Philosophical Methodology in Modal Epistemology

Dana Goswick, University of Melbourne

Published online: 30 January 2012
© Dana Goswick 2012

Abstract

This paper examines the legitimacy of two common methodologies within philosophy: thought experiments and conceptual analysis. In particular, I examine the uses to which these two methodologies have been put within modal epistemology. I argue that, although both methods can be used to reveal conditional essentialist claims (e.g. necessarily: if x is water, then x is H₂O), neither can be used to reveal the de re essentialists claims (e.g. x is water and x is essentially H₂O) they’re often taken to reveal.

Introduction

This paper examines the legitimacy of two common methodologies within philosophy: thought experiments and conceptual analysis. In particular, I examine the uses to which these two methodologies have been put within modal epistemology. I argue that, although both methods can be used to reveal conditional essentialist claims (e.g. necessarily: if x is water, then x is H₂O), neither can be used to reveal the de re essentialists claims (e.g. x is water and x is essentially H₂O) they’re often taken to reveal.

Post-Kripke’s 1970 Princeton lectures (published as Naming and Necessity, 1972) we’ve become blasé about essentialism, i.e. about de re modality. An object o has a property F essentially iff F is a non-trivial, necessary, context-independent, constant property of o itself, rather than of the way we talk about or represent o. We routinely speak of objects having properties essentially.¹ We’ve been quick – too quick, I’ll argue – to dismiss Quine’s worries about essentialism as resting on the silly error of not distinguishing the de re from the de dicto. There is a deeper Quinean challenge:

From Quine’s arguments a certain kind of challenge is set up for those that want to defend de re modality, and the possibility of an epistemology of it. The challenge to the defender of de re modality is to make sense of de re modality by providing a principled reason for
preferring one way of referring to an object over another way of referring to the same object. (Vaidya 2007, 15)

Or, as the metaphysician might put it, to find a principled reason for claiming it is one object (i.e. the one that is essentially F) rather than another (the one that is only accidentally F) that exists at spacetime region r and to find a methodology which grounds our knowing this. In what follows, I’ll assume the metaphysical question has been resolved in favor of essentialism. That is, I’ll assume the notion of de re modality really is coherent and that there really are objects that have properties essentially. My concern will be with the epistemological question of how we come to know that an object which is essentially F exists. How – given that we know an object which has F exists at spacetime region r – do we come to know that it’s an object which has F essentially rather than an object which has F only accidentally?

There are a plethora of articles concerning how we come to know that o has F essentially and a variety of approaches to securing modal knowledge.2,3 In what follows, I focus on two influential methodologies: thought experiments and conceptual analysis. According to the thought experiment methodology – most prominently advocated by Kripke (1972) –, we discover o is essentially F by imagining various scenarios and discovering that we can’t imagine a scenario in which o is not F. According to the conceptual analysis methodology – advocated by Bealer (1987, 1999, and 2002) –, we discover o is essentially F by discovering that to be o an object must fall under the concept c and falling under concept c requires being essentially F.

An adequate modal methodology must capture both the fact that (i) o exists, ( and the fact that (ii) in order to be o, an object must be essentially F, ( )

Showing only that (i) o exists, independently of showing which properties it has accidentally and which it has essentially, isn’t sufficient because o might exist but be F only accidentally or not at all. Showing only that in order to be o an object must be essentially F isn’t sufficient because it’s compatible with the non-existence of o and, hence, can’t by itself secure the existence of objects which have essential properties. I will argue that, although both the thought experiment methodology and the conceptual analysis methodology satisfy (ii), neither method satisfies (i). Thought experiments fail to satisfy (i) because there are systematic cases of false positives which the method is unable to rule out. In other words, the method will yield that there’s an object o which is essentially F, even when there’s only a different object, o*, which is not essentially F. Conceptual analysis succeeds in demonstrating (ii) only at the cost of undermining (i). We’re left with the claim that any object which satisfies concept c is essentially F, but the reasons we have for thinking there’s an object o which satisfies concept c have been undermined.

Method One: Thought Experiments
According to the thought experiment methodology, if one wants to know whether o is essentially F, one tries to imagine a scenario in which o is not F. If one cannot imagine such a scenario, then one concludes that o is essentially F. Prima facie, one might balk at the thought experiment method:

“If one thinks [that] things have essences quite independently of how we think or talk about them, then the role of imagination in modal inquiry will seem quite mysterious. It will be unclear why we should think that there is even a relation of good evidence . . . between our imaginations and the modal structure of the world” (Sidelle 1989, 88).

