Mystery and the evidential impact of unexplainables

Dominik Klein & Matteo Colombo

Abstract   What's the evidential impact of learning that something is a mystery? To answer this question, we first explicate the notion of a mystery in terms of unexplainability. After distinguishing different ways in which something can be unexplainable, we develop a test to evaluate the evidential impact of two distinct types of unexplainables: symmetrical and a-symmetrical unexplainables. We argue that only a-symmetrical unexplainables have evidential impact. We finally clarify how our explication of mysteries as unexplainables complements existing accounts of abduction and contributes to the literature on the mystery of consciousness.
Keywords: Explanation; Confirmation; Unexplainability; Mystery; Credences; Consciousness

1 Introduction

We are surrounded by mysteries. Free-will, intentionality, mental causation, personal identity and the origin of life: these are some of the phenomena that philosophers have called 'mysteries.' Several scientific puzzles have been characterised as mysteries too. Examples include the Big Bang, the asymmetry of matter and antimatter in the observable universe, the evolution of sex, the evolution of language, and the existence of conscious experience, which "is at once the most familiar thing in the world and the most mysterious" (Chalmers 1996, 3).

Although so many phenomena and events have been characterised as mysteries, the concept of a *mystery* remains opaque. Epistemologists and philosophers of science have paid no attention to questions like: What does it mean that something is a mystery? Whence does a mystery derive its mysteriousness? What is the evidential impact of learning that something is a mystery? Given the apparently widespread usage of the concept of mystery, addressing these questions should bear on several on-going debates in philosophy and science, and will help illuminate a neglected aspect of our epistemic practices.

In this paper, we explicate the concept of mystery in terms of *unexplainability*. After regimenting the question of how mystery should relate to credence (Section 2), we provide a taxonomy of mysteries as unexplainables (Section 3). We then present a test mapping these types into two cases—i.e., symmetric and a-symmetric unexplainables—and we argue that only a-symmetric unexplainables have evidential impact (Section 4). We conclude by putting these results into a broader philosophical perspective, illustrating their significance in light of the "mystery of consciousness" (Section 5).

2 Mysteries as unexplainables

The term 'mystery' (Greek: *musterion*: something hidden, secret, or closed) is used in a variety of ways in a wide range of contexts.[1] Generally, predicating of something that it *is a mystery* expresses

---

[1] Tuggy lists five different meanings of "mystery" in the literature on the doctrine of the Trinity: "[1]…a truth formerly unknown, and perhaps undiscoverable by unaided human reason, but which has now been revealed by God and is known to some… [2] something we don't completely understand… [3] some fact we can't explain, or can't fully or adequately explain… [4] an unintelligible doctrine, the meaning of which can't be grasped….[5] a truth which one should believe even though it seems, even after careful reflection, to be

some lack of knowledge or understanding. For example, when I say 'String theory is a mystery to me' I am expressing that I do not understand, or do not know much about string theory. When physicists say 'The Big Bang poses several mysteries', they are typically expressing that the Big Bang theory raises several puzzles in cosmology and astrophysics, some of which have been resolved while others are still awaiting a solution.

In particular, the term 'mystery' is often used to express a lack of knowledge in reply to a request for explanation. Suppose, for example, that Georgi asks Zhasmina why the expansion of the Universe has been accelerating, and Zhasmina asserts in response that the acceleration of the expansion of the universe is a mystery (to her). Zhasmina's reply can plausibly be understood as saying that she does not possess an explanation for why the expansion of the Universe has been accelerating. If explanations are answers to why-questions (van Fraassen 1980; Colombo 2016; Skow 2016, Ch. 2), then Zhasmina is asserting that she does not possess an answer to Georgi's why-question. Since "possessing an answer" is just knowing the answer, Zhasmina is asserting that she does not know the answer to Georgi's why-question. She may lack relevant knowledge of physics and cosmology, or ignore that the Universe's expansion is accelerating. While Zhasmina does not possess an answer to Georgi's why-question, a cosmologist may possess such an answer. The cosmologist may not possess an answer right away, but she may acquire it with sufficient research, experiment, and observation. In any case, Zhasmina's assertion in reply to Georgi's why-question should have no evidential impact on Georgi's beliefs.

On a different understanding, however, what Zhasmina is asserting is not just that she does not know the answer to Georgi's why-question. Zhasmina may be asserting that nobody presently possesses an explanation for why the expansion of the Universe has been accelerating. She may express that nobody will ever possess this explanation; or she may even express that no human can in principle possess this explanation, and so that no human could possibly answer Georgi's question.

