The Two-Dimensional Argument for Materialism

Jessica Wilson*

Eclectic Approaches to Causality and Explanation
December 17, 2021

1   Consciousness and explanatory gaps

My topic today concerns the status of consciousness as physically acceptable or unacceptable. As Chalmers (1996) notes, consciousness takes many forms:

> Conscious experiences range from vivid color sensations to experiences of the faintest background aromas; from hard-edged pains to the elusive experience of thoughts on the tip of one's tongue; from mundane sounds and smells to the encompassing grandeur of musical experience [...] All these have a distinct experienced quality. (4)

Here I will speak generically of consciousness (or associated mental features) with such ostensively defined experiences in mind. Various anti-materialist/anti-physicalist arguments appeal to or rely on insuperable predictive or explanatory gaps between consciousness and physical goings-on:

- 'In-principle unpredictability' arguments, whereby even a mathematical archangel could not predict qualitative features of conscious experience (Broad 1925, 71)

- Knowledge arguments, whereby complete knowledge of the physical facts fails to generate complete knowledge of the consciousness facts (Nagel 1974, Jackson 1986)

- Conceivability arguments, whereby explanatory gaps enable the conceivability of zombies—physical duplicates of creatures like us, but lacking in conscious mentality (Levine 1983, 359)

There are initial difficulties with these explanatory gap arguments:

- Various complex systems are clearly physically acceptable, though having features that are relevantly insuperably unpredictable or unexplainable in lower-level physical terms.

- Conceivability-based arguments advance beyond mere explanatory failures, but on the face of it, conceiving tracks epistemic rather than metaphysical possibility.

Correspondingly, what's needed are independent reasons for thinking that explanatory gaps or the conceivability of, e.g., zombies, has such import.

---

*Department of Philosophy, University of Toronto; jessica.m.wilson@utoronto.ca

2 Chalmers's two-dimensional argument against materialism

The primary advance of Chalmers's (1996, 1999, 2009) 'two-dimensional argument against materialism' lies in his situating the conceivability of zombies in an independently motivated framework—'epistemic two-dimensionalism', or E2D. The motivating line of thought is as follows:

- On E2D, certain facts about meaning, taken to be a priori accessible, can be used to identify certain facts about modality, expressing what is genuinely metaphysically (im)possible.

- It is commonly assumed that the a priori access to meanings proceeds by way of conceiving.

- So, commitment to the E2D strategy for gaining access to modal truth, and to implementing this strategy via a conceiving-based epistemology of meanings, provides an independent basis for taking the conceivability of zombies to have anti-physicalist metaphysical import, as a case-in-point of a systematic connection between conceivability and metaphysical possibility.

Chalmers's argument then proceeds as follows:

1. It is conceivable that there's a world physically exactly like ours, but lacking consciousness.

2. If the world described in (1) is conceivable, then it is metaphysically possible. (E2D)

3. If the world described in (1) is metaphysically possible, then consciousness is 'over and above' the physical, and materialism/physicalism is false.

4. So consciousness is over and above the physical, and materialism (physicalism) is false.

While Chalmers's argument improves on other gap arguments, it nonetheless fails. My focus will be on undercutting premise (2). The more general plan is as follows:

- I'll present E2D in more detail, highlighting its attractiveness.

- Drawing on work with Stephen Biggs, I'll suggest that E2D is better implemented by appeal to an abduction-based rather a conceiving-based epistemology of the meanings at issue.

- One might wonder whether abduction would be suited for implementing E2D, given the need for access to meanings to be a priori. Drawing on other work with Biggs, I argue that abduction is an a priori mode of inference.

- I'll argue that when E2D is implemented using abduction rather than conceiving, the prima facie weight of considerations push against zombies' being metaphysically possible, contra Chalmers, and consonant with materialism.

3 Epistemic two-dimensionalism

Chalmers's appeal to E2D takes as its starting point Kripke's (1972/80) undermining of the traditional supposition that modal truths are (always) a priori accessible:

- Kripke argues that many necessary truths about individuals and natural kinds can be known only a posteriori. For example, it is a posteriori that water is identical to H_2O; nonetheless, Kripke plausibly asserts, given that water is H_2O, then water is necessarily H_2O.

