Illusionism on the Brink of Disillusionment

Ariel LaFayette
University of Toronto

Abstract
Distilling a theory of “qualia”—with at least some degree of consensus—has been a grand point of contention in the philosophy of mind since the term qualia were introduced. In this paper, I will focus on one area where qualia realists and anti-realists come to head-to-head; how does philosophy sufficiently account for qualia’s constitutive property of “intrinsic subjectivity”? First, I will summarize David Chalmers’ meta-problem to contextualize this aspect of the qualia debate. Second, I explain how Keith Frankish’s theory of “illusionism” is a tenable solution to resolving the meta-problem from an anti-qualia realist perspective. Yet, when it comes to addressing qualia’s constitutive feature of intrinsic subjectivity, illusionism faces an insurmountable obstacle due to its methodological commitment to third-person empiricism. I argue the preceding point by analyzing how illusionism confronts two challenges: (1) Philip Goff’s real-acquaintance hypothesis, accompanied with Goff’s theory of panpsychism; (2) a modern adaptation of Ludwig Wittgenstein’s “private language” thought experiment. My purpose is to argue that Frankish’s theory of illusionism is ill-equipped for dealing with these philosophical challenges.

Introduction
The term “qualia” is referenced in a wide range of philosophical arguments which maintain the elements of conscious experience that cannot be reduced to neuroscientific explanation. The classic definition of qualia refers to the intrinsic, subjective, and ineffable components of conscious experience (Frankish 2012, 2-3). The subjective component is “what-it’s-like” to undergo a conscious experience, emphasizing the subject’s direct access to their own conscious experience. Qualia realists claim that even if we provided the most sophisticated neuroscientific explanation of our conscious experiences, this would fail to encapsulate qualia (Nagel 1974, 435). Therefore, qualia are irreducible as such.
The philosopher Keith Frankish defends the potential for third-person empiricism, such as developments in neuroscience, to “explain away” the ontological reality of qualia. In response to the qualia realist’s convictions, Frankish claims they are cast under a cognitive “illusion” that qualia properties exist. For example, he argues that the claim qualia are accessed via “introspection” will eventually be explained as a cognitive illusion (Frankish 2016, 12-13). In this sense, Frankish respects that qualia realists claim that it “feels like” there are qualia properties, but he denies their ontological existence.

The aim of my paper is to show how Frankish’s proposed method for explaining away qualia through neuroscience hits a dead-end when attempting to dismantle qualia’s constitutive property of intrinsic subjectivity. The crux of my argument is that Frankish’s commitment to third-person empiricism ultimately prevents him from directly addressing the definition of intrinsic subjectivity. My paper will show how this problem emerges in two different contexts. First, I will show how Frankish’s theory of illusionism faces against Philip Goff’s “real acquaintance hypothesis” and panpsychism, wherein the property of intrinsic subjectivity plays a central role. Second, to further drive the point that qualia’s property of intrinsic subjectivity is irreconcilable for illusionism, I will be re-imagining Wittgenstein’s famous “private language” thought experiment from a modern perspective. Lastly, I will present and respond to a potential counterclaim by Frankish, in which he defends his treatment of qualia realism for being compatible with David Chalmers’ meta-problem.

Qualia & the Meta-Problem
A discussion about qualia in the 21st-century seems to always begin with reference to Chalmers’ seminal hard problem—how philosophy can account for conscious experience if our scientific understanding of the functional, dynamical, and structural properties of the brain cannot? (Chalmers 1995, 200). Chalmers’ critics note that his phrasing of the hard problem is misleading because he relies on the a priori stipulation that brain processes and conscious experiences are ontologically distinct. As a result, Chalmers problematically dismisses the sizable literature arguing that our gaps in theories of consciousness will be overcome through advancements in neuroscientific research (Schier & Carruthers 2017). With respect to that hypothesis, many anti-qualia realists prefer the qualia-neutral perspective
Chalmers adopts in the phrasing of his meta-problem: “why do we ‘feel’ that qualia exist?” Unlike the hard problem, the meta-problem does not insinuate that qualia properties must necessarily exist in order to address the discrepancy between our scientific understanding of brain processes and our conscious experiences. At the same time, Chalmers’ meta-problem maintains that qualia realist arguments—such as the introspective claim that qualia exist—must be reckoned with in any theory of consciousness. To the advantage of both qualia realists/anti-realists alike, embarking on the meta-problem will either: (a) shed light on how to navigate the hard problem, or (b) illuminate an avenue for a neurophysiological explanation that thoughtfully addresses qualia realist concerns (Chalmers 2018).