On some of these interpretations, it is not obvious that Zhasmina's answer should have no evidential impact on Georgi's beliefs, and in particular on his belief that the expansion of the Universe has been accelerating. After all, learning that no established theory in physics can *ever* explain this phenomenon may reasonably lead Georgi to doubt that the Universe's expansion is actually accelerating. Georgi may revise his beliefs and conclude that the recorded rate of cosmic expansion is probably mistaken, merely apparent, or illusory. When and how, then, should learning Zhasmina's answer impact Georgi's beliefs? More generally, when and how should learning that something is a mystery impact one's belief that that something is the case?

The tools of Bayesian epistemology offer one way to sharpen, regiment and answer these questions (Bovens and Hartmann 2003). Suppose that p is some proposition that you believe, and *Prob* (p) your subjective, probabilistic credence that p is the case. You are not initially ready to doubt that p—after all, you do believe that p. Yet, you ask "Why (or how) is it the case that p?" to a source Z, whom you believe to be very knowledgeable, reliable, and sincere. Z replies to you that if p is true at all, then why p is a mystery. In providing you with this answer, Z does not give you any direct information about the truth or falsity of p. The only information you receive concerns the mysteriousness of p in case p is true. Because in this context the term 'mystery' is used to express a

---

impossible and/or contradictory and thus false" (Tuggy 2003, 175–6). As it will be clear, our contribution in this paper is focused on [3].

lack of knowledge in reply to a request for explanation, we can say that if the proposition p that you believe is a mystery, then p, in some sense, cannot be explained,[2] that is: p is, in some sense, *unexplainable*.[3]

There are several ways in which Z's reply can impact your beliefs. Your belief about what a mystery is can change; your belief that the source Z is knowledgeable, reliable, and sincere can change; your degree of confidence in some shared background theory that governs people's explanatory practice can change; and, lastly, your degree of confidence that p can change. Here we focus on the latter question: Should your credence in p change, after you learn that if p, why p is *unexplainable*?[4] Put differently, the question we set out to answer is the following: Assume you learn that p, if true, is unexplainable without receiving any further new information about the truth or falsity of p. Should this lead you to (rationally) change your credence in p?

This question concerns a specific type of evidence that can count against some belief of yours. In particular, this question relates in interesting ways to the issue of whether explanatory considerations have confirmation-theoretic import or not, which is central to debates about the status of abductive reasoning (Harman 1965; van Fraassen 1989; Douven 1999; Lipton 2004).[5] The basic idea underlying explications of abduction is that the explanatory success of a hypothesis with respect to an *explanandum* is part of the evidence that should lead one to raise her credence that the explanatory hypothesis is true (Douven 2011). However, discussions about the status of abduction have paid no attention to the complementary question of whether explanatory *failure* with respect to an *explanandum* is part of the evidence that should impact one's degree of credence in the

---

[2] Our analysis does not hinge on any particular account of explanation. We assume that all *explanantia* are propositions or sets of propositions. An explanans may describe some law, argument pattern, mathematical proof, cause, or mechanism that offers an explanation of the *explanandum* of interest. We also assume that *explananda* are propositions describing phenomena, observations, or events. *Explananda* are not assumed to be true.

[3] It may be argued that this way of setting up the problem assumes that the *principle of sufficient reason* is false. If you accept this principle, you will accept that: For every p, there must be an explanation why p is the case. In other words, if you accepted the principle of sufficient reason, then you should think that the testimony that *p, if true, is a mystery* implies the falsity of p. However, the connection between *explainability* and the principle of sufficient reason is in fact more complicated than this. One could consistently endorse the principle of sufficient reason and the claim that something is unexplainable. In a weak sense of *unexplainability*, one might say: "I endorse the PSR and so ultimately there must be an explanation of p, though I will never be able to explain it" (i.e., p is unexplainable for me). In a stronger sense (of an explanation being epistemically inaccessible to creatures like us), someone could consistently maintain: "I endorse the PSR and so ultimately there must be an explanation of p, though I also believe that the explanation is beyond human ability to discover / know / comprehend." We are grateful to a referee for pointing this out to us.

[4] We interpret the *if* in "If p, why..." subjunctively, as expressing that "If p is or was the case, then why or how so is or would be a mystery." It should be emphasised that this statement is not *logically* entailed by the falsity of p.