- Kripke's results might be seen as undercutting the prospects of our having much, if any, a priori modal knowledge about broadly scientific goings-on.

That would be undesirable, since much theory and practice assumes we can know modal truths independently of, or at least prior to the end of, empirical inquiry. So we might hope that, post-Kripke, the link between necessity and a priority could be restored, to some considerable extent.

This is the promise of E2D, per Chalmers (1996, 2009), Chalmers and Jackson (2001), and others. E2D refines and generalizes Frege's (1892/2010) claim that there are two kinds of meaning: sense, which is a priori accessible to a competent speaker, and reference, which may not be so accessible. E2D similarly aims to characterize two aspects of meaning, represented as 'intensions': functions from possible worlds to extensions. Following Chalmers's exposition:

- Intensions are functions from either scenarios (centered worlds, incorporating a subject's perspective) or scenario-world pairs to extensions.

- The 'primary' intension of an expression E takes each scenario s to the extension of E in s on the supposition that s is actual. The 'secondary intension' takes each scenario-world pair $< s, w >$ to the extension of E in w given that s is actual. Primary intensions track extensions in worlds "considered as actual"; secondary intensions track extensions in worlds "considered as counterfactual", holding fixed which world is actual.

- The E2D strategy supposes that many natural kind expressions of the sort entering into a posteriori necessities are associated with primary and secondary intensions, so understood.

- The promise of E2D ultimately lies in the suggestion that primary and secondary intensions are connected so as to provide a basis for a priori knowledge of a wide range of modal truths.

Thinking in terms of a two-dimensional matrix, with worlds "considered as actual" along the vertical, and worlds "considered as counterfactual" along the horizontal, assists in understanding the E2D strategy. For example, the two-dimensional matrix for the intensions associated with the term 'water' would look something like the following:

'water'	H_2O-world	XYZ-world	. . .
H_2O-scenario	H_2O	H_2O	. . .
XYZ-scenario	XYZ	XYZ	. . .
. . .	. . .	. . .	. . .

The entries encode that our access to appropriate intensions provides a basis for our knowing a priori that *if* water is actually H_2O, *then* water is necessarily H_2O; and *if* water is actually XYZ, *then* water is necessarily XYZ; and so on. Hence while it is a posteriori that water is necessarily H_2O, the empirical contribution here and in other cases is limited to discharging the antecedent of a conditional known a priori—compatible with our having a priori access to much modal knowledge.

So far, so good. The question remains which epistemology of intensions (meanings) should be taken to be operative in implementing this strategy. This is usually assumed to involve conceiving:

> Conceivability is the only guide to necessity; our concepts, and the intuitions about possibility that derive from them, provide our only grip on modal claims. (Gertler 2006, 205)

Given E2D and a conceiving-based epistemology of intensions (CEI), we are in position to generate the conclusion of Chalmers's two-dimensional argument against materialism:

- CEI implies that 'zombie' has a positive extension at some world if one can conceive of a zombie world; E2D coupled with CEI implies that, if one can conceive of a zombie world, then zombies are metaphysically possible.

- But as per premise 1 (granted here), one can conceive of a zombie world. So zombies are metaphysically possible, which along with premise 3 (also granted here) implies that consciousness is not physically acceptable, contra materialism/physicalism.

4 Conceiving vs. abduction-based epistemologies of intensions

Why suppose that E2D must proceed via CEI? Why not take the epistemology of intensions/meanings to proceed via inference to the best explanation, as per an abduction-based epistemology of intensions (AEI), on which various abductive principles—plausibility, compatibility with other beliefs, unifying power, fruitfulness, ontological and ideological parsimony, and so—might come into play in identifying the extensions of concepts or terms in non-actual scenarios?

Plausibly, AEI has not been taken seriously as a basis for implementing the E2D strategy on grounds that, if the strategy is to do the job of reforging an a priori accessible link between meaning and modality, the operative epistemology of intensions must involve an a priori mode of inference; and while conceiving is a priori, abduction is not.