I think now of Nixon. I reflect rather carefully on him. This isn’t a method which gives me access to his accidental properties. What reason do we have to think it’s a method which gives me access to his essential properties? Perhaps empirical investigation is the appropriate method to ascertain an object’s accidental properties and rational reflection on thought experiments is the appropriate method to ascertain an object’s essential properties.4

Kripke suggests that we can learn, via thought experiments, what the essential properties of ordinary objects are. We can learn, for instance, that Nixon is essentially animate and that Elizabeth II has her parentage essentially:

Supposing Nixon is in fact a human being, it would seem that we cannot think of a possible counterfactual situation in which he was, say, an inanimate object. (Kripke 1972, 46)

How could a person originating from different parents, from a totally different sperm and egg, be this very woman? One can imagine, given the woman, that various things in her life could have changed. . . But what is harder to imagine is her being born of different parents. It seems to me that anything coming from a different origin would not be this object. (Kripke 1972, 113)

That Kripke intends to make a link between our thoughts and the modal structure of the world is clear. He is not asserting the mere psychological fact that we cannot think of a scenario in which Nixon is an inanimate object. He’s citing the fact that we can’t do so as evidence that Nixon is essentially animate. If Kripke’s target were our concepts, the thought experiment methodology would be in good shape. We can learn about our concepts via reflecting on thought experiments. But Kripke’s target is not our concepts. Kripke’s target is externally existing mind-independent objects. According to Kripke, we engage in various thought experiments which involve the object Nixon: could he have been a puppet, could he have been female, could he have been a mathematician. By doing so, we learn what the object’s (i.e. Nixon’s) essential properties are. As such,
the thought experiment methodology is dependent on our ability to distinguish objects that have exactly the same nonmodal properties, but differ with respect to their essential properties.

Consider Nixon and Snixon. They have exactly the same nonmodal properties, but they differ with respect to their essential properties. Nixon is essentially animate, whereas Snixon is only accidentally animate. Suppose that, in fact, it’s Nixon who actually exists and whom you’re reflecting on when you engage in Kripke’s thought experiment. Nixon is essentially animate, so the modal structure of the world matches the results you ascertain via your thought experiment. But you don’t know that the object in front of you – the object that is, in fact, Nixon – is essentially animate. It’s a happy accident that the object in front of you is essentially animate. You might not have been so lucky. It might have been Snixon that actually existed, rather than Nixon. Presented with Snixon, you would have drawn exactly the same conclusions you drew when presented with Nixon. In particular, you would have (falsely) judged that Snixon was essentially animate. This undermines the thought experiment methodology. We don’t take a red/green color-blind person’s claim that the ball is green to be good evidence that the ball is green because we know that – even if the ball were red rather than green – he would still say it was green. Likewise, we shouldn’t take Kripke’s thought experiment to reveal that there’s an object which is essentially animate because – even if the accidentally animate object Snixon rather than the essentially animate object Nixon were the target of our thought experiment – the thought experiment methodology would still lead us to say that there’s an object which is essentially animate.

There are two ways an advocate of the thought experiment methodology can respond to the Nixon/Snixon case: (1) He can agree object pairs such as Nixon/Snixon – i.e. objects which have the same nonmodal properties but differ with respect to their essential properties – are metaphysically possible, but can argue they don’t undermine the thought experiment methodology. (2) He can deny that object pairs such as Nixon/Snixon are metaphysically possible, i.e. he can argue that, for any two objects x and y, if x and y have exactly the same nonmodal properties, then they have exactly the same essential properties. I’ll consider each response in turn.

Consider, first, the response that, although object pairs such as Nixon/Snixon are metaphysically possible, the possibility that it’s Snixon we are reflecting on, rather than Nixon, doesn’t undermine the conclusion of our thought experiment, i.e. the conclusion that the object in front of me is essentially animate. A methodology need not be infallible to deliver knowledge. The look-and-see method for ascertaining whether there’s a horse in the field is a perfectly acceptable methodology despite the fact that it sometimes yields false positives, e.g. in cases where a distant cow is misidentified as a horse, in cases where I’ve had too much to drink, in cases where my ornery friends have placed horse statues in the field, etc. None of these false positives is particularly troubling because (i) they’re not that common and (ii) we know how to avoid them.
Regarding (i), if my friends systematically placed horse statues in fields I drove by, the look-and-see method would not be a good method for ascertaining if there was a horse in the field. But such occurrences are rare. Most of the time I think there’s a horse in the field, I think this because there really is a horse. The false positives are rare enough that they don’t undermine the methodology. Regarding (ii), false positives which are in principle avoidable don’t undermine a methodology. In fact, the look-and-see method sometimes fails, but it does so because I’m lazy about using it or because someone is tricking me rather than because of some inherent flaw in the methodology. If I avoid drinking too much and only claim there’s a horse when I’m standing three feet away from the object I identify as a horse, then the look-and-see method is pretty darn reliable.