Qualia & Intrinsic Subjectivity
According to Chalmers, anti-qualia realism’s response to the meta-problem requires a robust explanation for qualia realists’ long-held intuitions. My paper will zero in on Chalmers’ metaphysical intuition that consciousness is intrinsic to an individual’s subjective experience (Chalmers 2018). According to Chalmers, the metaphysical intuition is indebted to René Descartes’ argument that I can doubt all the external relationships in the physical world, but I cannot doubt the fact that I am conscious. Therefore, the datum I can be most certain of is the fact that I am conscious (Descartes 1641).

We can see Descartes’ legacy in contemporary theories claiming that qualia are “intrinsically subjective.” To claim a property is “intrinsic” means that its existence can be grasped independently from all other extrinsic properties. Attributing the feature of “intrinsic” to subjective qualia properties implies that qualia must necessarily be grasped by the subject directly via first-hand experience (Langton & Lewis 1998). Qualia theories of this kind draw upon Bertrand Russell’s insight into the limitations of the material sciences—which are restricted to structural-functional explanations—for its inability to explain intrinsic natures (Russell 1927). Inspired by Russell’s views, Chalmers states, “the problem [referring to the hard problem] is hard precisely because it is not a problem about the performance of functions (Chalmers 1995). The problem persists even when the performance of all relevant functions are explained.” Positing that “qualia are intrinsic” resolves Chalmers’ hard problem that our structural-functional brain
processes fail to account for subjective experience since qualia serve the role as a “missing puzzle piece” that is non-structural-functional. Simultaneously, the qualia-as-intrinsic hypothesis addresses Russel’s concern that structural-functional relationships as described in the material sciences cannot account for intrinsic natures.

The qualia-as-intrinsic hypothesis may lead to two distinct ontologies: a branch of Russellian monism claiming that qualia are the only intrinsic properties of substances; a branch of Cartesian dualism postulating that physical substances have non-experiential intrinsic properties of which we may have no knowledge, alongside experiential intrinsic properties of which we can be certain of (i.e., one’s conscious experience) (Menon & Siddarth 2017). The qualia realist perspectives discussed in this paper opt for the first.

The Illusionist Position
Frankish’s theory of “illusionism” agrees with Chalmers’ claim that anti-qualia realists must provide a neuroscientific explanation for why qualia theories are so pervasive. The important difference is that illusionism insists that qualia properties are not “real” in an ontological sense of the term (Frankish 2016, 14). The illusionist approach is motivated by the success of third-person empiricism in dissolving our fallacious convictions about the natural world. When using the term “third-person,” I am referring to neurophysiological explanations that analyze qualia from a perspective that is external to the subject of experience. Looking back at the history of science, there is a strong case to be made for conducting science from a third-person perspective. Most notably, the shift from first-person empirical methodology, such as Aristotle’s Ptolemaic worldview, is what allowed invaluable paradigm shifts as Galileo’s Heliocentric worldview (Goff 2020).

In the following analogy, I will bolster Frankish’s conviction that adopting a third-person perspective when investigating qualia is advantageous. Now, imagine that you are looking out towards the horizon of the desert. In the far distance, you believe that you see a pool of water. Perhaps to your surprise, a physicist comes along and explains to you how your eyes are actually fooling you. They explain how when sunlight passes through two layers of air with distinct temperatures, the two air masses collide and appear like a mirror. Thus, the supposed
“lake” is actually a reflection of the sky above. This fascinating optical illusion is called “mirage” and is essentially analogous to how illusionists hypothesize about qualia. Because we are so radically misled by our immediate, perceptual representations of the natural world, how are we to be certain about our first-person intuitions about qualia-as-intrinsic? Frankish doubts the epistemic reliability of qualia impressions, like how Descartes considered that an “evil demon” could be deceiving his impressions of the external world. Thus, Frankish proposes that theories of consciousness should be informed by neuroscience, such that our theories of consciousness are no longer “radically at odds with that of the physical sciences” (Frankish 2016, 24).

The Panpsychist Challenge
My first criticism is that illusionism’s reliance on third-person empiricism does not adequately address the qualia realist belief that subjects of experience have inherent access to the nature of qualia. To provide a concrete example of the impenetrability of the qualia realist’s argument, I will be examining a qualia theory posed by Philip Goff, a contemporary qualia realist.

