s brain would not in any way enable us to know what it is like to be that bat. So even if we have a complete neurophysiological trace of a bat chasing a fly, we still have no way of knowing what it is like to be that bat observing that fly. Thus objective knowledge can not provide us with access to the bat’s qualia. This of course translates to you and I. I cannot know your qualia and you can not know mine. Qualia are in some sense perfectly subjective. The inverted spectrum thought experiment also argues that it is not necessarily the case that my qualia are like yours. The experiment goes something like this: suppose that when I see red I get the same experience as when you see violet. In fact, suppose that for the entire color spectrum my experience is completely inverted with respect to yours. What we both call “red,” for example, will be a quale (singular of qualia) of violet for me, and red for you. This experiment shows that there is no logical necessity that our subjective experiences are the same for the same stimulus. It raises the question: is there any way to know that we

experience the same things? These thought experiments capture the difficult issues with regard to qualia that cognitive science does not have the tools to address.

Why are cognitive scientists concerned with the issue of qualia? There are at least two important reasons: the first is that qualia are directly tied to consciousness, which many claim is directly related to behavior and action. The second is that in many ways qualia represent the brass ring. If we can scientifically characterize subjective feelings such as pain, fear, or the smell of baking bread, we are very close to understanding a central aspect of human cognition. Unfortunately the current tools of cognitive science are not up to the task. Current solutions either leave the difficult problems totally untouched or discount them as fictitious, both of which are inadequate.

### Proposed Solutions to the Qualia Problem

The fact that we lack an adequate account of what qualia actually are does not mean that there have not been approaches from a cognitive science perspective to solving the qualia problem. The two more prominent solutions are eliminativism and the qualia vs. content argument. Each of those will be considered here, beginning with eliminativism. Eliminativism, in particular "eliminative materialism," is most strongly associated with Churchland (e.g. 1988, 1989). The eliminativist perspective on qualia mirrors the eliminativist arguments for essentially all mental phenomena for which eliminativists currently lack a strong neural explanation: qualia do not need to be explained because they do not really exist. The belief that there are qualia is a vestige of the "folk psychological" viewpoint which will ultimately be eliminated when we develop a complete neuroscientific account of mind/brain. The concept of qualia—along with a host of other mind/brain concepts—will simply not be a part of this more mature understanding of neuroscience.

There are (at least) two problems with this proposed solution to the qualia problem. First, there is no guarantee whatsoever that the problem will simply go away if we do the right things in neuroscience. What, exactly, is the basis for the belief that every single extant mind/brain concept will necessarily be eliminated, other than the fact that some concepts in some other fields have been eliminated? This question has never been satisfactorily answered. Furthermore, the eliminativist camp has yet to even approach letting the world know which things and which things will not be eliminated—it has been implied that all current concepts of mind/brain will be eliminated, but this seems unreasonable. Certain concepts in other disciplines have been eliminated by lower-level reductions (e.g. impetus) but others have not (e.g. force, time). Despite what eliminativists (and members of Congress) might believe, problems tend not to simply vanish when vague promises of later solutions are proposed.

Second, this neuroscientific promissory note provides little satisfying information. By what will qualia be eliminated? For that matter, what would a neuroscientific account of whatever will replace qualia even look like?

When will such an account be developed? In the wake of proposing the elimination of qualia and just about everything else in mental life, eliminativism leaves just as many unanswered questions as it started with.

A somewhat more constructive approach to the qualia problem is the content vs. qualia argument offered in slightly different ways by Akins and Tye. Akins's paper is an attempt to rebut the "what is it like to be a bat?" argument. To make the argument, Akins (1994) has the reader consider the following thought experiment: A neuroscientist travels to the future and returns with the authoritative, future-neuroscience-approved "film" of what it's like to be a bat. This film contains:

a kaleidoscopic display of vibrant colour forms. Swirling and pulsating in three dimensions, the colored forms dance across the screen, colliding and dispersing, suddenly appearing or vanishing. That's all. That, I claim, is what it's like. (p. 262)

Obviously, something is amiss here—how can that be all there is? Akins's claim is that the entire qualia enterprise is mistaken just for the same reason that the bat film is nonsensical, because it assumes that "separation of our conscious experience into two parts, the representative and qualitative aspects" is possible. The point of the bat film is that the qualitative aspects of the bat's experience are meaningless to us because we lack the perceptual and representational capacities of the bat. There is no qualia problem because the mere idea of separation of the quality of the experience from the content of the experience is nonsensical. However, the fact that the qualia would be unintelligible hardly constitutes proof that the qualia do not exist.

