Too Many Omissions, Too Much Causation?

Petersson, Björn

Published in:
Maurinian Truths

2019

Document Version:
Publisher's PDF, also known as Version of record

Maurinian Truths

Tobias Hansson Wahlberg | Robin Stenwall (Eds.)

Maurinian Truths – Essays in Honour of
Anna-Sofia Maurin on her 50th Birthday

Maurinian Truths

– Essays in Honour of
Anna-Sofia Maurin on her 50th Birthday

LUND
UNIVERSITY

Tobias Hansson Wahlberg | Robin Stenwall (Eds.)

Painting by Jean-Louis Maurin

Photo by Jesper Heimerson

Department of Philosophy

Lund University

ISBN 978-91-88899-53-8 (print)

ISBN 978-91-88899-54-5 (digital)

Printed in Sweden by Media-Tryck, Lund University
Lund 2019

Media-Tryck is an environmentally certified and ISO 14001:2015 certified provider of printed material. Read more about our environmental work at www.mediatryck.lu.se

MADE IN SWEDEN

Foreword

Anna-Sofia has recently turned fifty. This is occasion for celebration and compilation of a *Festschrift*!

Before becoming a professor in Gothenburg, Anna-Sofia was a member of the philosophy department here in Lund, and a well-respected one at that. We have to confess that as PhD-students we were a bit afraid of Anna-Sofia because she was our senior and gave a somewhat critical, fiery and flamboyant impression. But, as anyone who has had the pleasure to know Anna-Sofia can attest, the above is not a list of her shortcomings but of her strengths as a philosopher, colleague and friend. Not only did she force us to sharpen our arguments and look at the problems from different angles, but her outspokenness, which was (and still is) equally directed across the academic hierarchy, convinced us of her integrity, honesty and loyalty. And as is typical of persons with these qualities: once exposed to, they are the most addictive and rewarding.

The topic of this *Festschrift* is metaphysics. Its title is *Maurinian Truths*. Maurinian truths, as we understand them, are about the underlying reality undergirding, grounding or making true the more familiar Moorean truths (cf. Maurin 2002; Moore 1925/1959). Given that we metaphysicians are highly fallible, a more accurate title would perhaps have been *Maurinian Truths - and Possibly some Maurinian Falsehoods*. Or simply, *Maurinian Falsehoods*. But we wanted the title to reflect our, *qua* metaphysicians, optimistic (but undoubtedly overconfident) state of mind. Of course, yet another

possibility would have been to choose the conditional title *If Maurinian Truths*; but as of now we haven't fully worked out the consequences of the papers (if true). (And, taken as a compound unit, they clearly entail too many consequences; some of the papers, interestingly enough, flatly contradict each other.)

All in all, the *Festschrift* consists of eighteen papers written by Swedish and international Anna-Sofia fans. The papers are all quite short. We have tried to keep them around 2000 words (give or take a few hundred words, depending on the compliance of the author). We hope you, Anna-Sofia, will find them stimulating and enjoyable!

The Editors

Acknowledgements

The printing of this Festschrift was financed by the Department of Philosophy, Lund University. This work is also partly supported by Riksbankens Jubileumsfond (The Swedish Foundation for Humanities and Social Sciences) [grant number P16-0541:1].

References

Maurin, A-S. (2002) *If Tropes*, Dordrecht: Kluwer.

Moore, G. E. (1925/1959) A Defence of Common Sense. In *Philosophical Papers* (pp. 32-59), London: Allen and Unwin.

Content

No Trope has Another Trope as Part

Peter Simons

1. Tropes as simple

Anna-Sofia Maurin's *If Tropes* (2002) is an impressive sustained defence of trope ontology. As a co-adherent of this ontology, I make no issue with her arguments for tropes or her rebuttal of counterarguments, but am assuming that trope ontology is, even if not – as we both believe – correct, at least a strong contender for fundamental ontology. Instead I shall be examining a thesis that Maurin proposes in her book in the expository second Chapter, Section 2.2 (2002: 11–15). This is the proposition

Tropes are simple.

In the section in question, Maurin defends her thesis against arguments by Chris Daly to the effect that a trope has to be complex, namely as a state of affairs of a substrate instantiating a universal. While having no complaints about her successful rebuttal of Daly, it seems to me that a more moderate position, that *some* tropes are complex, needs further consideration, and that is my topic here. Specifically, I shall be considering the question whether one trope can have another trope as a (proper) part. If it can, then by the Weak Supplementation Principle of mereology (Simons 1987: 28), such a trope has at least two disjoint parts, and is therefore complex.

2. Tropes *are* parts

The standard tropist view of concrete particulars (hereafter: *concreta*) is that they consist at least mainly of tropes. The more parsimonious (and elegant) bundle theory holds that concreta consist ultimately of tropes and nothing else. The adverb 'ultimately' is necessary because most familiar concreta have other concreta as parts. The bundle theory applies in the first instance to concreta with no other concreta as parts: atomic concreta. Candidate examples are fundamental particles such as electrons and quarks. Whether they really are good examples is not the point: what the tropist holds is that whatever the atomic concreta are, they are trope bundles. How the bundles are held together is a matter of debate, but is not here crucial. What bundle theorists all hold is that no further parts apart from tropes go to make up atomic concreta. A few tropists such as Martin (1980) hold that tropes alone do not a concretum make: a concretum requires a substrate as a further constituent to hold it together. Such a substrate theorist does not deny that tropes are parts of concreta, only that they are not the only parts.

3. Tropes can be extended

Tropes are employed in ontology to explain the qualitative and quantitative natures of things. For example, a fundamental particle typically has a mass, electric charge, spin, magnetic moment, and a few other properties. Each of these, for the tropist, is a trope. According to physics, fundamental particles are not point-sized, but extend over a small region. It is therefore reasonable to assume that the same goes for their tropes. If particles are not basic, but either

fields or processes are more basic, then these are more obviously extended; in the case of processes, in time as well as space. Only if there are point-sized atomic concreta would their tropes also be unextended. But it is extremely doubtful that there are such concreta.

Being extended over a region which has other regions as parts does not imply that a trope extended over this region has parts corresponding to the parts of the region. Simple tropes may extend over the region as a whole, and they are then candidates for being *extended simples* (Simons 2004).

A more familiar example would be colour. Suppose an object's surface is uniformly coloured: it might be a red tomato or sheet of white paper. If we accept that the whole surface has a single colour trope of the appropriate kind, then this trope would have a spatial extension with that area.

4. Does an extended trope have other tropes as parts?

Cut a uniformly red tomato in half. Ignoring the internal sections exposed by the cut, each of the halves has a uniform red surface which was previously part of the surface of the whole tomato. These two half-surfaces have their own colour-tropes if the whole tomato did. The two parts of the tomato's surface were that colour before the cut, so it seems reasonable that their colour-tropes existed before the cut and did not come into being with the cut. If so, then it seems likewise reasonable to conclude that the tropes of the two parts of the tomato's surface before the cut were parts of the trope of its whole surface, from which it follows that a trope can have another

trope as a proper part. If this is right, then some tropes are not simple, but have other tropes – here of the same kind – as parts.

Suppose we wish to maintain the view that all tropes are simple in the face of this argument. What options are there?

The first is the Aristotelian view that the two partial tropes only come into existence when the cut is made, and at that time the trope of the whole surface ceases to exist, so at no time is one trope part of another. The trope is a trope of the maximal surface of that colour, it is itself an extended simple, and its existence *qua* extended simple explains the uniformity of the surface colour. On this account, the parts of the surface do not have their own personal colour-tropes, but rather borrow their colour from the whole trope. This however goes against the tropist idea that when several particulars are exactly alike in some respect, it is because each particular has its own trope, where these tropes are simply exactly alike. Turning the point around, suppose we accept (I do not, but let it pass) that there is a single maximal object consisting of all the objects having that particular colour. Then only it would have a colour-trope and all the smaller ones would get their colour from it. But this is hardly different from saying that the unique whole trope is a universal, going against the grain of trope nominalism. This first option is then, if not inconsistent, certainly uncongenial to a tropist.