If object pairs such as Nixon and Snixon (i.e. objects which have the same nonmodal properties but have different essential properties) are metaphysically possible, then the thought experiment methodology differs from fallible, but reliable, methodologies – such as the look-and-see methodology – in that the possibility of false positives is both (i) common and (ii) unavoidable. Suppose there are possible objects pairs that have the same nonmodal properties, but different essential properties. Then, there will be many such possible object pairs. But, if there are many such possible object pairs, then there’s always a worry that the result of any particular thought experiment is a false positive. Any time a thought experiment purportedly led to the conclusion that an object was essentially F, the critic could point out, “Wait a minute, o* isn’t essentially F and there’s no way for you to tell if you were presented with o or o*, so there’s no way for you to tell that the object you were presented with is essentially F”. If the advocate of the thought experiment methodology admits the metaphysical possibility of Nixon/Snixon object pairs, he’s in trouble. The thought experiment methodology is only a reliable guide to an object’s essential properties if there’s a 1-1 link between an object’s having nonmodal properties F1-Fn and its having essential property F.

Consider, hence, the response that object pairs such as Nixon/Snixon are impossible, i.e. because objects which have exactly the same nonmodal properties have exactly the same essential properties. Some will think there are clear instances of objects which have the same nonmodal properties, but different essential properties (e.g. Gibbard’s Lump1 and Goliath). But, of course, the way is always open for one to deny this by denying the existence of Lump1 or of Goliath (or of both). What can the advocate of the thought experiment methodology say to defend the claim that objects which have exactly the same nonmodal properties have exactly the same essential properties? He could appeal to thought experiments – e.g. Can you imagine two objects which have exactly the same nonmodal properties, but different essential properties? – but such thought experiments don’t yield the result that Nixon/Snixon cases are impossible. Many coinciding objects’ theorists think that Lump1 and Goliath have exactly the same nonmodal properties even though they differ with respect to their essential properties. Those of us who aren’t coinciding objects theorists can surely imagine that Lump1 and
Goliath both exist. Hence, thought experiments seem to count in favor of – rather than against – the claim that there can be objects which have exactly the same nonmodal properties, but different essential properties. The advocate of the thought experiment methodology could appeal to privileged kinds, e.g. perhaps Nixon is somehow more natural than Snixon and the world only contains objects which are appropriately natural. I won’t here argue that the advocate of the thought experiment methodology can’t save the thought experiment methodology from false positives by arguing that objects such as Snixon (i.e. an object which has the same nonmodal properties as Nixon, but differs with regard to its essential properties) are impossible. But this is no easy task and, absent such a defense, we have no reason to think we can use the thought experiment methodology to figure out what object o’s essential properties are. The problem with this method is that – absent arguments that rule out the existence of objects which have the same nonmodal properties but differ with regards to their essential properties – for every case in which we conclude that the object in front of us is essentially F, we would have drawn the same conclusion even had the object in front of us been only accidentally F. Given this, a thought experiment’s yielding that the object in front of me is essentially F doesn’t lead to my knowing that the object in front of me is essentially F.

Method Two: Conceptual Analysis

The conceptual analyst attempts to gain knowledge of de re modal facts via combining a priori knowledge concerning what objects have to be like to satisfy our concepts with a posteriori knowledge concerning which concepts are actually satisfied. I’ll argue that, by building so much in to what’s required for an object to satisfy concept c (e.g. water), the conceptual analysis methodology undermines the reasons we have for thinking any object actually satisfies concept c (e.g. for thinking there is any water). In other words, although the conceptual analyst fairs well with regard to claim (ii) that in order to be o, an object must be essentially F, he fairs poorly with regard to claim (i) that o exists.