5 The a priority of abduction

We offer many positive and defensive motivations in support of abduction's being a priori. Here we briefly sketch the line of thought in our (2017), and highlight some historical and contemporary precursors of our view. Please see our various papers for a more fullsome exposition.

5.1 Abductive principles and ceteris paribus clauses

We assume that (belief in) the epistemic value of abduction is justified, and argue that this justification is a priori. To start, consider a specific abductive principle—say, Parsimony, according to which, ceteris paribus, one should choose the theory involving the fewest fundamental type-level ontological commitments. Why think that Parsimony's epistemic value (hence that of abduction, as partly constituted by Parsimony), depends on experience? Two lines of thought are on offer:

1. Parsimony's epistemic value depends on whether the world is parsimonious.

2. Parsimony's epistemic value depends on whether parsimonious theories are more likely to be true.

Re (1): how many fundamental kinds (for example) does a world have to contain in order to be unparsimonious? Two? Forty-two? A billion? This question is misguided, since Parsimony is a principle or a norm guiding theory choice, not a descriptive fact about the world. Even if there is a sense in which a world with a billion kinds is unparsimonious, this fact would be irrelevant to Parsimony's epistemic value. It would remain, in such a world, that in theorizing one should not posit ore fundamental kinds (say, a billion and one, at that world) than are needed.

Re (2): Here the idea seems to be that we could gain evidence about whether theories satisfying Parsimony were more likely to be true. That's incorrect, however:

> [W]e could never gain empirical evidence, however 'indirect', to this effect. Parsimony, like all abductive principles, requires that 'other things be equal'. Consequently, no empirical evidence could, even in principle, distinguish between a world in which theories satisfying Parsimony were more likely to be true, and one where this was not the case. If there were such empirical evidence—if, for example, experiments were to indicate

that the world contained more fundamental kinds than our best theory implies—then the ceteris paribus condition in Parsimony would not be met: one theory would be explanatorily better than the other (*vis-à-vis* another abductive principle). We could never be in empirical position to know, then, that the actual world is not cooperating with Parsimony. (743–4)

Generalizing, we suggest that common supposition that abduction is a posteriori reflects a failure to appreciate that the ceteris paribus clauses associated with abductive principles operate to shield them from empirical disconfirmation.

Note that it is commonly recognized that failure to possess all the relevant data can lead to false belief via a priori modes of inference. For example, were one to initially apply modus ponens to some premises later revealed to be false, that would show not that modus ponens is a posteriori, but rather that one was working with faulty (incomplete or inadequate) data.

- Similarly, we maintain, for abduction: that new evidence might undermine the conclusion of a given abductive inference goes nowhere towards establishing that abduction is a posteriori.

- On the contrary, any new evidence would, at best, show that other things were not in fact equal as among the candidate theories.

Objection: Even if no empirical evidence could bear on Parsimony's epistemic value, our justification for this belief might be a posteriori, in depending on a contingent but empirically inaccessible fact. If, say, some kinds have hidden fundamental metaphysical essences, even though an application of Parsimony given all the empirically accessible facts would deem them non-fundamental. In this case, Parsimony would lead us astray, and belief in the epistemic value of Parsimony would be a posteriori unjustified, even if we never could be in position to know this. We respond:

1. Rather than interpret the case as showing that belief in Parsimony's epistemic value is justified a posteriori, in depending on how the world contingently is, one can rather interpret the case as showing that belief in Parsimony's epistemic value, while justified a priori, is defeasible (see Casullo 2003, Summerfield 1991, Thurow 2006).

2. Supposing that there is some principled non- empirical motivation for thinking that such metaphysical facts are in place, then accommodating these facts would be part of the ceteris paribus conditions under which Parsimony would be properly applied, such that the presence of such distinctions could not, even in principle, undermine justified belief in the epistemic value of Parsimony. After all, the ceteris paribus condition in abductive principles such as Parsimony doesn't advert only to the proper accommodation of empirical facts, but more generally to any facts that we have reason to believe obtain.