A second option is to deny that there are extended surface colour-tropes at all, but that surface colour results from the aggregation of other qualities, not themselves perceptible, which are located among the small parts of the surface. Colour is then not a basic property, and uniform colour results from a (near-)uniform distribution of the micro-tropes resulting more macroscopically in colour. This has the advantage of being closer to the scientific account of what constitutes surface colour, namely surface reflectance, which turns on the

natures of the molecules at or near a concretum's surface, with whatever tropes of their own that this involves. The part–whole relations required here are not among tropes themselves, but among the smaller concreta which are parts of the larger body.

Aggregation is in any case a feature that should command assent in other cases. Take the mass of a body, such as our tomato. It results from the aggregation of the masses of its basic parts, and is quantitatively the sum of those masses minus the small amount of binding energy it takes to hold the body and its parts together. There is no need to postulate an additional total mass-trope of the body as a whole. Similar remarks apply to other summative characteristics such as area and volume – if their account requires tropes (I am not sure it does).

5. Gestalt-tropes

Some characteristics of concreta appear not to derive from aggregation of characteristics of their parts, but to characterise the whole directly, albeit requiring some partial foundation in the concretum's parts. If there are tropes that account for such characteristics, it is appropriate to call them *Gestalt-tropes*, indeed as a tropist interpretation of Ehrenfels's original *Gestaltqualitäten* (1890). We consider two plausible examples: shape, and melody (the latter being Ehrenfels's own Machian example). Note that we are not here talking about our *perception* of shape or of melody, but of shape and melody themselves.

Consider a spherical object, such as a snooker ball. While its mass and colour arise by aggregation, its spherical shape does not derive from the shapes of its parts, but characterises the ball as a whole. It

is spherical because of the relative spatial dispositions of its smaller parts, but for the parts to be so disposed is just what it is for the whole to be spherical. As and when an object's parts are disposed in *that* way, it is spherical. But its being spherical is a resultant distinctive feature characterising it as a whole, and there seems no reason not to count that feature as a trope, which is then a Gestalt-trope.

A *very* simple case of a melody being played would be a non-pianist (myself, for example) picking out a tune with one finger on a piano, Sir Hubert Parry's *Jerusalem*, say. That same tune can be played many times, by many people, on different instruments and in different keys. What makes the various concrete events playings of the tune *Jerusalem* consists in the intervallic and durational interrelations of the successive tones produced. As in the case of spatial shape, the overall character of the melody accrues to it as a whole, and again there is no reason not to take this character as a distinctive Gestalt-trope.

Nevertheless, the existence of such Gestalt-tropes in no way supports the idea that one trope can be part of another, precisely because the Gestalt-trope characterises the whole object and not its parts or its tropes.

6. Fractal tropes?

A more intriguing and potentially worrying possibility would be shapes or other Gestalten that are self-similar or fractal. If an object had a shape which is a *fractal* Gestalt-trope, then it would have parts which are exactly similar to that of the whole, so would have infinitely many exactly similar Gestalt-tropes "inside" itself. But once again they would be (exactly similar) *tropes of parts*, not tropes *that*

are parts of other tropes. The case is similar to that of a spherical ball and one of its hemispheres: the spherical shape characterises the whole ball, the hemispherical shape characterises half of the ball, *not* the shape of the whole ball. The difference in the case of fractals is that some parts are exactly similar in shape to their wholes, and so on down *ad infinitum*. The possibility that one Gestalt-trope be part of another only appears plausible if we forget there is an object with the shape and imagine that all there is to the fractal object is the shape – which of course is how mathematicians describe things.

There are no exact fractals in nature: despite the existence of objects which approximate fractals to a few degrees, such as snowflakes, proper fractals are infinitely complex, and therefore mathematical, not natural, so trope theory does not realistically apply. How mathematical abstracta and tropes might consort together is, for a nominalist, an uninteresting and merely hypothetical matter.

7. Conclusion

There may be other cases lending some plausibility to the idea that one trope can be a proper part of another. I have not been able to think of any – but that does not mean there are none. At any rate, for the cases that I have surveyed, there is no convincing argument for one trope's being part of another. So we may continue to hold, as a reasonable hypothesis, that tropes are indeed simple.

References

Ehrenfels, C. von (1890) *Über „Gestaltqualitäten". Vierteljahrsschrift für wissenschaftliche Philosophie* 14, 242–292.

Martin, C. B. (1980) Substance Substantiated. *Australasian Journal of Philosophy* 58, 3–10.

Maurin, A-S. (2002) *If Tropes*. Dordrecht: Kluwer.

Simons, P. M. (1987) *Parts*. Oxford: Oxford University Press.

—— (2004). Extended Simples: A Third Way between Atoms and Gunk. *The Monist* 87, 371–384.

Tropes are Gunky – On the Relations Between Trope Theory, Monism, and Particularism

Fredrik Stjernberg

There is an argument showing that trope theory, as developed and described in for instance Williams (1997) and Maurin (2002), is problematic, or at least that tropes turn out to be complicated in an unexpected manner (Stjernberg 2008). The argument shows that tropes cannot be atoms – hence (perhaps) not the "elements of being", as D. C. Williams put it. This argument can be given a twist, to show that tropes are *gunky*, in the technical sense described in Schaffer (2010). This would provide a new argument for *monism*, at least for what Schaffer calls priority monism: "Priority monism … is the doctrine that exactly one concrete object token is basic, and is equivalent to the classical doctrine that the whole is prior to its (proper) parts" (Schaffer 2018, Introduction).

First, the argument.[1] Second, the connection with gunk, and finally a few thoughts about the situation. The end result is perhaps

[1] What follows is in essence the argument in Stjernberg (2008). I have changed my mind a bit about how to react to the argument, however.

not as clear-cut as I have led the reader to think. More work would have to be done here.

1. Tropes and atomicity

How are we to explain that different objects can share one and the same property? This is a problem with a long history, going back to Plato. According to trope theory, a sentence like "John is happy" is not to be analysed as attributing a property to some object (John) that can have several different properties. Trope theory suggests that the correct analysis is by talking about abstract particulars (tropes), in this case John's happiness. Then our test sentence is true iff that trope is an element in two sets: the set making up John (explaining why we are talking about *John's* happiness) and the set of happiness tropes (explaining why we can say that both John and Lisa can be happy). The two sets serve different purposes. One set makes up the object we are talking about, the other is a set of similar tropes, as John's happiness, Lisa's happiness, and so on. The first set is also called the *concurrence set* of tropes, *Con*{*a*}; the second set is the *similarity set* of tropes, *Sim*{*a*}. *Con*{*a*} makes up the individuals we are talking about (since *Con*{*a*} consists of John's happiness, his tallness, his colour, and so on), and *Sim*{*a*} is the set of tropes that together make up *F*.

A bit more formally, we can provide truth-conditions for a sentence "*a* is *F*" in the following way:[2]

[2] An analysis along these lines is provided in Bacon (1989).

1. "*a* is *F*" is true iff there is a trope *b* such that *b* is an element in the intersection of *Con*{*a*} and *Sim*{*F*}

"John is happy" is true iff there is a trope *b* which can be found in both set set of tropes making up John and in the set of tropes comprising all the happinesses in the world. Now consider the sentence "*a* is not composed of tropes", which in effect says that *a* is an atom, perhaps the ultimate trope, something not consisting of further tropes. Then the truth conditions for that sentence would be:

2. "*a* is not composed of tropes" is true iff it is not the case that "*a* is composed of tropes" is true

leading to

3. "*a* is not composed of tropes" is true iff it is not the case that there is a trope *b* such that *b* is an element in the intersection of *Con*{*a*} and *Sim*{composed of tropes}

The property *not being composed of tropes* can also be expressed as *being atomic*,[3] so we have the equivalence

4. "*a* is atomic" is true iff it is not the case that there is a trope *b* such that *b* is an element in the intersection of *Con*{*a*} and *Sim*{composed of tropes}

or we could go for a direct analysis of what "*a* is atomic" might mean:

5. "*a* is atomic" is true iff there is a trope *b* such that *b* is an element in the intersection of *Con*{*a*} and *Sim*{atomic}

Combining these two equivalences, we get

6. there is a trope *b* such that *b* is an element in the intersection of *Con*{*a*} and *Sim*{atomic} iff it is not the case that there is

[3] Can it really? More on this question below.

a trope *b* such that *b* is an element in the intersection of *Con*{*a*} and *Sim*{composed of tropes}

If the property *being atomic* is the same as the property *not being composed of tropes*, then *Sim*{atomic}=*Sim*{not being composed of tropes}. We can call this set *A*. Then we get

7. There is a trope *b* such that *b* is an element in the intersection of *Con*{*a*} and *A* iff it is not the case that there is a trope *b* such that *b* is an element in the intersection of *Con*{*a*} and *A*

We can call *Con*{*a*} *B*. Then we get

8. There is a trope *b* such that *b* is an element in the intersection of *B* and *A* iff it is not the case that there is a trope *b* such that *b* is an element in the intersection of *B* and *A*

This is the paradoxical

9. $\exists b(b \in A \cap B) \leftrightarrow \in \neg\exists b(b \in A \cap B)$

which spells bad news for trope theory. The question is, how bad? At the very least, this shows that more has to be done, that the proposed truth-conditions for sentences of the type "*a* is *F*" cannot be the whole story.