[5] Hempel (1965) distinguishes between "explanation-seeking why-questions" and "reason-seeking why-questions." He writes: "An appropriate answer to the former will offer an explanation of a presumptive empirical phenomenon; whereas an appropriate answer to the latter will offer validating or justifying grounds in support of a statement... any adequate answer to an explanation-seeking question 'Why is it the case that *p*?' must also provide a potential answer to the corresponding epistemic question 'What grounds are there for believing that *p*?'"(335). In the light of Hempel's distinction, our topic can be reformulated as follows: Suppose that an adequate answer to an explanation-seeking why-question about *p* cannot be provided. Does this fact give us information relevant to answer the corresponding reason-seeking why-question?

explanandum. In what follows, we bring to the fore this question, and explicate whether and under what conditions it is epistemically rational to revise your credences in such a way that:

$$Prob \text{ (p | p is unexplainable)} < Prob \text{ (p)}$$

We begin by distinguishing different senses in which a proposition can be said to be unexplainable. It is worth pointing out that our treatment is not intended to cover all cases, in which ordinary people call something a 'mystery' in order to express their lack of knowledge about whether a proposition has been successfully explained, or to express, more generally, their lack of understanding about certain events and phenomena. Our treatment is intended to apply to some epistemically interesting cases, widely discussed in philosophy, theology, and science, that involve a modally robust notion of *unexplainability*.

3 A taxonomy of unexplainables

In a general sense, to say that a proposition p is unexplainable is to say that no epistemic agent in our world can ever adequately answer the explanation seeking question 'Why (or how) p?'

Many ordinary uses of 'mystery' refer to the fact that an individual or a community do not see a way to explain an accepted proposition p, given their current explanatory resources. Such resources may not conflict with p; they may even make p more likely. Because the explanatory resources available at that point in time fall short of providing an adequate answer to why p, p remains unexplained at that time. But saying that a proposition is unexplained is different from saying that the proposition is unexplain*able*.

Although the proposition that the tides are higher during a full moon was unexplained till the 1600s, it was explainable. It was, in fact, explained when Newton published his *Principia* in1687. Newton showed that the gravitational laws he provided explain, among other phenomena, the correlation between the behavior of the tides and the position and phase of sun and moon. So, given some point in time, if a proposition p does not have an adequate explanation at that time, it does not follow that p is unexplainable. While many ordinary uses of 'mystery' refer to events or phenomena that remain unexplained given our current resources, for our purposes we can leave this usage on the side, and focus on a modally more robust usage.

If nobody happens, has happened, and will ever happen to possess an adequate answer to the question 'Why p?', then p *could* be unexplainable. There are at least three cases to distinguish here. Just two of these cases concern the modally robust sense of unexplainability, in which we are interested.

In the first case, p remains unexplained because nobody has and will ever have the interest, the motivation, the curiosity, or the material and theoretical resources to find out about why p is the case. The proposition p is not genuinely unexplainable, but, as a matter of contingent fact, nobody will ever possess an answer to 'why p?'. Thus, p will never be adequately explained. This is, of course, not a case of genuine unexplainability, as an adequate answer to 'Why p?' is actually epistemically accessible to creatures like us.

In the second case, nobody has possessed and will ever possess an adequate answer to 'Why p?', because any such answer contingently lies on one side of human cognitive capabilities. Given the kind of epistemic agents we are, with bounded cognitive and reasoning capacities, an adequate answer to the question 'Why p?' is epistemically inaccessible to us in this case, as this answer lies beyond our cognitive reach.

For example, Chomsky (1975, Ch. 4; see also Hauser et al 2014) has famously argued that how and why our linguistic computations and representations evolved may well remain unexplainable to creatures like us. Another famous example includes the mystery of consciousness, where some have argued that the answer to the question 'Why (or how) do brain processes produce conscious mental states?' is in principle inaccessible to epistemic agents like us (McGinn 1991). If these authors are right in their diagnosis, then our cognitive capacities would not be fit to answer, or to even comprehend adequate answers to these questions. Hence, why some mental states are conscious, or why language evolved would forever remain unexplained to creatures like us.

The third case concerns answers to why questions that are epistemically inaccessible, because of the nature of the *explananda* they seek to explain. Some of these *explananda* correspond to brute phenomena or facts that have no explanation. For example, there may be no answer for why the laws of nature have the form they do. But if there is in fact no explanation for p, then nobody can answer the question 'Why p?'[6]

Some other *explananda* do not correspond to brute phenomena, but are incoherent with some shared standard of explanatory reasoning or with some established background theory. Shared standards of explanatory reasoning and established background theories are routinely used by members of an epistemic community to ask why-questions, to answer why-questions, and, more generally, to distinguish adequate from inadequate explanations.