   And if the metaphysical facts are in-principle inaccessible not just to empirical but to rational investigation, then they can simply be rejected as useless posits.

5.2 Four roles for experience in justification

Given that the epistemic value of abduction is not a posteriori, we then argue that it is moreover a priori, by attention to the roles experience may play in inferential reasoning.

Consider a claim p. There are four ways in which experience might play a role in the course of a particular belief in p coming to be justified. Experience might play a role in

1. acquiring the concepts required to entertain p,

2. acquiring the evidence required to justify belief in p,

3. justifying belief in the epistemic value of the mode of inference used to justify belief in p, or

4. acquiring or learning to deploy the mode of inference used to justify belief in p.

Can a belief for which experience plays an ineliminable role along one or more of (1)–(4) be justified ('entirely') a priori? As we discuss in our 2017a, it depends on which role is at issue:

- Re (1): it is commonly maintained that belief in p can be justified a priori even if experience is needed to acquire the concepts required to entertain p. For example, belief in 'sisters are siblings' can be justified a priori even if we need experience to acquire the concepts expressed by or meanings of 'sister' and 'sibling.'

- Re (2): it is commonly maintained that belief in p cannot be justified a priori if experience is (at all) needed to acquire the evidence supporting p. This is the sense in which any reliance on experience or empirical facts suffices to render the associated justification a posteriori. For example, belief in 'water is H_2O' is commonly taken to be justified a posteriori, on grounds that justifying this belief requires, among other things, acquiring empirical evidence to the effect that water and H_2O are spatiotemporally coextensive.

- Re (3): it is commonly maintained that belief in p cannot be justified a priori if experience is needed to justify belief in the epistemic value of the mode of inference required to justify belief in p. (It is this role which is typically offered as indicative of abduction's being a posteriori.)

- Re (4): although the role of experience in an agent's acquiring or learning how to deploy a given mode of inference is not much discussed, it seems reasonable to maintain that this role is relevantly similar to that of (1): in each case, experience contributes to belief (or supposition) formation, not to justification per se. For example, students often need encouragement to think in an appropriately imaginative way about what is possible; but that such imaginative 'training up' is required in order to engage in suitably competent conceiving is not taken to undermine the status of conceiving as an a priori mode of inference.

Correspondingly, only (2) and (3) are such that the playing of this role in the justification of a given belief would render that justification a posteriori.

So, whether an abductively justified belief p is justified a priori turns on whether experience enters into its justification via either role (2) or role (3):

- Experience does not enter into abductively justifying the belief in p via role (3): as previously, the justification of the epistemic value of abduction does not rely on experience.

- Re (2): to start, though abduction often operates on empirical evidence (e.g., 'water and H_2O are coincident') to produce an a posteriori justified claim (e.g., 'water is H_2O'), the inferential transition can be encoded in conditionals—e.g., 'if water and H_2O are coincident, then water is H_2O'—which are justified via a hypothetical form of abduction.

  Such hypothetical abductive inference is akin to suppositional reasoning in conditional proof, enabling identification of what would be the best explanation of the antecedent were this to obtain/be true. Since such conditional beliefs may be justified without the antecedent's being believed or true, abduction here operates independently of any claim justified through experience. Accordingly (and since abduction's epistemic value does not rely on experience), such abductively justified conditional beliefs are 'entirely' a priori.[1]

---

[1] Hawthorne (2002, 252) makes a similar point, suggesting that abduction can deliver a priori justification for belief in a conditional whose antecedent describes an 'experiential life history' and whose consequent is whichever theory best explains some aspect of that life history.

- Hence abduction is an a priori mode of inference, even when operating on (non-hypothetical) empirical evidence. When abduction operates on empirical evidence, it does not produce beliefs with (completely) a priori justification, any more than modus ponens does. But the contribution of empirical evidence in such cases of abductive inference is simply to discharge the antecedent of a conditional whose justification is entirely a priori. This line of thought is familiar from discussion of the conditional a priori basis of a posteriori modal claims.