In fact, the above formalized argument is a bit of window dressing. There is a more direct and simple argument leading in the same direction.

Assume that something is atomic. Then it has no parts – after all, that is what "being atomic" means. But if it has no parts, the trope-theoretic analysis doesn't work, since it expressly analyses predications by distinguishing parts in objects. What is to be done here?

One way out is to try to distinguish various senses of "being atomic". Perhaps physical atomicity and metaphysical atomicity can

and should be distinguished. The most basic, and for the trope theorist most relevant, sense is arguably a metaphysical sense, since trope theory is a theory about predication, not a contribution to physics. Then metaphysical atomicity would mean that something has no metaphysical parts, or no parts that surface in an analysis. Given the above argument, this notion is in trouble – at the very least, we have to say that there are no metaphysical atoms. The notion of being atomic is simply contradictory. Hence atomism, or particularism, is false. Let's see where this thought leads.

The trope-theoretic version of this kind of thought says that atomism is false, and that everything is made up of tropes, *ad infinitum*. In this case, there are no basic building blocks in the world. This can perhaps sound like an *ad hoc* response to the above argument: it is the only way for the trope theorist to avoid the unpleasant consequences of the argument. But perhaps we can say more to support this view? After all, the debate between metaphysical monism and particularism has deep historical roots.

2. Monism on the cheap?

The above argument is quite close in spirit to that of Schaffer's *argument from gunk* (Schaffer 2010). Gunk is something non-atomic, something where every part of it has parts of its own, and so on, lacking ultimate parts. Bringing two lines of thought together, we can now say that *tropes are gunky*:

(1) Either the ultimate parts must be basic at all worlds, or the ultimate whole must be basic at all worlds.

(2) There are gunky worlds without ultimate parts (and hence no ultimate parts to be basic at those worlds).

(3) The ultimate whole must be basic at all worlds.[4]

This is clearly a valid argument, so the truth of (10) and (11) are what remains to discuss. A full discussion would have to be the subject of some other paper.

Tropes are gunky, so atomism cannot be true, so monism must be the true theory about the world. At least I think this is a new argument for monism. But it is conditional times two: *if* trope theory is the best way to handle the one-over-many problem, and *if* truth-conditions for sentences of the form *Fa* are to be given as suggested in the previous section, then tropes cannot be atoms. And if the world cannot contain atoms, there are no particulars making up the world. Hence monism is true. Determination basically is a top-down affair – the world, a whole, is that which determines all the other truths.

Against this, a particularist can say many things. One is to attack the assumptions behind the conditional reasoning. This is not the place to do that. Another is to try and turn the argument around. This is perhaps a more promising strategy, or at least what I will look (very briefly) at in the next section.

3. … or perhaps not

Turning the argument around is the idea that we can argue to show that the world is *junky* – the idea that every whole is part of some

[4] This is the argument as presented in Schaffer (2018) § 3.2.3.

bigger whole (Bohn 2009). If this is the case, then there is no ultimate whole that grounds everything else. If the world is gunky, no ultimate particulars exist, and if the world is junky, no ultimate wholes exist. Why would anyone think that there could be no ultimate wholes? Again, a fuller discussion would have to wait for some other time.

The most intriguing way forward here is that talk about ultimate wholes appears to talk about *everything*, in some unqualified way which puts us on a sure path to paradox. What is this complete whole that is supposed to be prior to its parts? This would have to be an absolute infinity: nothing greater of which it is a part could be conceived. But now we must tread carefully.

The set theoretic paradoxes that keep threatening to arise when we try to talk about everything can perhaps – perhaps – be tamed, but the way they keep popping up should at least give the over-confident monist pause for thought.[5] Talking about everything has costs of its own.

Back to the drawing board.

References

Bacon, J. (1989) A Single Primitive Trope Relation, *Journal of Philosophical Logic* 18, 141–154.

Bohn, E. D. (2009) Must There Be a Top Level? *Philosophical Quarterly* 59, 193–201.

Cartwright, R. (1994) Speaking of everything, *Nous* 28(1),1–20.

Filin Karlsson, M. (2017) *All There Is. On the semantics of quantification over absolutely everything*. Göteborg: Acta Universitatis Gothoburgensis.

[5] For a useful introduction, see Filin Karlsson (2017). See also Cartwright (1994). Priest (2002) has much of interest in his discussion of the set-theoretic paradoxes, especially in chapters 8 and 11.

Maurin, A.-S. (2002) *If Tropes*, Dordrecht: Kluwer Academic Publishers.

—— (2018), Tropes, *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2018/entries/tropes/>.

Priest, G. (2002) *Beyond the Limits of Thought*, Oxford: Oxford University Press.

Schaffer, J. (2010) Monism: The Priority of the Whole, *Philosophical Review* 119(1), 31–76.

—— (2018), Monism, *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2018/entries/monism/>.

Stjernberg, F. (2008), Trope theory and atomic objects, *Logic and Logical Philosophy* 17, 275–281.

Williams, D.C. (1997), On the elements of being: I. In Mellor, D.H., and Oliver, A., (ed.), *Properties*, Oxford: Oxford University Press.

The Bottom-Up, the Top-Down, and the Maurinian Analysis of Similarity Classes

Ingvar Johansson

1. Introductory words about philosophical discussions

When relaxing, after an intense discussion with a philosophical opponent, one may happily conclude: "Nice, I don't have to rethink anything." Or, one may end up a bit distressed saying: "Damned, I have to revise some of my views." Anna-Sofia Maurin is an opponent of mine with respect to the existence of universals; and we have had a journal exchange about it (Johansson 2007, 2010; Maurin 2007, 2009, 2010). I am since long quite relaxed about the outcome: "Let's agree that we disagree."

However, when this state entered my mind, it was accompanied neither by the reaction "Nice" nor by "Damned." Instead it was: "Interesting, I need not revise my belief in universals, but there is a lacuna in my defense that I ought to think more about."

One lacuna I became aware of already when reading what made me start the discussion, namely (Maurin 2007). In my defenses of universals, I should have added some words about the distinction between vicious and benign infinite regresses; not written as if all are vicious. Later, because of the discussion, I also realized that I ought to say more about the relation *exact similarity (resemblance)*, since it plays an essential role in my argument for the existence of universals. Now I will do so.

For brevity's sake, I will in what follows often abstract "exact" away from "exact similarity (resemblance)," and talk simply of *the similarity (resemblance) relation*.

Also, I will focus on where the relation has two distinct relata, neither more nor less. Since the binary similarity relation is transitive, it can be used to analyze similarity classes with more than two members. On the other hand, since the relation predicate "similar" is reflexive, it can be applied to a single entity, too, as in "*a* is similar to *a*." In such cases, however, we meet only empty tautological truths, and they cannot be used to analyze anything.

Soon, I will say some explanatory words about the expressions "bottom-up" and "top-down," but in the next paragraph I use them unexplained.

Similarity classes are constituted by a similarity relation in tandem with the members of the class. In my opinion, the similarity relation has to be understood as emerging bottom-up from the members of the class. Resemblance nominalists, to the contrary, make the members of the class emerge as such top-down from the similarity relation to its relata. Maurin tries with her trope theory to steer a middle course between these to her Scylla and Charybdis in the ontological ocean.