According to Goff’s “real acquaintance hypothesis,” individuals have direct, epistemic access to the intrinsic nature of qualia in virtue of “being” in an experiential state (Goff 2015, 3). For instance, Goff believes that one cannot be in the state of pain and not have epistemic access to the essence of pain. Therefore, the experience of pain is the essential property of pain and constitutes the “real definition” of pain. In Goff’s real acquaintance hypothesis, we see the influence of Descartes’ argument that one can doubt the external referent of experience—such as the neurophysiological basis of the pain experience—but we cannot doubt that we are having a conscious experience (Menon & Siddarth 2017, 411). In the same way, Goff claims that simply being in an experiential state of pain allows one to grasp the intrinsic nature of that state.

I will be considering ways that Frankish could possibly respond to Goff’s “real acquaintance hypothesis” based on the information he provides (Goff 2015, 3). First, Frankish might be tempted to refute that only subjects with “introspective mechanisms” possess the capacity to have perceptual illusions of qualia (Frankish 2016, 14). In this sense, Frankish’s refutation relies on cases where our perceptual faculties cause misleading misrepresentations of the nature of the world (e.g., the
Goff’s particular perspective on qualia can be broadly categorized under the umbrella term “panpsychism.” In brief, panpsychism encompasses a large sum of theories throughout the history of philosophy. Pansychist views profess that consciousness is fundamental and ubiquitous in the natural world. The most contentious element of panpsychism is the claim that inanimate objects are conscious. Panpsychists will go as far as claiming “thermostats are conscious,” an assertion that challenges how we colloquially use the term “conscious” in association with human cognition (Chalmers 1995). I will omit a detailed history of panpsychism because it is irrelevant to my following point about illusionism’s inability to adequately address panpsychism’s principles according to the merit of its own methodology.

Although Frankish’s addressment of panpsychism is brief, I argue that it shows an important way that illusionism fails to sufficiently address the essential components of qualia directly. Frankish simply claims that the neuroscientist’s inability to vindicate their assertions—such as thermostats having conscious experience—is justification for accusing them of fabricating a “fiction” (Frankish 2016, 32). Here, Frankish’s justification is problematic because it is antithetical to his anti-first-person intuitional approach. Supposedly, the key advantage of relying upon third-person neuroscientific explanations is to overcome the fact that our introspectively derived impressions about our conscious experiences are systematically misleading. After all, Frankish’s argument relies
heavily on the assertion that our attitudes towards theories of consciousness cannot rest on superficial attitudes about what theories of consciousness ought to be like. Instead, we ought to be persuaded by the theory that is the most “rationally compelling” (Frankish 2016, 37). With respect to that rationale, it is hypocritical for the illusionist to fight the panpsychist intuition by merely retorting their own intuition that anti-qualia realism can rely on advancements in neuroscience. The hypocrisy lies in the fact that, instead of relying on his methodology to dismantle “intrinsic subjectivity” as a cognitive illusion, Frankish seems to merely protest that intrinsic subjectivity is not a conveniently discernable concept for the purpose of neuroscientific research.

According to my analysis, the qualia realist who holds intrinsic subjectivity would have the upper hand against Frankish in the panpsychist debate. When the qualia realist claims “intrinsic subjectivity,” they are not only making a case about the ontology of qualia, but they are also simultaneously designating the constraints on how philosophy can accommodate qualia in a scientific worldview. From that perspective, de-bunking qualia realism is not a matter of acquiring a more detailed understanding of cognitive mechanisms, as Frankish would hope (Frankish 2016, 37). Adhering to the definition of “intrinsic subjectivity,” a qualia realist would say that even if a neuroscientist were to provide the most complete, neurophysiological explanation of our conscious experience, they would be failing to grasp his point about that conscious experience is necessarily grasped by the subject’s first-person experience, and not by any other means (Goff 2017, 7). To that extent, Frankish’s hands are tied, unless he is able to provide a neuroscientific account about why panpsychists are being systematically deceived about their beliefs that thermostats are conscious.

I argue that even if Frankish attempted to directly address the component of intrinsic subjectivity, his proposed explanation would fail to be compelling compared to the panpsychist. Recall, for illusionism to prevail against qualia realism, the anti-qualia realist must explain why our conscious experiences seem to have an additional quality that feels to be “intrinsic subjectively” but is actually an illusion. Frankish generally addresses the challenge of explaining our perception of qualia in neuroscientific terms as the “illusion problem.”
When speaking of the illusion problem, Frankish provides a general schema for how philosophy can interpret the qualia realist’s claims about the nature of conscious experience. When the qualia realist insists upon the ontological distinctiveness of consciousness, Frankish insists they are deceived by systematic illusions caused by their perceptual faculties. He uses the term “quasi-phenomenal property” as a place-holder term for the cognitive process underpinning the qualia realist’s beliefs about qualia (Frankish 2016, 16). For example, when qualia realists claim that their visual experience of red has the property of “phenomenal redness,” the illusionist says that physical quasi-phenomenal properties “trigger” introspective representations of phenomenal redness. In other words, the qualia realist’s perception of phenomenal redness is an illusory perception that subsists on top of their perceptual visual experience of red.