5.3 Precursors

The view that abduction is a priori may be non-standard, but the position has notable historical and contemporary precursors. In brief (see Biggs and Wilson 2019):

- Kant took synthetic a priori truths (including most philosophical claims) to be justified via an ampliative mode of inference—plausibly, abduction.

- Carnap took knowledge of conceptual content to proceed via explicitly abductive 'explication'.

- In addition to Hawthorne (2002), Cohen (2010), and Wedgwood (2013), several philosophers, including Bonjour (1998), Swinburne (2001), and Peacocke (2003), have offered reasons aiming to establish the a priority of certain ampliative modes of inference or abductive principles.

6 The advantages of implementing E2D via AEI rather than CEI

The E2D strategy is better coupled with AEI than with CEI. The rough idea is as follows:

The E2D strategy requires expressions to have a priori-accessible extensions in every possible scenario. But in cases of conceptual underdetermination, associated expressions lack fixed antecedent extensions in some scenarios. Such underdetermination is widespread, threatening the strategy, unless overdetermination can be overcome. Conceiving alone, understood (as per Chalmers, Jackson) as not involving abductive considerations, cannot do so. However, abduction, being ampliative, can do so.

There is in fact pervasive underdetermination in our concepts from several sources, reflecting that

- applications of natural kind predicates frequently depend on arbitrary factors (M. Wilson 1982 and 2006);

- predicate vagueness is pervasive, and the application of such predicates to borderline cases is typically determined by arbitrary contextual and/or psychological factors (Raffman 1994);

- scientific record supplies many cases of underdetermination in natural kind terms—'acid', 'mass', 'planet', etc.—as reflected in dispute over applications of terms to new cases.

There aren't any realistic prospects of overcoming this underdetermination via conceiving alone:

- A deeper lesson of Wilson's cases and of Raffman's discussion is that the influence of such factors cannot be foreseen, at least not by conceiving alone. There is simply no fixed extension to 'conceive', even taking relevant circumstances into account.

- As we know, the applications of natural kind expressions such as 'acid', 'mass', and 'planet' has been heavily informed by abductive considerations; but the operative notion of conceiving does not involve such abductive or ampliative resources.

By way of contrast, AEI is up to the task of overcoming conceptual underdetermination, in the cases of natural kind expressions and more generally:

- When considering whether to apply an expression in a given scenario, abductors can consider not only historical accident and psychological variability, but also any non-demonstrative rational grounds that might push towards one extension rather than another.

- Hence abduction, unlike conceiving, can be productive (it is ampliative, after all), allowing those identifying intensions to consider how the concept *should* be applied, given the usual abductive principles, even in cases of conceptual underdetermination.

- Since abduction can, in a rational way, go beyond what expressions antecedently encode, AEI has the potential to overcome conceptual underdetermination, extending applications of natural kind expressions to new scenarios or situations, as the E2D strategy requires.

7   Revisiting the metaphysical possibility of zombies

I now return to Chalmers's two-dimensional argument against materialism, and to whether the independently desirable E2D framework supports the metaphysical possibility of zombies:

> The question is: are there physical and functional duplicates of our world in which the intension of 'zombie' has a non-null extension—that is, worlds where there are zombies?

On AEI, answering this question will proceed by attention not just to what we can conceive but to any relevant considerations. Different 'theories' of the intensions at issue will rate better or worse, depending on how well these do at accommodating or satisfying these considerations:

- A theory on which 'zombie' has a non-null extension at some world scores positively in accommodating the seeming conceivability of zombies.

- On the other hand, since a non-null extension implies that consciousness is Strongly emergent or otherwise physically unacceptable, the associated theory of intension of 'zombie' scores negatively in being less ontologically parsimonious than one on which 'zombie' has a null extension at every world (compatible with conscious states being identical with/realized in physical states).

- A theory on which 'zombie' has null extension at every world also scores positively in treating conscious states in unified fashion, as (like nearly all phenomena) physically acceptable.

- A theory on which the extension of 'zombie' is non-null at some world is less systematic and explanatory so far as accommodating that conscious states can have physical effects, for it is unclear how non-physical and physical goings-on can interact; and though I think this can be made sense of via the notion of a fundamental mental interaction, such an account requires further ontology and ideology.