2. Maurin, the trope nominalist, and I, the immanent realist

The central difference between Maurin's ontological views and mine can be stated thus: she claims that there is no need to postulate property and relation universals, tropes can do the work that universals do for realists, whereas I think that we cannot make ontological sense of the universe in which we live without postulating immanently existing universals. She is a trope nominalist of sorts; first put forward in (2002). I am an immanent realist of sorts; best presented in (2014). I defend the existence not only of property and relation universals, but of tropes and relation-tropes, too. I am not denying her view that there are tropes, but she is denying mine that there are universals. A universals-plus-tropes ontology can also be found in philosophers such as E. Husserl, J. Cook Wilson, D.W. Mertz, E.J. Lowe, and J. Heil.

In her construction of what from my point of view is a substitute for property universals, Maurin claims that property universals, correctly conceived, are *equivalence classes of tropes held together by a similarity relation*. Her complete ontology, therefore, contains not only tropes, but *in some sense* classes and similarity relations, too: "all entities that exist besides tropes are constructed from tropes […] saying that there is nothing but tropes is to say that tropes are the only metaphysically *fundamental* entities" (Maurin 2002: 5).

Maurin uses the class concept as if it is not in need of discussion; unlike Quine, though, she cannot possibly regard classes as universals. However, I leave the ontology of classes outside of this paper, and focus here only on the similarity relation.[1]

[1] I discuss Quine in my (2016), and my view of classes is put forward in (2015).

3. The notions of bottom-up and top-down

The metaphors bottom-up and top-down can be exchanged for more precise concepts, but depending on background assumptions such precisifications may differ ontologically. I would nonetheless like to say some words about how I look upon the metaphors.

In Platonist ontologies, the top-down notion is fitting. All identities/essences/natures down in our spatiotemporal world are regarded as derived from eternal entities up in a mind-independently existing world of ideas. This-worldly entities are claimed to get their qualitative identities thanks to a kind of participation in essences existing in another realm. Such ontologies I see no reason to bother about.

However, the top-down notion comes naturally also in ontologies that posit a creative mind as a top-level entity. In such ontologies, a mind is regarded as being able to create and project identities down on pre-given entities that in themselves lack such an identity. For instance, a mind may create and project a resemblance down on two entities that in themselves are not similar. Call such a view *subjective emergentism*.

Carefully seen, however, this view has a consequence that places it outside the topic of this paper. In subjective emergentism, the seemingly binary similarity relation is turned into a *ternary* relation. There is no longer a direct and unmediated binary similarity relation between two relata. Instead, there is a three-term relation that brings in also a mind. Such similarities can be represented by the three-term expression "similarity(relatum$_1$, relatum$_2$, a mind)."

The kind of minds and ternary relations just mentioned are absent from Maurin's analyses, and also from those of resemblance nominalists'. Maurin says: "unless otherwise indicated, the resemblance

we are dealing with is *objective* and *primitive*" (2002: 80). Therefore, I confine my discussions to the truly binary similarity relation.

In common sense, the bottom-up conception of relations comes naturally. Normally, we discuss possible relations by investigating the relata of the presumed relations. With respect to the similarity relation, this means that if the relata have the same kind of property, then they are similar; otherwise not. As far as I am concerned, common sense may well be ascribed an implicit belief in property universals. This does not mean that I think that common sense is a reliable source for finding truths. To the contrary. But I happen to believe that a non-commonsensical proof can be given to the effect that there are both tropes and property universals (Johansson 2014).

Contemporary analytic philosophy contains attempts to find an asymmetric relation that might be used to characterize the asymmetry contained in both the bottom-up and the top-down conceptions of similarity. This is not the place to discuss them, but I will briefly mention two proposals and my preferred alternative.

First, if the notion *in virtue of* is taken as primitive, then in the bottom-up approach the similarity relation exists in virtue of the nature of the relata. In the top-down approach, the natures of the relata exist in virtue of the similarity relation.

Second, if a relation of *metaphysical grounding* is taken as ontologically basic (and not as defined by an in-virtue-of relation), then in the bottom-up approach similarity relations are regarded as grounded in the nature of their relata, whereas in the top-down approach the natures of the relata are regarded as grounded in the similarity relation.

Third, I have since the mid-80s worked with a notion somewhat similar to the ones mentioned. I have taken it from Husserl, and I call it *one-sided existential dependence* (Johansson 2004: ch. 9.3). Us-

ing this notion, the bottom-up approach claims that similarity relations are one-sidedly dependent on their relata, which means that similarities cannot possibly exist without their relata, whereas each of these relata can exist even if the relation does not exist. The top-down approach, conversely, claims that neither relatum can exist with its nature if the similarity relation does not exist, but the similarity relation can in some way exist without the relata.

Using a wide concept of explanation, one may say that in the bottom-up approach the similarity relation is meant to be explained by the nature of its relata, and that in the top-down approach the natures of the relata are meant to be explained by the similarity relation.

4. Maurin, I, and resemblance nominalism

I take resemblance nominalism to be definable as essentially containing the top-down approach. "In short, in Resemblance Nominalism resemblance among concrete particulars is not *explained*, but is used to explain the properties of these particulars" (Rodriguez-Pereyra 2002: 26).[2] The general approach may then be specified in different ways by different philosophers. One philosopher may try to use the in-virtue-of relation, another may use grounding, and a third one

[2] Despite calling himself a resemblance nominalist according to the characterization presented, Rodriguez-Pereyra also makes the astonishing claim that there are no resemblances! Only as-if resemblances: "although there are resembling particulars, there is no entity over and above them that is their resemblance. […] But for ease and simplicity of exposition I shall talk throughout the book about resemblance as if it were an entity, that is, a relation" (*ibid.*: 62). This statement simply contradicts his view that resemblance facts are objective (*ibid.*). As-if facts are subjective mind-phenomena. As has been noted before, his position is "an ontological version of the Indian rope trick" (Svennerlind 2008: 135n).

one-sided existential dependence. I think all three kinds of resemblance nominalism founder when confronted with the following simple line of thought.

Most ontologists accept in their argumentation far-reaching thought experiments, and I am no exception. Let us first imagine a two-trope world where there are only two tropes, t and $t^{\#}$, and no similarities whatsoever; call it world-(A). Let us then in thought enter another two-trope world, world-(B), where $t^{\#}$ has been exchanged for t^{*}, and where t and t^{*} are exactly similar. This world contains not only two tropes, but a similarity-trope, too. Third, let us imagine a one-trope world, world-(C), by deleting t^{*}; it cannot possibly contain a similarity relation with two distinct relata.

In the move from world-(A) to world-(C), it seems unproblematic to think of t as being the same in all the three worlds, and therefore as having the same nature in all three. What, then, are here the options for resemblance nominalists? For them to claim that t has no nature in world-(C) seems absurd. And for them to claim that t has a nature because "t is similar to t" makes their approach to relations with distinct relata superfluous. The sentence "x is similar to x" is true for all entities independently of their similarity to other entities. However, to argue for a superfluous analysis is absurd. Therefore, on either option, resemblance nominalism falls prey to a *reductio ad absurdum*.

What, then, does Maurin say about resemblance nominalism?

> [C]ompare the standard view [of tropes] with a view with which it is often confused: resemblance nominalism. On trope theory, tropes have the same nature if they resemble each other, and they resemble each other (or not) *in virtue of the (primitive) nature they each 'have'* (or 'are'). According to resemblance nominalism, on the other hand, two objects have the natures they do in virtue of the resemblance relations which obtain between them. This means that, whether

> they resemble or not, is not decided given the existence and nature of the objects themselves. Rather, the pattern of dependence is the other way around. […] Perhaps for that reason, resemblance nominalism has no explicit proponent among the trope theorists. (2018: endnote 28)

When, in the next section, I discuss Maurin's analysis of the similarity relation, I can do so quite confident that both of us find resemblance nominalism incredible.

5. The Maurinian analysis: invoking pseudo-additions and haecceity-natures

Let me use an example to make the discussion more concrete. Look at the spots within the square brackets: [● ●]. Each spot contains at least both a color-trope and a shape-trope. The color-tropes (being black) may be regarded as existing only in our perceptual fields, but the shape-tropes (being circular) may be ascribed a mind-independent existence. Now, whatever kind of tropes the reader would like to regard as being acceptable, let's call the left trope *t* and the right one *t**. The tropes are exactly similar, *S*.