What is significant is that Frankish’s construal does not consist of showing how quasi-phenomenal properties are real. Rather, Frankish merely posits that when the qualia realist claims that they have “direct access” to the nature of qualia, they have no way of knowing that what they are referring to is a quasi-phenomenal property. Essentially, Frankish’s accusation can be summarized in the following question: how would the qualia realist know the difference between a phenomenal property (i.e., qualia) or a quasi-phenomenal property if they are cast under an illusion? (Frankish 2016, 9-20).

I agree with Frankish’s accusation that the anti-qualia realist would not be able to tell the difference if a quasi-phenomenal property is triggering their beliefs about qualia rather than a phenomenal property. However, that same logic can be used against illusionism as well, as Frankish himself cannot prove that quasi-phenomenal properties are causing systematic illusions about qualia. To that extent, the qualia realist’s argument that actual qualia are what cause beliefs about qualia are relatively more parsimonious. In comparison, the illusionist would have to make concessions to explain why our cognitive mechanisms cause us to have systematic illusions about non-existent qualia. Frankish makes no concrete attempt to address this problem of why we would have systematic illusions about the nature of our conscious experience.

The Private Language Conflict
In my preceding argument, I showed that Frankish does not
provide a compelling case that qualia’s constitutive property of intrinsic subjectivity can be explained as a cognitive illusion. In order to close the case that Frankish could not in fact pin-down intrinsic subjectivity on any other account, I will be re-imagining Ludwig Wittgenstein’s infamous “private language” thought experiment (Wittgenstein 1953, §244-§271). My purpose in this endeavor is to provide a convincing case that Frankish cannot simply rely on advancements in neuroscience to provide a coherent neuroscientific explanation of qualia’s intrinsic subjectivity as an illusion.

With reference to the “private language” thought experiment, let us imagine that a neuroscientist enters a room full of philosophers, each holding a box that cannot be opened. In order to make some sense of the situation, the philosophers agree on the assertion, “there is something called a ‘beetle’ inside of my box.” As the neuroscientist is a staunch adherent to third-person empiricism—over and above all—she believes that the philosophers cannot prove that there are “beetles” inside of any of their boxes. After reading Frankish, her game-plan is to discover the complex array of neural correlates that underpin the philosopher’s “beetle” illusion. Her trusted tool is the fMRI machine, as hopefully, that will allow her to map the philosopher’s first-hand reports about “beetles” onto their brain activity. Assuming she succeeds in finding these neural correlates, perhaps then the philosophers will concede that their “beetle” is nothing more than a complex array of neural firings.

Unfortunately for the neuroscientist, her enterprise is precluded by the philosophers’ definition of what beetles are. As she attempts to analyze the neural correlates of the philosopher’s belief propositions, the content of the philosopher’s first-hand reports proclaims that she cannot possibly see what lies inside the various boxes. Specifically, her problem is that the philosophers have agreed this concept is “intrinsic subjectively”—analogous to how philosophers conceive of qualia as exclusively accessible to the subject of experience. What exactly can the neuroscientist find the neural correlates of then? The challenge for the neuroscientist is that her only vehicle of explanation is the fMRI machine, but that does not equip her with the capacity to make a philosophical argument that third-person empiricism should be trusted over and above the philosopher’s claims. If her research program is under the supervision of Frankish, then her mode of explanation
is restricted within the bounds of what neural correlates can explain.

Of course, many philosophers have given strong theoretical arguments against the beetle-in-the-box dilemma that support Frankish’s convictions, but they do not rely on specific neuroscientific evidence whatsoever. Most infamously, the anti-qualia realist, Daniel Dennett, claims the beetle-in-the-box problem justifies canceling the “language game” altogether. Specifically, Dennett proclaims that qualia—like the beetle—has derived its meaning based on how it has been constructed via philosophical argumentation (Dennett 1998, 4). His interpretation scathingly undermines qualia’s “intrinsic” quality because language is inherently relational (i.e., extrinsic). On this basis, we can eliminate this concept from our ontology because we have no objective, third-person proof that qualia exist.