While explanatory standards and background theories may change over time, and may vary across different communities, inquirers will share at least some explanatory standard and some relevant background theory with their informants. These standards and background beliefs plausibly include the principle that contradictory propositions cannot both be true in the same sense at the same time, the principles that causes explain effects but not vice versa and that properties of parts explain aggregate properties but not vice versa, and the proposition that no event in nature can exceed the productive power of nature. *Explananda* that describe miraculous phenomena are obviously incoherent with these principles.[7] When an *explanandum* p is incoherent with some shared standard of explanatory reasoning or background theory, the question 'Why p?' cannot be adequately answered given that standard or theory.

---

[6] This is the only sense of unexplainability in our taxonomy, which plausibly precludes acceptance of the principle of sufficient reason. It does so, because brute facts are those that have no explanation, as distinct from those that merely cannot possibly be explained by creatures like us. It is worth noting that people may rationally ask why-questions about brute facts, as they may not know that those facts are in fact brute. Even when people *know* that a fact is brute, they may still be interested in figuring out why it is brute. Sometimes, however, it might be misleading for us to go telling people that a fact, which is known to have no explanation, is a mystery. One might argue that it is misleading, because when we say that something is a mystery we are not merely claiming that it's unexplainable, but also that there may be an explanation—which could indeed not ever be knowable/comprehensible to anybody. We are grateful to a referee for helping us clarify these complications about the relationships between mystery, brute facts, and the principle of sufficient reason.

[7] St. Thomas Aquinas (*Summa Contra Gentiles*, III) says: "those things are properly called miracles which are done by divine agency beyond the order commonly observed in nature *(praeter ordinem communiter observatum in rebus)*." This characterisation of a miracle is entailed by Hume's famous definition that a miracle is "a violation of the laws of nature" (Hume 1748/2000). Voltaire (1764/1901, p. 272) provides a stronger definition of a miracle as "the violation of mathematical, divine, immutable, eternal laws. By the very exposition itself, a miracle is a contradiction in terms: a law cannot at the same time be immutable and violated."

In short, there are a variety of reasons that may underly the modally robust sense of unexplainability we are interested in. These roughly fall in two classes. In the first class, unexplainability is due to the subject matter of p and its relation to us as epistemic agents. Unexplainables in this class might describe a brute fact, or some topic whose explanation is inaccessible to epistemic agents like us. The second class contains those cases, where unexplainability is triggered by the *explanandum*'s relation to relevant explanatory standards and background beliefs. Unexplainables in this class might, for example, be logically incoherent, as the laws of logic are part of the background standards, or violate the principle that nothing comes from nothing.

4 Unexplainables and credence

There is a simple test that distinguishes two, modally robust types of unexplainability. Recall that your reliable, knowledgeable, and sincere source Z informs you about the unexplainability of p while remaining silent about the truth value of p. The information you receive is that:

(a) If p is true, then 'Why p?' is unanswerable.

Since p is a proposition, we can form its negation *not*-p, and ask 'Why *not*-p?' Z could give you this answer:

(b) If *not*-p is true, then 'Why *not*-p?' is unanswerable.

Depending on the content of p, there may or may not be a connection between the statements (a) and (b) as asserted by Z.[8]

To illustrate, suppose that you acquired the belief that there is more landmass on the northern hemisphere of planet Earth than on its southern hemisphere. You're not ready to doubt it, but you want to know why that is the case. Assume that a knowledgeable and sincere source tells you that the question 'Why is there more landmass on the northern than on the southern hemisphere?' is in principle unanswerable. Nobody may ever explain why the majority of land mass is in the northern hemisphere. If this question is unanswerable, also the converse of the question will be unanswerable: even if it held true, nobody would ever be able to explain why it is *not* the case that the majority of land mass is in the northern hemisphere. In short, the content of the *explanandum* proposition allows for a symmetry in unexplainability; and a knowledgeable, reliable and sincere source should be prepared to assert both (a) and (b).

For the other extreme, consider instead some other proposition such as the proposition that Jesus materially turned water into wine at the marriage of Cana. Many people believe that this proposition is true, though it apparently clashes with some of their fundamental theoretical and logical commitments. While they do not doubt it, they are also ready to inquire for an explanation: 'Why or how is it that Jesus did that?' Not only all knowledgeable chemists, but also most lay people—including many that believe that proposition—will subscribe to the statement that 'If it is true that Jesus transformed water into wine at the marriage of Cana, then that is unexplainable.'

---

[8] Recall again that we read the *"if p…"* subjunctively. In particular, there is no logical connection between (a) and (b).