With respect to the example, Maurin and I do on an *abstract semantic level* share the following views:

(i) *t* and *t** are two different tropes, and as such each has a nature;

(ii) they are related by a similarity-trope, *S*;

(iii) necessarily, if *t* and *t** exist, then *S* exists;

(iv) necessarily, if *S* exists, then there are two relata, *t* and *t**;

(v) *t* can exist even if *t** does not, and vice versa, which – together with (iv) – entails that *t* and *t** can exist separately even if there is no *S*.

The differences between us pop up when we specify the term "trope nature" in (i) and the second occurrence of "exist" in (iii). I start with the latter.

Maurin says: "Resemblance may be understood in one of two ways: either as a pseudo-addition or as a genuine relation-trope" (2005: 138). For me, the second occurrence of "exist" in (iii) has the same sense as the first, which means that for me *S* is a genuine relation-trope that exists just as much as the tropes do. Maurin, on the other hand, allows my use of "exist" to have two senses: *t* and *t** genuinely exist, but *S* only pseudo-exists. The presumed latter fact is what allows her to reject the bottom-up approach that statement (iii) may seem to imply. According to Maurin, there really is no "up," only a "pseudo-up." To her, *S* is not only (as for me) less ontologically fundamental than *t* and *t**, *S* is merely an ontological *pseudo-addition*. She says:

> This will leave us with the third and final view [which is Maurin's option], a view according to which exact resemblance should be ontologically characterized as a *pseudo-addition* to our basic metaphysical scheme. (2002: 94)

She is from the start well aware of the fact that the term "pseudo" may in ontological contexts sound odd: "the invocation of 'unreal' additions to solve philosophical problems may seem nothing short of magic" (*ibid.*). And then she goes on to defend her view (2002: 109–115). After criticism, she briefly defends the notion of pseudo-

addition again (2005: footnote 23). For good criticism, see (Svennerlind 2008: 129–142). In her SEP entry "Tropes," she uses the notion in passing: "the trope theorist had the option of treating the relation [resemblance] as a 'pseudo-addition'" (2018: sect. 3.2). To my mind, her argumentation proves only that *S* is less fundamental than *t* and *t**.

Next the term "trope nature." To me, a trope's nature is a universal, since I regard a trope as an instance of a universal. To Maurin, on the other hand, the nature of a trope is as particular as the trope itself. Moreover, she regards them as identical: "tropes are nothing but their particular nature" (2002: 93). At least in everyday talk, the term "nature" has connotations that make it refer to something repeatable. My use conforms to this, Maurin's does not. From medieval scholasticism I will now borrow the term "haecceity," which means thisness in contrast to universality. I will use it, for the sake of verbal clarity, to call Maurin's trope natures *haecceity-natures*. That is, natures that, necessarily, can be the nature of only one entity.

Looking at the history of European philosophy, one finds some entities that are regarded as having a haecceity-nature. First, Plato's mind-independently existing Ideas. There cannot be two Ideas that are exactly similar. Then there are Aristotle's the Unmoved mover and God in scholastic philosophy. All these entities have this feature: *nothing else can possibly resemble them exactly*. Exact similarity between trope natures, however, is absolutely central to Maurin's trope theory: "similarity relations do not afford us, but are rather a consequence of, the nature of individual tropes" (2010: 47; my translation). In spite of the fact that the haecceity-natures of tropes differ from the other ones mentioned, Maurin never explicates her notion of trope nature, i.e., a trope's particularity. After some reasoning, she

concludes: "we must therefore simply accept the particularity of the trope as primitive" (2002: 21).

Conclusion: there is a lacuna in Maurin's defense of the view that each and every trope has its own haecceity-nature.

6. Concluding Russellian words about discussions of relations

Bertrand Russell, who was a firm believer in the existence of both property and relation universals, put forward the following claim in 1924:

> The question of relations is one of the most important that arise in philosophy, as most other issues turn on it: monism and pluralism; the question whether anything is wholly true except the whole of truth, or wholly real except the whole of reality; idealism and realism in some of their forms; perhaps the very existence of philosophy as a subject distinct from science and possessing a method of its own. (Russell 1956: 333)

If what Russell says is true, then I hope to have shown that his list of opposing positions can definitely be augmented by the one between immanent realism and trope nominalism.

Acknowledgements

For paper-improving comments I would like to thank Kevin Mulligan, Robin Stenwall, and Christer Svennerlind.

References

Johansson, I. (2004 [1989]) *Ontological Investigations*, Frankfurt: Ontos Verlag.

Johansson, I. (2007) Maurin och Russells regress, *Filosofisk Tidskrift* 2/2007, 60–62.

Johansson, I. (2010) Maurins regresser: en replik, *Filosofisk tidskrift* 1/2010, 45–46.

Johansson, I. (2014) Proof of the Existence of Universals—from a Fallibilist, In Bonino, G., Jesson, G. & Cumpa, J. (ed.), *Defending Realism. Ontological and Epistemological Investigations* (pp. 51–61). Berlin: de Gruyter,.

Johansson, I. (2015) Collections as One-and-Many—On the Nature of Numbers, In Lapointe, S. (ed.), *Themes from Ontology, Mind, and Logic. Present and Past. Essays in Honour of Peter Simons* (pp. 17–58). Leiden: Brill Rodopoi.

Johansson, I. (2016) Against Fantology Again, In Zaibert, L. (ed.), *The Theory and Practice of Ontology* (pp. 25–43). London: Palgrave Mamillan.

Maurin, A-S. (2002) *If Tropes*, Dordrecht: Kluwer.

Maurin, A-S. (2005) Same but Different, *Metaphysica* 6(1), 129–145.

Maurin, A-S. (2007) Infinite Regress: Virtue or Vice, In Rønnow-Rasmussen, T. et al., *Hommage à Wlodek*. URL = <https://www.fil.lu.se/hommageawlodek/index.htm>.

Maurin, A-S. (2009) Russells regress: en replik, *Filosofisk tidskrift* 3/2009, 44–54.

Maurin, A-S. (2010) Svar på svar, *Filosofisk tidskrift* 1/2010, 47–48.

Maurin, A-S. (2018) Tropes. In Zalta, E.N. (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition), URL = <https://plato.stanford.edu/archives/sum2018/entries/tropes/>.

Rodriguez-Pereyra, G. (2002) *Resemblance Nominalism*, Oxford: Clarendon Press.

Russell, B. (1956) *Logic and Knowledge* (ed. Marsh R.C.), London: Allen & Unwin.

Svennerlind, C. (2008) Examination of A-S. Maurin's Ontology, In Svennerlind C., *Moderate Nominalism and Moderate Realism* (pp. 77–159). Gothenburg: Acta Universitatis Gothoburgensis.

A Little Dialogue on Extensionality (and music)

Jeroen Smid

Anna: So we agree that objects are bundles of tropes?

Sofia: Trope bundle theory is ontologically very parsimonious and it nicely avoids Bradley's regress. It definitely has a lot going for it (Maurin 2002, 2010, 2011).

Anna: But?

Sofia: Well, trope theory concerns one area of metaphysics but there are other areas, too. If my favourite theory in one of those other areas doesn't fit with trope bundle theory, I might give up on tropes.

Anna: "*If*…" Well, trope theory fits well with extensional mereology, especially if tropes are literally *parts* of an object. For it makes sense then to say that having the same part (i.e., tropes) is sufficient for identity. We can then have a mereological trope bundle theory.

Sofia: I thought about this, too; but I do have a worry. For if instead we have good reasons to deny extensionality, we might have to say that different objects can *share* a single trope. The statue and the clay, for example, might have one and the same mass trope although they are distinct. But tropes, it seems, are not shareable. There are other ways to respond…

Anna (interrupting): Can we please not talk about the statue and the freaking clay! The example bores me and honestly, I don't think it shows two objects can have exactly the same parts (Smid 2017).

Sofia: Fair enough! And regardless of the question whether non-extensional mereologies are compatible with (mereological) trope bundle theories, I like fussing about extensionality.