Unlike Dennett, Frankish does not outright dismiss the qualia realist’s feeling of qualia’s intrinsic subjectivity. He claims that, despite qualia not being “real,” our intuitions about qualia substantially reflect how our cognitive mechanisms evaluate conscious experience (Frankish 2016, 15). When confronting Chalmer’s meta-problem, Frankish believes that the anti-qualia realist must explain why the qualia realist holds the principle of intrinsic subjectivity from a neurophysiological perspective. Frankish is determined to confront such qualia concerns because, quite obviously, Dennett’s dismissal thirty years prior was unsatisfying for the qualia realists. In fact, philosophers continually re-imagine thought experiments like the “beetle-in-the-box” as a justification to push qualia realism more aggressively. They might consider, “if public language is extrinsic and relational, then how will I ever determine if what I refer to as a “beetle” is the same as yours?” According to qualia realists, questions about intersubjective experience are important and worth pursuing. Yet if the qualia realist claims the subject’s privileged access to their own qualia ultimately bears the truth of the matter, then the neuroscientist is blocked from offering further insight.

**Illusionism & The Meta-Problem**
Although I maintain that illusionism fails to address qualia’s constitutive component of intrinsic subjectivity, there is an advantage to Frankish’s willingness to accept qualia’s
constitutive features in his argument. Particularly, when facing Chalmers’ meta-problem, illusionism seems like the most agreeable anti-qualia realist position. Unlike Dennett who dismisses the epistemic reliability of the qualia realists’ claims, illusionism thoughtfully considers the constitutive features of qualia as described by the qualia realist. Frankish achieves this common ground by willfully accepting the core arguments like Chalmers’ metaphysical intuition. Returning to my mirage analogy, Frankish does not simply tell the person standing in the desert: “you do not actually see a lake in front of you because your experience does not exist as such.” Instead, Frankish argues that there ought to be a robust, causal explanation that does justice to the verisimilitude of the subject’s conscious experience of the mirage.

Since Frankish conveys a non-dismissive attitude towards qualia realism beliefs (e.g., qualia are irreducible), Chalmers says that he would identify as an illusionist if he was forced to pick another position (Chalmers 2018, 8). Appealing to qualia realists is integral because, quite obviously, they are the ones who need convincing. To be clear, I am not implying that Chalmers is the righteous authority who gets to make the final call about which solution to the meta-problem is ultimately “right.” Rather, I am identifying a key explanatory advantage of illusionism in its ability to advance the discussion by agreeing with the qualia realist’s definition. In this sense, Frankish could refute my argument by claiming that it is unsatisfactory for anti-qualia realists to dismiss the qualia realist’s reports about what their conscious experience feels like at face value. Frankish voices this motivation in his response to his fellow anti-qualia realists (e.g., Nicholas Humphrey, Peter Mandik)—as he argues that anti-qualia realists present an “inclusive attitude” towards the qualia realist’s claims about the nature of conscious experience (Frankish 2016, 16).

In response, I argue that Frankish’s “inclusive attitude” is also his greatest shortcoming, as he is unable to explain how his methodology can penetrate the concept of intrinsic subjectivity. Frankish’s whole theory of illusionism is based on the premise that neuroscience may eventually explain our beliefs about conscious experience (Frankish 2016, 16). Since Frankish does not have a neuroscientific explanation for why qualia realists believe that qualia are intrinsically subjective, he must rely on his intuition that qualia are not causing these beliefs. To that
extent, his response is equally as unscientific and unfounded in third-person empirical evidence as the qualia realist position that he criticizes. In other words, Frankish’s own perspective on qualia-as-intrinsic is not substantiated by the neuroscientific explanations which he holds as the golden standard for theories of consciousness.

Conclusion
In summary, I argue that Frankish’s theory of illusionism is ill-equipped for dismantling intrinsic subjectivity as a constitutive component of qualia. First, I explained the history of the qualia realism debate and why Frankish is compelled to confront Chalmer’s metaphysical intuition that qualia are intrinsically subjective. I argued that Frankish’s illusionist methodology fails to properly address qualia’s component of intrinsic subjectivity in two different contexts: Goff’s theory of panpsychism, a theory of consciousness that holds intrinsic subjectivity as a core feature, and my re-imagining of Wittgenstein’s “private language” thought experiment.

Ultimately, Frankish’s fatal error is his willingness to accept whatever qualia intuitions are thrown upon him and to merely rely on advancements in neuroscience to address them. The tragic fate for anti-qualia realism is that advancements in neuroscience may eventually provide substantive explanations for every other constitutive feature of qualia (e.g., ineffability, irreducibility). Nevertheless, the illusionist’s explanation for intrinsic subjectivity will always seem like it is missing the point of what “intrinsic subjectivity” truly means.
Bibliography