However, few people would be prepared to support the corresponding assertion (b), that 'If it is *not* true that Jesus transformed water into wine, then why so is unexplainable.' The content of p is such that there is an asymmetry between (a) and (b). If p is true, then why or how p is unexplainable. But if *not*-p is true, then why or how so is *not* unexplainable, since an answer to this question coheres with a shared standard of explanatory reasoning like the principle that no event in nature can exceed the productive power of nature. So, in this case, a knowledgeable, sincere and reliable source should be prepared to assert (a) but not (b).

The upshot is that the question 'If not-p were true, would not-p be unexplainable?' allows us to distinguish between two classes of unexplainables: *symmetric* unexplainables, where both p and not-p, if true, are unexplainable, and *asymmetric* unexplainables, where only p is unexplainable.

These two classes map to the distinction between the two cases of genuine unexplainability we identified above. When the unexplainability of *p* is due to limitations of humans' cognitive capacities, then these limitations hinge on the subject matter of *p*, not its truth value. If *p,* if true, is unexplainable, then also *not-p i*s and we have a symmetry between (a) and (b). Instead, when the unexplainability of a proposition does not depend on limitations of cognitive capacities but on logical and explanatory relations between the proposition and some shared explanatory standard or background belief, then this cannot at the same time be true of *not-p*. If p is logically or explanatorily incoherent with some shared explanatory standard or background belief, then not-p is tautological given that standard or background belief. And if both *p* and *not-p* simultaneously conflict with some shared explanatory standard or background belief, then the latter must be internally inconsistent or explanatorily incoherent.[9] In short, symmetric unexplainables correspond to subject matters inaccessible to epistemic beings like us while asymmetryic unexplainables arise from an incompatibility between p and some relevant reasoning standard or background belief.

4.1 How should unexplainables impact your credences?

The symmetric-asymmetric distinction is relevant to addressing how you should update your credence in p after learning that p is unexplainable. Armed with the tools of Bayesian epistemology, we posit that a rational agent's credence function *Prob*( ) satisfies the axioms of probability, and that agents' belief update is carried out via Bayesian conditionalization. Rational agents' credences should be such that their probabilities of mutually exclusive propositions should sum up to at most one, and mutually incompatible propositions should receive a joint probability of zero.

Consider asymmetric unexplainables. Let us denote the set of relevant standards of explanatory reasoning and background theories by S, where S is a conjunction of many, interrelated, individual propositions. The information that 'If p is true, then why or how p is unexplainable' yet that it is *not* the case that ' If *not*-p is true, then why or how *not*-p is unexplainable' entails that we are in the second of the two classes we distinguished above. The unexplainability of p cannot be grounded in the subject matter of p and in its relation to us as

---

[9] In the case of logical inconsistency, this is related to a well-known principle of classical logic. Whenever we can derive a contradiction from a set of propositions *F* together with *p* and we can derive a contradiction from *F* together with *not-p*, then we can derive a contradiction from *F* alone. So if both *p* and *not-p* are contradictory with F, then F must be internally inconsistent. The argument roughly goes as follows. Assume *F*. By logical reasoning, either *p* or *not-p* must be true. If *p* is true, we have *F* and *p*, which, we assumed, together derive a contradiction. If *not-p* is true, we have *F* and *not-p* which again derive a contradiction together. Thus, assuming *F*, we can *always* derive a contradiction. This exactly says that *F* is inconsistent. In case of incoherence, a similar argument applies.

epistemic agents alone. Rather, the unexplainability of p is grounded in its relationship to S. Learning that if p is true, then why so is unexplainable, conveys the information that p and S are mutually incompatible, or at least, that p is very unlikely given S. By the laws of probability, the agent should revise her credences in light of this information, in such a way that the conjunction p & S gets assigned a probability of zero,[10] that is:

$$Prob \text{ (p \& S| p is unexplainable)} = 0$$

Where *Prob* (*x*| p is unexplainable) stands for the probability of *x after* revising with the new information that p is unexplainable. By the same argument, the revision will have to ensure that the probabilities of p and S sum up to at most one, since p and S are mutually exclusive:

$$Prob \text{ (p | p is unexplainable)} + Prob \text{ (S | p is unexplainable)} \leq 1$$