Anna: Let's first get clear about parthood. When I talk about parthood I mean *proper parthood*: the relation holding, for example, between my hand and me. And by 'part' I mean *proper or improper part*, i.e. the thing itself.

Sofia: So, proper part is an irreflexive, transitive, and asymmetric relation.

Anna: Best to stay neutral (for now) on asymmetry—some philosophers think distinct objects can be proper parts of each other (Thomson 1983, 1998; Cotnoir 2010, 2016).

Sofia: For example, the statue and the…, err, sorry, I mean, a bottle and its plastic are then mutual parts?

Anna: Exactly. So shall we say that proper parthood is irreflexive and transitive?

Sofia: But the asymmetry of proper parthood then follows: if x is a proper part of y and y is a proper part of x, then—by transitivity—x is a proper part of x. But this is impossible given irreflexivity.

Anna: Yes, of course. So, maybe we should say that either proper parthood is irreflexive and transitive (and thus asymmetric) or it is irreflexive and restrictedly transitive?

Sofia: Restrictedly transitive?

Anna: Yes: if x is a proper part of y and z is not identical with y and z is a proper part of x, then z is a proper part of y.

Sofia: That seems fine. And, maybe as a third option, we should allow for proper parthood to be transitive, but neither irreflexive nor asymmetric.

Anna: I doubt we are still talking about a single notion of parthood…

Sofia: If necessary, we can formulate different theories for different parthood relations. In any case, whether proper parthood is a strict partial order or something weaker, we need some remainder principle that does justice to the idea that a proper part of an object is not the whole object. If you take away a (proper) part of an object, there should be something that remains.

Anna: Plenty to choose from:

> *Quasi-Supplementation*: If x is a proper part of y, then y has two parts, w and z, such that w and z do not overlap.
>
> *Weak Supplementation*: If x is a proper part of y, then y has a part z that does not overlap x.
>
> *Strong Supplementation*: If x is not a part of y, then some part of x does not overlap y.

Sofia: I assume that 'overlap' simply means sharing an improper part?

Anna: Yes, so I overlap myself; Skåne overlaps Sweden because it is a proper part of Sweden; and your flat overlaps with your neighbours' flat because they share a wall.

Sofia: A rather thin wall…

Anna: At least you have a *stuga*.

Sofia: True that. Anyway, do we have to determine "The Only True Supplementation Principle" before we can address the question of extensionality?

Anna: Nope. Let's just agree that any counterexample to extensionality must respect at least one of these supplementation principles. Otherwise the counterexample proves too much.

Sofia: Why?

Anna: Well, even those who deny extensionality think at least one of these three principles must be true. (Simons 1987: 116; Cotnoir 2016: 125–127; Gilmore forthcoming) Which makes sense, because it seems that we are only willing to say that x is a proper part of y if x is not *all* of y: something besides x should also be part of y.

Sofia: Sure. So we consider a counterexample to extensionality legit only if it does not violate all three supplementation principles.

Anna: And, to be on the same page, the extensionality principle states:

> *Extensionality*: If x and y have proper parts and they have all the same proper parts, then $x = y$.

Sofia: Moreover our counterexamples shouldn't concern sculptures.

Anna: Please!

***

Sofia: I think I found some counterexamples (Walters 2019: 36).

Anna: Do tell.

Sofia: So, consider the words 'no' and 'on'. They are distinct, yet have the same parts.

Anna: I assume you are talking about the word types here, not their tokens since the tokens have different (token) parts.

Sofia: Of course.

Anna: Hmm, but then consider the word type 'feel'. If 'ee' is a proper part of 'feel', then 'e' is also a proper part of 'ee'. But this example then violates both *Quasi* and *Weak Supplementation* since 'ee' only has a *single* proper part, 'e' (Varzi 2019: sec. 3.2).

Anna: But *Strong Supplementation* is not violated by this example?

Sofia: Well, since everything that is part of 'ee' overlaps 'e' we get—by the contrapositive of *Strong Supplementation*—that 'ee' is a part of 'e'. That's just wrong.

Sofia: OK, but what about lexemes, for example, the word 'bank' can mean 'financial institution' or 'land alongside a river' (Walters 2019: 36)?

Anna: Sure, the word 'bank' is homonymous and can thus express two lexemes, both spelled 'bank' but one meaning FINANCIAL INSTITUTION and the other meaning RIVER BANK. But what are the *parts* of those lexemes?

Sofia: Surely, the letters are parts of the lexemes.

Anna: I think you are confusing words and lexemes here. The words 'go', 'going', and 'went' are all forms—inflictions—of a single lexeme, GO. Lexemes are instead, I think, best regarded as *simple* objects; the meaningful units of a language that can be inflicted by different words of that language. However, suppose instead that letters are parts of lexemes. In that case, we end up in the same situation as before, right? If letters and lexemes are types, then all supplementation principles fail because such types allow for objects that have a single proper part.

Sofia: What if we consider the tokens instead? So that every word token of 'bank' expresses two lexemes.

Anna: Well, then there is still no problem with *Extensionality*: we have four letter tokens that are part of a single word token—'bank'—that happens to have two meanings.

Sofia: This is proving to be more difficult than I thought.

Anna: I know. But you inspired me. Consider works of art (Walters 2019: 36).

Sofia: Are *you* going to be the one to bring up a sculpture after all?

Anna: Of course not! I was thinking of music. Consider two pieces of music that are, note for note, exactly the same. They are, however, distinct pieces of music because different composers wrote them and, as such, they have different qualities.

Sofia: Not sure I agree it is possible that *two* distinct pieces of music can be note for note the same (Dodd 2007). But suppose it is possible, for sake of the argument.

Anna: Well, this is sort of the whole argument: the pieces of music are composed of the same notes (i.e., they have the same parts) but they are nonetheless distinct.

Sofia: OK, bear with me, I am going to get metaphysical… what *is* a piece of music?

Anna: Aristotle, here we come…

Sofia: I mean, is a piece of music a performance, or the fusion of its performances? Is it its score? Or is it an abstract type, existing independently of its performances and score?

Anna: Does it matter?

Sofia: Of course! Say we identify a piece of music with its performances and thus say that, since the two pieces are distinct, performances of the one are not performances of the other, and *vice versa*. By parity of reasoning the pieces of music then don't have the same parts either. The first few bars played of one performance are only part of that piece of music that is being performed, but not of the other piece.

Anna: Can't I say that every performance is a performance of *both* pieces of music? In that case they do have the same parts.

Sofia: Not so fast. Why did you say the pieces of music were distinct?

Anna: Well, they might have different aesthetic qualities: since the composers may be from different traditions, one piece may be *weirder*, *more controversial*, or *more upsetting* than the other (Levinson 1980: 11).

Sofia: *A maiore ad minus*.

Anna: *Omnia dicta fortiora si dicta Latina.*

Sofia: Sorry. I mean: if the pieces of music are distinct, then some of their parts are, too. It's not that the difference between the pieces (*if* there is one) is noticed only at the level of the whole. The part consisting of all but the final three notes is probably also *weirder*, *more controversial*, or *more upsetting* than the analogous part of the other piece. Hence, you should say that they do not have the same parts after all. And I could say the same of one particular note. That note may, in the first piece, have a different quality than in the second piece, because of the place the whole piece occupies in a musical tradition. So even on the smallest level there is a difference in their parts (Cf. Varzi 2008: 115ff).

Anna: What if I identify a piece of music instead with its score?

Sofia: So the scores have the same notes, but they are nonetheless distinct?

Anna: Yes, exactly.

Sofia: But how is this a counterexample to extensionality? I mean: if we talk about tokens of scores then fine, piece A is distinct from piece B. But no note written on one score is also, as a token, written on another score. So, again, they do not have the same parts.

Anna: Yeah, so let's say scores are types, not tokens.

Sofia: Well, then we end up in the same boat as with word types: a motif consisting of two or more of the same notes will have a proper part violating all three supplementation principles.

Anna: OK. This turns out to be much harder than I thought… *Fika?*

Sofia: Fika!

Acknowledgments

Thanks to Vetenskapsrådet, i.e. the Swedish Research Council, for generously funding my research; grant number 2017-06160_3. And thanks to David Liggins for discussion and also to Tobias Hansson Wahlberg for comments on an earlier draft.