Tye (1991) presents a similar argument directed at visual qualia. Arguments have been made that the difference between certain visual experiences (e.g. the difference between seeing blue and seeing red) is a difference in the qualia of the two experiences. Tye argues that in all such examples, there is in fact a difference in the information contents of the two experiences and that the experiences feel different not because there are any qualia involved, but because of this difference in content. He maintains that "[t]he 'felt' aspect simply cannot be divorced from the representational aspect" (p. 133). That is, the qualia for seeing red is perfectly correlated with the information that the visual stimulus reflects red light. In essence, this is the same argument as Akins's—qualia do not exist independent of content.

This is a much more clearly elucidated and convincing argument than the eliminativist argument. Rather than ignoring the problem and hoping it will go away, this solution attempts to locate the source of the problem in something familiar to the cognitive/computational perspective: information content. If the qualia issue is a content issue, it lends itself to analysis by the traditional methods of cognitive science. In that case, there is nothing that is in principle unexplainable about qualia and an account, though we may not have it in hand, should certainly be reachable.

While this is a compelling perspective, it still falls short of the goal of eliminating or explaining qualia. While it

may be the case that all intra-individual differences in qualia are associated with differences in content, this does not provide a guarantee that the subjective feel of the experiences is the same for different people. Consider again the inverted spectrum problem; that is, persons A and B can both accurately discriminate and name different colors but their subjective experiences are reversed. That is, "seeing red" for A feels like "seeing violet" for B and vice versa. The content perspective does not eliminate this possibility. Both A and B have the same rods and cones (i.e. Akins's perceptual machinery) and both A and B have the same color distinctions and categories (i.e. representational content). As Tye suggests, within an individual, the qualia are perfectly paired with informational content. Person A always has a "red" quale when seeing a red object—the information content determines which quale. Person B also always has a "red" quale when seeing a red object for the same reason. However, there is no way to guarantee that the subjective feel of seeing red is the same for both A and B. Nothing in Tye's account guarantees inter-individual agreement on subjective feel. The inverted spectrum problem clearly remains, even if we assume that a particular quale is tied to particular information content for a given observer. The problem of inter-subjectivity is simply not solved by invoking associations between qualia and contents.

### Not a Cognitive Science Problem

While the association between content and qualia does not, as the authors claim, solve the qualia problem in that it does not explain what and where qualia are, the "content solution" has important implications with respect to the nature of qualia and inquiry in cognitive science. The failure of the content solution to actually do away with qualia clearly demarcates the boundaries between what is content and what is quale. That which has informational content is not quale, and, importantly, vice versa: that which is quale has no informational content. Content, specifically informational content, is the objective substance of cognitive science. Conducting experiments without this substance is futile because the experimenter can never know that the "subjective content" of his or her inquiry is remotely similar as the "inverted spectrum" problem clearly illustrates.

Examples may help illustrate. Consider again the inverted spectrum problem. When perceiving light of some wavelength, both person A and person B, there is indeed informational content to the sensation of seeing the color, which would be something like "light of wavelength X." This is the information content of the seeing experience, and both A and B would get this information from the sensation—just as Tye points out. However, in the case of inverted spectra, A and B will **not** have the same subjective feel of the experience. What "feels" red to A may well "feel" blue to B, despite the identical information content of the experience. We define qualia by what they are not: qualia are what is left of sensations after the objective information content is removed. As Tye suggests, there may

be a perfect correspondence between which qualia will be felt when certain information is conveyed, but this does **not** eliminate the qualia.

We do not deny that sensations have informational content—dropping a hammer on one's foot makes this immediately clear.<sup>1</sup> Information about an impact, such as rough estimates of the weight, hardness, and velocity of the object striking the foot, is quit available, and this is indeed information in the objective sense. If one were to hook up sophisticated neuron recording devices, a third-party observer could also likely get much of this information based on which neurons fire and at what rate. But this observer would miss the "feel" of the hammer striking his or her foot. The sensation contains two things: objective information, and something that is **not** objective information. The latter are qualia.