If we turn our attention to your prior credences, then it seems plausible that, before being informed about the unexplainability of p, you should have a high degree of credence in S. After all, S consists of background standards and theories deeply engrained in processes of explanatory reasoning that allow you to ask, answer, and reason about why-questions. Furthermore, your prior credence in p should be sufficiently high. After all, if you already disbelieved that p, then it would be unlikely that you will inquire about why p. Together these two considerations imply that your prior credences are such that:

$$Prob \text{ (p)} + Prob \text{ (S)} > 1 \qquad (1)$$

One independent way to see why this should hold is through a threshold theory of belief. If *t* is a fixed threshold for converting probabilistic credences into belief simpliciter, then you believe *p* (simpliciter) if and only if $Prob(p) > t$.[11] for some *t > 0.5*. Assume that you believe (simpliciter) in your background theory S, that is: $Prob(S) > t$. If you enquire about *p*, you will most probably also believe p, hence $Prop(p) > t$. Hence $Prop(p) + Prop(S) > 2t > 1$, thus validating(1). But of course, your reason for asking might be that you are somewhat suspicious about *p*. In that case, you might prefer to suspend judgment about p, prior to your question i.e. you neither belief *p* nor not-*p*. The latter implies that $Prob(not\text{-}p) < t$ and hence $Prob(p) > 1\text{-}t$. But then again we get that $Prop(p) + Prop(S) > t + 1 - t = 1$, in line with (1). The only case where the sum of $Prob(S)$ and $Prob(p)$ is smaller or equal to 1 is when you actively disbelieve *p*, you believe that not-*p*. In all other cases, in particular when suspending judgment about p, the sum $Prob(S) + Prob(p)$ will be strictly larger than 1. Since it seems plausible that epistemically rational agents would typically

---

[10] For simplicity, we focus on the case where p and S are incompatible. Where p is highly improbable given S, i.e., *Prob* (p|S & *p is* unexplainable.) ≈ 0, a similar analysis applies. By Bayes's theorem, we have that *Prob* (p | S & *p is* unexplainable) = $Prob$(p & S/ *p is* unexplainable)/$Prob$(S/ *p is* unexplainable). Since the denominator of the right hand side is smaller than one, the left side can only be close to zero if also the nominator of the right side is. Thus, upon learning that *Prob* (p|S & *p is* unexplainable) ≈ 0, revision has to ensure at least that *Prob* (p &S / *p is* unexplainable) ≈ 0. This triggers very much the same arguments as the case "*Prob* (p &S / *p is* unexplainable) = 0" treated here.

[11] The existence of such a threshold follows, e.g., from the Lockean Thesis of belief (Leitgeb 2014).

refrain from asking the question 'Why p?' to a knowledgeable, reliable, and sincere source if they actively believe p be false, the assumption is justified that:

$$Prob \ (p) + Prob \ (S) > 1$$

In this way, learning that p and S are mutually incompatible (or that p is very unlikely given S) should lead epistemically rational agents to revise their degrees of credence to ensure that:

$$Prob \ (p \mid p \text{ is unexplainable}) + Prob \ (S \mid p \text{ is unexplainable}) \leq 1$$

Such revision can be done by reducing your credence in S, the credence in p, or both.

When should you revise your credence in S? Let us stipulate that your inquiry 'Why p?' is not the first explanatory question you ask. Plausibly, you have engaged in a significant number of explanatory endeavours before, along with other members of your community. Since your explanatory standards and background theories S have played some role in many of these endeavours allowing you to successfully acquire answers to several why-questions, it will be reasonable for you not to give up your credence in S too easily. After all, many of your current beliefs have been acquired by relying on these standards of explanatory reasoning and background theories. The standards and theories are so implicated with many of the things that you believe that to abandon them would be to give up on or jeopardise many propositions that you currently take to be true. Revising your credence in S may then trigger a significant amount of epistemic labour, involving far reaching epistemic reassessment and possible epistemic losses. It is surely reasonable not to revise S too easily, given that far less radical alternatives are available (Quine 1951).

Another reason why you should not give up S too easily is that you share these reasoning standards and background theories with other members of your epistemic community. It is the social character of these epistemic standards and theories that allows you and other members of your community to engage in joint explanatory endeavors. So, decreasing your credence in S will compromise future, joint explanatory inquiries with members of your epistemic community in default of a new set of shared standards or background theories. Thus, the social embedding of S provides additional reason not to lower your credence in S.

Lastly, the facts that S is shared with other members of your community and governs your explanatory practices bear on how firm your belief in S should be. If members of your community have relied on S to acquire many beliefs and in particular many new explanations, then the epistemic adequacy of S and of its individual elements should have been already extensively tested. Whenever members of your epistemic community evaluated the quality of an explanation acquired through S, they have put under test the reliability of S or of some of its elements. To the extent that S successfully passes this kind of testing time and again, belief in S becomes firmer. It would then be unreasonable to significantly revise your firm belief in S, rather than revising credence in p, after you learn that the latter belief of yours is incoherent with S (cf. Skyrms 1980).