References

Cotnoir, A. J. (2010) Antisymmetry & Non-Extensional Mereology, *The Philosophical Quarterly* 60(239), 396–405.

——— (2016) Does Universalism Entail Extensionalism? *Noûs* 50(1), 121–32.

Dodd, J. (2007) *Works of Music: An Essay in Ontology*, Oxford: Oxford University Press.

Gilmore, C. (forthcoming) Quasi-Supplementation, Plenitudinous Coincidentalism, and Gunk. In Garcia, R.K. (ed) *Substance: New Essays*, Philosophia Verlag.

Levinson, J. (1980) What a Musical Work Is, *The Journal of Philosophy* 77, 5–28.

Maurin, A.S. (2002) *If Tropes*, Dordrecht: Kluwer Academic Publishers.

——— (2010) Trope Theory and the Bradley Regress, *Synthese* 175(3), 311–326.

——— (2011) An Argument for the Existence of Tropes, *Erkenntnis* 74(1), 69–79.

Simons, P. (1987) *Parts: A Study in Ontology*, Oxford: Oxford University Press.

Smid, J. (2017) Material Constitution Is Ad Hoc, *Erkenntnis* 82(2), 305–25.

Thomson, J.J. (1983) Parthood and Identity across Time, *The Journal of Philosophy* LXXX (4), 201–220.

——— (1998) The Statue and the Clay, *Noûs* 32, 148–73.

Varzi, A. C. (2008) The Extensionality of Parthood and Composition, *The Philosophical Quarterly* 58 (230), 108–133.

——— (2019) Mereology. In Zalta, E.N. (ed.) *Stanford Encyclopedia of Philosophy*, Spring 2019 Edition. <https://plato.stanford.edu/archives/spr2019/entries/mereology/>.

Walters, L. (2019) Are The Statue and The Clay Mutual Parts? *Noûs* 53(1), 23–50.

Armstrong's Truthmaker Argument for the Existence of States of Affairs Revisited

Robin Stenwall

Consider David Armstrong's famous argument for why contingent predications require the existence of states of affairs in order to be made true (1997: 113-16).[1] Let some particular *a* instantiate universal *F*, and assume that there are possible worlds in which *a* is not *F*. Is there an entity which makes it true that *a* is *F*? In view of Truthmaker Necessitarianism (i.e. the principle according to which a truthmaker necessitates the truth of the proposition(s) it makes true),[2] the truthmaker cannot be *a* itself. The particular might not have been an *F*, in which case there will be possible worlds where *a* exists but where it is false that *a* is *F*.[3] Nor can the truthmaker be the pair of *a* and *F*-ness, since both entities could exist and it still not be

[1] What follows is a more elaborate version of one of the arguments given in Stenwall (2016).

[2] I will assume that propositions are the primary bearers of truth, although nothing essential hinges on this assumption.

[3] For illustrative purposes, I will engage in talk about possible worlds without thereby ontologically committing myself to such entities.

true that *a* is *F* (given that properties cannot exist unexemplified, *F* could be instantiated by some other particular *b*). Consequently, the argument goes, there has to exist something in addition to *a* and *F*-ness to make the proposition that *a* is *F* true. And if we are to believe Armstrong, this is the state of affairs *a's being F*, since in every possible world where that state of affairs exist, it will also be true that *a* is *F*.

With this in mind, let us follow Armstrong and take the proposition that the ball is red to be made true by the state of affairs, *the ball's being red*. Unlike the ball itself or the joint existence of the ball and redness, there are no possible worlds where that state of affairs exists, but where it fails to be true that the ball is red. The question is whether the necessity that the state confers is somehow a consequence of (or essentially related to) the truthmaking abilities of that state. Armstrong thinks that it is. For the state, he reasons, would not qualify as a truthmaker if it did not necessitate the truth of the proposition in question (2004: 5-7). But why is that? Why can't the necessity that the state confers rather be a consequence of its being both a necessary and sufficient condition for the transworld identity of states of affairs that they contain the same constituents organized in the same manner?[4]

Consider the case of the mad metaphysician, Milo. Milo thinks, contrary to Armstrong and very much contrary to fact, that the transworld identity of states of affairs does not supervene on their structural composition and, consequently, that what '*ball's being red*' refers to in this world could have been constituted by, say, the ball and blueness.[5] If Milo is correct, the non-structural state of affairs—

[4] Notice that this has to be assumed independently of Truthmaker Necessitarianism in order for Armstrong's argument to work.

[5] For the sake of brevity, I shall use the term 'non-structural states of affairs' to refer to states of affairs whose identity is not given by their structural composition.

call it '*Bob*'—would not necessitate the truth of the proposition in question, since there are worlds where *Bob* exists, but where it is nonetheless false that the ball *is* red. The non-structural state of affairs will only make the proposition true in worlds like the actual where *Bob* is constituted by the ball and redness. But Milo takes this to be a contingent fact. So there are, according to the mad metaphysician, possible worlds where the same exact state of affairs obtains, but where the proposition is false. Yet it is hard to see why this lack of necessitation on behalf of *Bob* would deprive the non-structural state of affairs the status of being a truthmaker for the proposition that the ball is red. For it is clear that the existence of *Bob* is enough *to make it true that the ball is red*—this, despite the lack of necessitation by that state of affairs. What goes on at possible worlds other than the actual is totally irrelevant to the question of what makes the proposition true in the world in which it is evaluated. That *Bob* could have been constituted by the ball and blueness has no bearing whatsoever on the truthmaking abilities of that state. Why should it? As long as the state of affairs makes it the case that the ball is red in this world, it would seem that there is no reason to take any otherworldly inhabitants into account.

Armstrong, however, thinks that a truthmaker whose existence merely contingently suffices to make a certain proposition true is 'incomplete':

> If it is said that the truthmaker for a truth could have failed to make the truth true, then we will surely think that the alleged truthmaker was insufficient by itself and requires to be supplemented in some way. A contingently sufficient truthmaker will [make it] true only in circumstances that obtain in this world. But then these circumstances, whatever they are, must be added to give the full truthmaker. (Armstrong 1997: 116)

But in the present case it makes no sense to require that *Bob* be supplemented with the actual circumstances. The non-structural state of affairs is only a truthmaker for the proposition that the ball is red under circumstances that obtain in this world, where the state of affairs is constituted by the ball and redness. But since by supposition it is true that the ball is red solely in virtue of *Bob*, it follows that there is nothing of relevance in addition to that state of affairs to be added. And if Armstrong means to say that something *has to exist* in this world that does not exist in a world where it is false that the ball is red,[6] and that this existence must be incorporated into the truthmaker, he begs the question. For then, to claim that the truthmaker is "insufficient by itself" amounts to nothing more than the claim that the truth of the proposition is not necessitated by that which makes it true (Skiles 2015: 737). But to infer that the "incomplete" truthmaker therefore be supplemented with the actual circumstances is to presuppose what the argument is meant to establish, namely Truthmaker Necessitarianism. We still need an independent argument showing that this supplement, assuming it exists, would have to be added to *Bob* in order for it to be true that the ball is red (Cameron 2008: 109-12). Until then, there is no reason to assume that truthmakers must necessitate the truth of that which they make true.