This distinction is of paramount importance to cognitive science. While there are probably few things that all practitioners of cognitive science agree on, the centrality of information processing is almost certainly one of them. Take, for example, the list of keywords for the 1996 Cognitive Science Annual Conference—almost every single one of these research categories implicitly or explicitly relies on information-processing accounts of one form or another. If one were going to select the single identifying characteristic of research in cognitive science, information processing would be an excellent candidate. It is the information in the stimuli that concern the experimental psychologist, it is the information that is handled in the programs of computer scientists, it is the information in language that concerns the linguist. Science in general is an enterprise of explaining the observed and objective information about the natural world. Thus, qualia do not fall under the domain of cognitive science. Qualia are precisely the aspects of our experience that do not have the kind of information content which is at the core of cognitive science. If there is no such objective informational content to study, it is surely impossible to make a science out of the endeavor. How could scientific method be appropriate in domains lacking informational content?

As pointed out earlier, qualia are a problem of inter-subjectivity. Issues regarding qualia fall into the domain of the philosophical phenomenologist, not the cognitive scientist. Cognitive science is materialistic in orientation, and clearly so in scope. Yet phenomenology and "things with no content" seem more along the lines of dualism. However, based on our current knowledge of the material, there may be no other approach, despite the fact that dualism has been rejected even by critics of traditional cognitive science such as Searle.

Even Searle, however, does not adequately answer the dualist. He claims that dualism has been "thoroughly discredited," and therefore should not be considered. We see only three possibilities for his dismissal of the dualist perspective. 1) He is terrified like his peers that there is something beyond the physical. 2) He is making a

<sup>1</sup> We would like to thank an anonymous reviewer for the suggestion that we conduct this particular experiment. However, we conducted it as a thought experiment only.

Lakatosian claim that dualism has ceased to provide interesting scientific problems to work on. 3) He is claiming that dualism has been discredited by the materialist positions which he himself tears down. Of these we believe the answer must be the second. "Dualism in any form is today generally regarded as out of the question because it is assumed to be inconsistent with the scientific world view." (p. 3) Unlike Einstein, however, who eliminated the concept of "aether" from scientific discourse by subsuming those aspects accessible to scientific investigation into the concept of "field," dualism continues to escape subsumption.

It is the subjective aspects of qualia that remain untouched by current attempts to solve the qualia problem. These attempts attack what David Chalmers (1995b) calls the easy problems of consciousness, such as "how can a human subject discriminate sensory stimuli and react to them appropriately?" (p. 81) The hard problem is "the question of how physical processes in the brain give rise to subjective experience" (p.81). The difficulty lies in the fact that normal scientific measures of objective fact fail to have an impact in this domain. In this respect Searle provides us with at least a starting point. He argues that the evidence for the nature of qualia will not be in the realm of the objective but in the equally real realm of the subjective. Thus objective scientific techniques will not suffice. It is for this reason that the qualia problem is a purely philosophical issue and not a cognitive science issue. For until we have the scientific tools to handle subjective evidence, the study of qualia will remain squarely in the realm of phenomenology.

This is not to say that the issues surrounding qualia are unimportant, but simply that they are not issues for cognitive science. We expect reactions to our conclusions to be varied. One of us (MB) believes that this means qualia are really of no concern, while the other (TG) believes this makes dualism and phenomenology attractive alternatives to mainstream cognitive science. We are in agreement, however, that qualia lie outside the domain of inquiry in cognitive science.

### Acknowledgments

We would like to thank Nancy Nersessian for getting the two of us interested in this topic at the same time, pointing us to some of these references, and initiating the discussions which ultimately led us to writing this paper.

### References

- Akins, K. A. (1994). A Bat without Qualities? In M. Davies & G. W. Humphreys, (Eds.) *Consciousness: Psychological and Philosophical Essays*. (pp. 258-273). Cambridge, MA: Blackwell.
- Chalmers, D. J. (1995a) The Puzzle of Conscious Experience. *Scientific American*, December, 80-86.
- Chalmers, D. J. (1995b) Explaining Consciousness: The "Hard Problem." Special issue of *Journal of Consciousness Studies*. Vol. 2. No. 3; Autumn 1995.
- Churchland, P. M. (1988). *Matter and Consciousness* (revised edition). Cambridge, MA: MIT Press.
- Churchland, P. M. (1989). *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*. Cambridge, MA: MIT Press.
- Nagel, T. (1974). "What is it like to be a bat?" *Philosophical Review* 4 LXXXIII: 435-450.
- Searle, J. R. (1992). *The Rediscovery of the Mind*. Cambridge, MA: MIT Press.
- Tye, M. (1991). *The Imagery Debate*. Cambridge, MA: MIT Press.