Conceptual analysis has recently been advocated by George Bealer. Bealer argues that we acquire knowledge of de re modal facts by having de re modal facts be the conclusions of arguments which have (a) a general a priori premise involving category concepts which contains the necessity, (b) a linking premise which subsumes a naturalistic concept under a category concept, and (c) an a posteriori premise which asserts the existence of an object which falls under the naturalistic concept mentioned in the linking premise (Bealer 1987, 296-297). According to Bealer, we know (a) general a priori premises because we can figure out – via ideal rational reflection – what the essential relations between a priori concepts are, we know (b) linking premises because we are able to figure out – via ideal rational reflection – which naturalistic concepts fall under which a priori category concepts, and we know (c) a posteriori premises via empirical investigation (ibid.). Thus, we have three premises, each of
which we appear to know, which can be used to construct a valid argument with a de re modal conclusion, viz.

**Water Argument**

P1: All objects which fall under the concept **compositional substance** have their chemical composition essentially.

P2: All portions of water are objects which fall under the concept **compositional substance**.

P3: Object o is a portion of water.

C: Object o has its chemical composition essentially.

P3 is problematic. If Bealer’s right about all the conditions an object must satisfy to fall under the concept water, then we can’t know a posteriori whether an object is a portion of water. In order to use conceptual analysis as a stepping stone to our knowing that there are objects in the world which have properties essentially, Bealer needs, also, to argue that we know some of the objects in the world are portions of water. He can’t rely on the standard reasons we have for thinking there’s water (“look, it’s coming out of the faucet”) because such reasons can’t distinguish water from swater, i.e. from an object which has the same nonmodal properties as water but which is only accidentally chemically composed of H2O. In other words, presented with swater, we would judge it to be water. But we would be wrong. Conceptual analysis only yields knowledge of the de re modal fact that there’s an object which is essentially F if supplemented by the claim that we know that there’s a portion of water: [(□Fx) & (Ǝx)(x = o)] → (Ǝx)(x = o & □Fx).

Given that we can’t distinguish water and swater, it’s difficult to see how we could know that o is water rather than swater. That is, it’s difficult to see how we could know that (absent such knowledge, the conceptual analyst is left with the conditional essentialist claim for all x, if x is a portion of water, then x has its chemical composition essentially, rather than with the de re essentialist claim there exists a portion of water x which has its chemical composition essentially.

**Conclusion**

Both the thought experiment methodology and the conceptual analysis methodology can be used to generate knowledge of conditional essentialist conclusions of the form (□Fx). Kripke’s thought experiment methodology licenses one to say, “Gold has atomic number 79 essentially” and “Men are essentially animate”. What it doesn’t licence one to say is, “That object there – o – is a portion of gold” and “That object there – o – is a man”. Likewise, Bealer’s conceptual analysis methodology licenses one to say, “Water is essentially H20”. What it doesn’t licence one to say is,
“That object there – o – is a portion of water”. In order to yield knowledge of the essential properties of actual objects both the thought experiment methodology and the conceptual analysis methodology must be supplemented by the existential claim (\( \exists x (x = o) \)). Given how high a bar the thought experiment methodology and the conceptual analysis methodology set for an object’s being o – it’s not enough that it just have certain properties, it must have them essentially – it’s no easy task to ascertain whether any actual object is object o. Until supplemented by such an account, neither the thought experiment methodology nor the conceptual analysis methodology can be used to generate knowledge of de re essential claims of the form (\( \exists x ) ( x = o & □ \)

References


1 Regarding speaking of objects having essential properties: “[O]nce we have discovered the nature of water, nothing counts as a possible world in which water doesn’t have that nature. Once we have discovered that water (in the actual world) is H2O nothing counts as a possible world in which water isn’t H2O” (Putnam 1975, 150); “Particular individuals have essences without which they would not ‘be what they are’ – would not exist” (Wiggins 1980, 125); “four-dimensional hunks have their boundaries essentially” (Heller 1990, 53); “Sandy: she has her origin essentially, so she could not have developed from a different zygote” (Paul 2006, 350); “on the Constitution View, things have their primary-kind-properties essentially” (Baker 2007, 221); “Are there essential differences even among exact duplicates? I suppose that there are, and that the most plausible place to locate such differences is . . . in different origins and different original parts” (Johnston Hylomorphism 659).