- One might respond by coupling a theory on which 'zombie' has a non-null extension at some world with the supposition that conscious states are epiphenomenal, as Chalmers sometimes suggests. But epiphenomenalism is implausible, given our experience, and unsystematic, since no other empirical goings-on are taken to be epiphenomenal. So this theory too suffers by way of comparison with one taking 'zombie' to have a null primary intension.

To be sure, the ultimate abductive comparison of theories of the intension(s) of 'zombie' will involve not just these, but many other considerations.

- Even so, it is clear that an abductive rather than conceiving-based approach is very far from indicating that 'zombie' has a non-null extension at some world, as is required on the E2D strategy if zombies are to be metaphysically possible.

- On the contrary, at present the all-things-considered weight seems to be on the other side.

References

Biggs, Stephen and Jessica M. Wilson, 2017. 'The A Priority of Abduction'. *Philosophical Studies*, 174:735–758.

Biggs, Stephen and Jessica M. Wilson, 2019. 'Abduction vs. Conceiving in the Epistemology of Modality'. *Synthese*.

Bonjour, Lawrence, 1998. *In Defence of Pure Reason)*. Cambridge: Cambridge University Press.

Broad, C. D., 1925. *Mind and Its Place in Nature*. Cambridge: Kegan Paul. From the 1923 Tanner Lectures at Cambridge.

Casullo, Albert, 2003. *A Priori Justification*. Oxford University Press USA.

Chalmers, David J., 1996. *The Conscious Mind*. Oxford: Oxford University Press.

Chalmers, David J., 1999. 'Materialism and the Metaphysics of Modality'. *Philosophy and Phenomenological Research*, 59:473–496.

Chalmers, David J., 2009. 'The Two-Dimensional Argument Against Materialism'. In Brian P. McLaughlin and Sven Walter, editors, *Oxford Handbook to the Philosophy of Mind*. Oxford: Clarendon Press.

Chalmers, David J. and Frank Jackson, 2001. 'Conceptual Analysis and Reductive Explanation'. *The Philosophical Review*, 110:315–60.

Frege, Gottlob, 1892/2010. 'On Sense and Reference'. In Darragh Byrne and Max Kölbel, editors, *Arguing About Language*, 36–56. Routledge.

Gertler, Brie, 2006. 'Consciousness and Qualia Cannot Be Reduced'. In Robert J. Stainton, editor, *Contemporary Debates in Cognitive Science*, 202–216. Malden, MA: Blackwell.

Hawthorne, John, 2002. 'Deeply Contingent a Priori Knowledge'. *Philosophy and Phenomenological Research*, 65:247–269.

Jackson, Frank, 1986. 'What Mary Didn't Know'. *The Journal of Philosophy*, 83:291–295.

Kripke, Saul, 1972/80. *Naming and Necessity*. Cambridge, MA: Harvard University Press.

Levine, Joseph, 1983. 'Materialism and Qualia: The Explanatory Gap'. *Pacific Philosophical Quarterly*, 64:354–361.

Nagel, Thomas, 1974. 'What is it Like to be a Bat?' *The Philosophical Review*, 83:435–50.

Peacocke, Christopher, 2003. *The Realm of Reason*. Oxford University Press.

Raffman, Diana, 1994. 'Vagueness Without Paradox'. *Philosophical Review*, 103:41–74.

Summerfield, Donna M., 1991. 'Modest a Priori Knowledge'. *Philosophy and Phenomenological Research*, 51:39–66.

Swinburne, Richard, 2001. *Epistemic Justification.* Oxford University Press.

Thurow, Joshua, 2006. 'Experientially Defeasible a Priori Justification'. *Philosophical Quarterly*, 56:596–602.

Wilson, Mark, 1982. 'Predicate Meets Property'. *The Philosophical Review*, 91:549–589.

Wilson, Mark, 2006. *Wandering Significance: An Essay on Conceptual Behavior.* Oxford: Clarendon Press.