While these considerations should make you cautious to lower your credence in S, we rely on a weaker assumption in our analysis of how unexplainability should impact credence. We allow that, after learning that p is unexplainable, you may lower your credence in S, as long as not all weight of revision is borne by S. Plausibly, at least some weight should be borne by your belief that p, which justifies the conclusion that:

$$Prob \text{ (p | p is unexplainable)} < Prob \text{ (p)}$$

We note that this result follows even if one rejects that S has any special epistemic status different from p such that one should be cautious in revising the credence in S. Several belief revision policies would prescribe to lower your credence in p upon learning that p is unexplainable, even when S is not granted special epistemic status.

Take *normalization*. This belief revision policy prescribes that the weight of revision should be borne equally by $Prob(p)$ and $Prob(S)$:

$$Prob \text{ (p|p is a mystery)} =$$

and likewise for S. Since we assume the denominator of the right side, $Prob(p) + Prob(S)$ to be larger than 1, it follows that $Prob(p \mid p \text{ is a mystery}) < Prob(p)$. So, normalization prescribes to lower your credence in p upon learning that p is unexplainable.

In summary, the asymmetry in the pattern of responses that the source Z should give to the questions of 'Why *p*?' vs. 'Why not-*p*?' brings to the fore a conflict between *p* and some shared background theory or explanatory standard *S*. This conflict provides you with evidence, which bears on the truth value of *p*. This evidence should at least lead you to lower your credence that p.

Let's finally turn our attention to symmetric unexplainables. In this case, a knowledgeable, reliable and sincere source is prepared to assert that both *p* and *not-p*, if true, are unexplainable. Hence, learning about the unexplainability of *p* does not inform you about some incompatibility between *p* and S[12]; and hence does not provide you with evidence to reduce credence in *p*.

As explained above, examples of the symmetric kind of unexplainability may include *explananda* such as the proposition that there is more landmass on the northern hemisphere of planet Earth than on its southern hemisphere, and the propositions that human linguistic computations and representations emerged from a uniquely human genetic endowment (Hauser et al 2014) and that consciousness emerges purely from brain activity (McGinn 1991). The *explanandum* proposition in these examples is neither self-contradictory, nor is it seemingly incoherent with some other shared standard of explanation or background theory. Where the unexplainability of a proposition *p* depends on limitations of human epistemic reach, your source's pattern of responses will not provide you with new evidence about the truth value of *p*, and so your credences should remain unchanged. After all, finding an answer to the question 'Why *p*?' and finding an answer to the question 'Why *not-p*?' would be equally hard epistemic tasks. Both answers would be equally epistemically inaccessible to us. Due to this symmetry, the information that p, if true, is unexplainable is unrelated to the truth value of p. If the unexplainability of *p* is unrelated to its truth value, you will have no reason to revise your prior credence that p.

5 The mystery of consciousness
Our discussion so far has offered an explication of *mystery* in terms of *unexplainability*, and an argument that, in some specific cases, learning that something is unexplainable should have

---

[12] S cannot, on pain of inconsistency, be incompatible with both *p* and *not-*p. For if it was, it would be the case that both, *S&p* and *S&not-p* entail a contradiction. By classical logical reasoning, then also *S&(p v not-p)* entails a contradiction. But *since (p v not-p)* is a tautology, this is equivalent to saying that *S* alone entails a contradiction, i.e. *S* is inconsistent. The argument for the case that *p* and *not-p* are merely highly improbable, given *S* is similar.

evidential impact. We conclude by clarifying how these results matter to several ongoing debates in philosophy.

The question addressed by this paper is closely related to issues about the nature and rationality of abduction, which assigns a special confirmation-theoretic import to explanatory considerations. In the last decades, the debate has focused on whether explanatory success is a (defeasible) mark of truth (Harman 1965; Van Fraassen 1989; Lipton 2004). No attention has instead been paid to the related questions of whether and under which conditions explanatory *failure* is a (defeasible) mark of falsehood. If our analysis of unexplainability is right, then there is a connection between persistent explanatory failure and falsehood. Specifically, we should lower our credence in propositions that we have reason to believe present an a-symmetric unexplainability. The case of consciousness illustrates this point.

McGinn (1991) claims that consciousness is a mystery[13] because we humans are simply not equipped to find an explanation for why or how brain states and activity give rise to conscious experience (see also Chomsky 1975). The point is not that the existence of consciousness would breach accepted standards of explanatory reasoning, neither is the point that consciousness cannot be given a functional characterisation. The point is that we humans lack the conceptual repertoire necessary for answering questions about the relation between properties of the brain and conscious experience. But if we lack these cognitive resources, then both the proposition that there is some neural property N such that N produces conscious experience, and the proposition that there is no neural property N such that N produces conscious experience will remain unexplainable. Possible facts about the relation between certain neural properties and consciousness will be epistemically inaccessible to us humans as are facts about the relation between the property of having a mass and inertia to pigs. We would face a case of *symmetric* unexplainability, where both *p* and not-*p* would remain unexplainable to us. If this is right, and because of this kind of epistemic inaccessibility, our prior credence in the proposition that there is some neural property N such that N produces conscious experience should not be affected by the new information that that proposition is unexplainable.