In addition to the thought experiment methodology and the conceptual analysis methodology discussed here, see also Yablo’s epistemic conceivability-based methodology (Yablo 1993, Roca-Royes 2011), Chalmers’ non-epistemic conceivability-based methodology (Chalmers 2002), Williamson’s conceivability-consistent-with-counterfactuals methodology (Williamson 2007), Peacocke’s principles of possibility methodology (Peacocke 2002). Although each of these methodologies focuses primarily on what it is for something to be possible, an account of what it is for o to be essentially F can often be derived from this, viz. for o to be essentially F is for it not to be possible for o to fail to be F.

4 See Bealer (1987) for a defense of the claims that (i) although we learn about an object’s accidental properties via empirical investigation, we learn about its essential properties via intuition, and (ii) thought experiments are a way of becoming clear as to what our own intuitions are.

5 See Gettier (1963) for a discussion of the difference between knowledge and happy accidents. The Nixon/Snixon case differs from Gettier’s cases in that, in a Gettier case, one is justified, whereas in the Nixon/Snixon case it’s not obvious that one is justified. One might argue that the metaphysical possibility that it’s Snixon rather than Nixon that one is interacting with undermines not just one’s knowledge that there’s an object which is essentially animate, but, also, one’s justification for thinking that there’s an object which is essentially animate. Nonetheless, there are interesting parallels between the Nixon/Snixon case and Gettier’s cases.

6 Similar to e.g. the way we don’t think the fact that I sometimes mistake far away horses for cows undermines my ability to use the look-and-see method to distinguish horses from cows.

7 Why think if there are any objects which have the same nonmodal properties but different essential properties, then there will be many objects which have the same nonmodal properties but different essentially properties? It would be odd to think there’s a pair of objects Nixon and Snixon which differ only with regards to their essential properties, whilst denying that there’s a pair of objects Obama and Sobama which differ only with regards to their essential properties. Nixon and Obama are both human beings. Surely if there’s an object which is just like Nixon save that it’s accidentally animate, rather than essentially animate, then there’s another object which is just like Obama save that it’s accidentally animate, rather than essentially animate. But, then, for every human being there is, presumably, an object that’s just like that human being save that it’s accidentally animate rather than essentially animate. Hence, either there are many objects which have the same nonmodal properties but different essentially properties or there are no such objects.

8 The only actual objects which can have the same nonmodal properties are coinciding objects, i.e. otherwise the objects would differ with regard to their nonmodal properties – one would be located at region r and the other would not. That coinciding objects theorists face a problem grounding the essential properties of coinciding objects is well known (see: Sosa, 1987; Heller, 1990; Burke, 1992; Zimmerman, 1995). What has been less often appreciated is that even those who deny that there are coinciding objects are faced with a worry about which, of the many metaphysically possible objects, actually exists (see: Bennett, 2004). For the non-coinciding objects theorist, at most either Nixon or Snixon exists, but the non-coinciding object theorist faces a task at least as daunting as the grounding task the coinciding objects’ theorist faces: that of arguing that it’s Nixon rather than Snixon that exists.

9 See Gibbard (1975)

10 See Burke (1994) for a defense of the claim that only Goliath exists. See van Inwagen (1990) and Heller (1990) for a defense of the claim that neither Goliath nor Lump1 exists.
11 See Koslicki (2008) for a defense of the claim that there are privileged kinds and that whether an object exists depends in part on whether it belongs to a privileged kind. See, also, Lewis (1983). Lewis introduces the distinction between perfectly natural properties and less natural properties. Although Lewis himself is a pluralist about the existence of objects, one can imagine a neo-Lewisian who accepts Lewis’ distinction between natural and unnatural properties and then goes on to argue only those objects which are ‘natural-enough’ exist.

12 I suspect, however, that it does lead to my knowing something else. Namely, that satisfying whatever concept is most salient to being Nixon (e.g. human being) requires being essentially animate. Thought experiments yield knowledge of conditional essentialist claims of the form (\(\exists x)(x = o \rightarrow \square Fx\)) they just don’t yield the de re essential claims of the form (\(\exists x)(x = o \& \square Fx\)).


14 See Roca-Royes (2011) for similar concerns about whether any actual objects satisfy Bealer’s concepts.

15 \([\neg \forall x(\neg Fx) \& \neg \exists x(x = o)] \rightarrow \exists x(x = o \& \square Fx)\]

16 There are various accounts in the literature with regards to securing our knowledge of the existential claim, (\(\exists x(x = o)\)) particular, plenitude (Bennet 2004), brutalism (Markosian 1998), and privileged-kinds (Koslicki 2008).