Chalmers's (1996) diagnosis for the mysteriousness of consciousness differs from McGinn's. According to Chalmers, consciousness is a mystery because of its incoherence with widely shared standards that govern our epistemic and explanatory practices. Consciousness cannot be explained like any other complex property (e.g., genetic dominance or inflation) in terms of physically realized functional properties, because it cannot be captured in physical or functional terms. Consciousness would depend on non-physical properties; and so, knowing all physical facts would not suffice for knowing everything there is to know to answer any why-question about consciousness. Furthermore, since consciousness cannot be reduced to physical properties, its explanation would be incoherent with the principle of the *causal closure of the physical*, according to which every event that has a cause has a physical cause. If this type of incoherence (with functionalism, physicalism, and the causal closure of the world) is the source of mysteriousness for consciousness, then the proposition that there is some neural property N such that N produces conscious experience, and the proposition that there is no neural property N such that N produces conscious experience are *not* equally unexplainable. The former but not the latter coheres with shared standards of explanatory reasoning, and we would face an instance of asymmetric

---

[13] Strictly speaking, consciousness itself isn't mysterious in our sense because it's not a proposition.

unexplainability. If this is right, then our prior credence in the proposition that there is no neural property N such that N produces conscious experience should be lowered after learning that this proposition is unexplainable. Put bluntly, learning that there is a genuine explanatory gap in consciousness studies gives us reason to be *sceptical* about the existence of irreducible *qualia*.

References

Aquinas, St. Thomas (1975), *Summa contra Gentiles*, James F. Anderson (trans.), Notre Dame: University of Notre Dame Press.

Bovens, L. and Hartmann, S. (2003). *Bayesian Epistemology*, Oxford: Clarendon Press.

Chomsky, N. (1975). *Reflections on Language*. New York: Pantheon Books.

Colombo, M. (2016). Experimental Philosophy of Explanation Rising: The case for a plurality of concepts of explanation. *Cognitive Science*. doi: 10.1111/cogs.12340

Douven, I. (2011) Abduction. In E. Zalta (Ed.), Stanford Encyclopedia of Philosophy. <http://plato.stanford.edu/entries/abduction/

Douven, I. (1999). Inference to the best explanation made coherent. *Philosophy of Science* (Proceedings), 66 , S424–S435.

Harman, G. (1965). The Inference to the Best Explanation. *Philosophical Review*, 74: 88-95.

Hauser, M. D., Yang, C., Berwick, R. C., Tattersall, I., Ryan, M. J., Watumull, J., ... & Lewontin, R. C. (2014). The mystery of language evolution. *Frontiers in psychology*, 5, 401.

Hempel, C. G. (1965). Aspects of scientific explanation. In C.G. Hempel, *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, New York: Free Press, pp. 331-496.

Hume, D. (1748) *An Enquiry Concerning Human Understanding*, Tom L. Beauchamp (ed.), New York: Oxford University Press, 2000.

Leitgeb, H. (2014). The stability theory of belief. *Philosophical Review*, *123*(2), 131-171.

Lipton, P. (2004). *Inference to the Best Explanation*, (2nd ed.), London: Routledge.

McGinn, C. (1991). *The Problem of Consciousness*. Oxford: Blackwell.

Quine, W. V. O. (1951) Two Dogmas of Empiricism. Reprinted in *From a Logical Point of View*, 2nd Ed., Cambridge, MA: Harvard University Press, pp. 20–46.

Skow, B. (2016). *Reasons Why*. Oxford University Press.

Skyrms, B (1980), Higher Order Degrees of Belief, *Essays in memory of F.P.Ramsey* Cambridge, CUP, 109-37.

Tuggy, D. (2003). The unfinished business of Trinitarian theorizing. *Religious Studies*, *39*(02), 165-183.

van Fraassen, B. (1989). *Laws and Symmetry*, Oxford: Clarendon Press.

van Fraassen, B. (1980). *The Scientific Image*. Oxford University Press.

Voltaire (1764). *Philosophical Dictionary*, in *The Works of Voltaire*, vol. 11, New York: E. R. DuMont, 1901.