Let me be clear about what I have been arguing. First of all, I do not claim that the identity of states of affairs does not supervene on their structural composition—it most certainly does. However, the

[6] Milo, of course, denies that truth supervenes on *existence* in that he allows for possible worlds that are exactly alike with respect to what exists but that nonetheless differ with respect to what is true. However, Milo is not mad enough to deny that truth supervenes on *being*. He is in agreement with the likes of Lewis (2001) and others who think that for any two worlds exactly alike with respect to what exists and which properties (and relations) each of those existents exemplifies are also exactly alike with respect to what is true.

reason we should reject a non-structural identity condition for states of affairs is not because it would make the truthmaking relation contingent, but because it would make transworld identification of states of affairs, if not unintelligible, at least difficult. To conflate the role of states of affairs to be identifiable across possible worlds with their role to provide grounds for truth muddles the discussion and gives importance to modality where none is merited. Secondly, I do not deny that certain truths are made true by states of affairs. There are independent reasons to suggest that the proposition that the ball is red is made true by the *ball's being red*, rather than, say, the ball itself. But this has everything to do with the fact that the ball *is* red and nothing to do with the *ball's being red* necessitating the truth of the proposition that the ball is red. Suppose, for the purpose of a *reductio*, that the truth of the proposition that the ball is red is made true by the ball. The ball is not only red; it is also of a determinate mass, volume, shape and so forth. So if the ball does indeed make it true that the ball is red, it also makes true various other (non-relational) predications involving the ball. But this seems wrong from the perspective of truthmaker theory. The idea that truth depends on being is the idea that different truths depend on different portions of reality. And what makes it true that the ball is red is not what make is true that the ball has the shape it has. What makes it true that the ball is red is that it *is* red, and what makes it true that the ball has the shape it has is that it *is* ball-shaped (cf. Rodriguez-Pereyra 2005: 23). It might be that Armstrong is correct to point out that this 'is' is not the 'is' of identity, but that of instantiation—a fundamental, non-relational tie between particular and property (1978: 108-11). If so, then the instantiation of a property by a particular is simply the state of affairs itself. Thus, to say that the proposition that the ball is red is true in virtue of the ball instantiating redness is just to say that the proposition is made true by the state of

affairs the *ball's being red*. The state cannot exist and it be false that the ball is red. This neither is nor should be denied. What is and should be denied is that this necessity is somehow a consequence of, or essentially related to, the state *making it true* that the ball is red. The ball instantiating the colour it actually has (whatever colour that may be) is, in the world in which the evaluation is made, sufficient to make the corresponding attribution true. In this world the ball exists and is red. And this is enough to make it true that the ball is red, irrespectively of whatever colour the ball happens to exemplify in any other world of evaluation.

This not only goes to show that Armstrong's truthmaker argument for the existence of states of affairs rests on a faulty assumption, but also that Truthmaker Contingentism can be extended to include not only the negative (Stenwall 2017) and general truths (Cameron 2005 and Briggs 2012), but also the atomic truths. What is more, necessitarianism has no explanatory role to play in arguing that states of affairs must be included in the ontological catalogue, and the necessity that states of affairs confer on truth can be explained away by the fact that their transworld identity supervenes on their structural composition.

Acknowledgements

I would like to thank Tobias Hansson Wahlberg, Anna-Sofia Maurin and Ylwa Sjölin Wirling for their comments on this paper and Riksbankens Jubileumsfond (The Swedish Foundation for Humanities and Social Sciences) [grant number P16-0541:1] for their financial support.

References

Armstrong, D. M. (1978) *Universals & Scientific Realism Vol. 1*. Cambridge: Cambridge University Press.

Armstrong, D. M. (1997) *A World of States of Affairs*. Cambridge: Cambridge University Press.

Armstrong, D. M. (2004) *Truth and Truthmakers*. Cambridge: Cambridge University Press.

Briggs, R. (2012) Truthmaking without Necessitation. *Synthese* 189, 11-28.

Cameron, R. (2005) Truthmaker Necessitarianism and Maximalism. *Logique et Analyse* 48, 43-56.

Cameron, R. (2008) Truthmakers, Realism and Ontology. In *Being: Contemporary Developments in Metaphysics*, ed. Robin LePoidevin, Royal Institute of Philosophy Supplement 83, 107-28.

Lewis, D. (2001) Truthmaking and Difference-Making. *Noûs* 35, 602-15.

Rodriguez-Pereyra, G. (2005) Why Truthmakers. In *Truthmakers: The Contemporary Debate*, ed. Helen Beebee and Julian Dodd, Oxford: Oxford University Press.

Skiles, A. (2015) Against Grounding Necessitarianism. *Erkenntnis* 80, 717-51.

Stenwall, R. (2016) Against Truthmaker Necessitarianism. *Logique et Analyse* 59, 37-54.

Stenwall, R. (2017) Causal Grounds for Negative Truths. *Philosophical Studies* 174, 2973-2989.

Bradley, Compositionality and Davidson

Martin L. Jönsson

Once upon a time, I wrote my Bachelor's thesis on Bradley's regress – an argument purporting to undermine, ultimately, the idea that there can be complex wholes (cf. Bradley 1893) – a topic that Anna-Sofia Maurin has concerned herself with, on several occasions (Maurin 2002, 2010, 2012), to the considerable benefit of the rest of us. When I, a few years later, got admitted as a Ph.D.-student, I ended up writing my dissertation (Jönsson 2008) on a quite different topic, the principle of compositionality – the thesis that the meaning of a complex expression is determined by its parts and its mode of composition.

Although seemingly related, I never explored the connection between these two topics in writing, since my transition between the two was far from continuous. In my dissertation I acknowledged only a prerequisite for the transition: the kind encouragement given to me in relation to my work on Bradley's regress, by, among others, Anna-Sofia. It seems fitting then, as a small token of my gratitude for her early encouragement, as well as later inspiring collegiality, for me to write something on the occasion of her 50th birthday about

the relation between Bradley's regress and the principle of compositionality.

1. Davidson's regress argument

A suitably narrow focus of attention, given the space allotted to me by the conscientious editors and my nexus of interest, is a brief, Bradleyan, argument in Donald Davidson's paper "Truth and Meaning". In what follows I will briefly relate three interpretations of this argument to compositionality arguments roughly in the same line of work. The mere juxtaposition will have to be enough; space prohibits assessing the strength of these interpretations (and thus Bradley's regress) in any detail.

In the paper, Davidson is concerned with how the meaning of a complex expression depends upon the meanings of its parts. The second paragraph of the paper reads as follows:

> One proposal is to begin by assigning some entity as meaning to each word (or other significant syntactical feature) of the sentence; thus we might assign Theaetetus to 'Theaetetus' and the property of flying to 'flies' in the sentence 'Theaetetus flies'. The problem then arises how the meaning of the sentence is generated from these meanings. Viewing concatenation as a significant piece of syntax, we may assign to it the relation of participating in or instantiating; however, it is obvious that we have here the start of an infinite regress. (Davidson 1967: 304)

Davidson here discusses a simple semantic theory *S* according to which the meanings of expressions and syntactic structures alike, are *entities* of some sort, where 'entity' is understood broadly enough to

include particulars (persons), properties, and relations. And he dismisses this theory, due to it generating a regress.

How is the regress generated? This is not obvious.

According to Davidson, the starting point of the regress is when *S* identifies the meaning of concatenation with the instantiation relation as a way to explain *DPC*, *Davidson's Primary Concern* (in 'Truth and Meaning'): how the meaning of a sentence is generated from the meanings of the parts of that sentence.

It is clear to see that to thusly identify the meaning of concatenation is not sufficient in itself to account for the meaning of 'Theaetetus flies'. In particular, the meaning of 'Theaetetus flies' cannot be identified with the meaning of concatenation, since 'Theaetetus flies' and 'Theaetetus weeps', and all other pairs of concatenated words, would then have the same meaning.

But why not say that the meaning of the sentence is *the complex of flying being related by instantiation to Theaetetus*? Call this version of *S*, S^{x}. This would avoid the former problem – 'Theaetetus flies' and 'Theaetetus weeps' would now have different meanings (since they involve the instantiation of different properties) – and we would have a straightforward explanation of *DPC*; the meanings of the parts of the sentence jointly build up the meaning of the sentence.

Does this proposal lead to regress? How?

1.1 Interpretation 1: consistency demands infinite complexity

William Lycan (2019: 4) discusses what seems to be Davidson's argument in some detail, and gives the following reason to think that something like S^{x} leads to a regress.

Since S^{x} identifies the meaning of concatenation with instantiation, and use instantiation to build sentential meanings, 'Theaetetus flies' really means the same as 'Theaetetus instantiates flying'. But

then consistency demands that the concatenation of 'Theaetetus' and 'instantiates flying' in 'Theaetetus instantiates flying' be assigned instantiation, or a similar entity, as its meaning. This entity would be part of the meaning of 'Theaetetus instantiates flying' and thus 'Theaetetus flies' since the two sentences are synonymous. But this means that 'Theaetetus instantiates flying' really means the same as 'Theaetetus bears the instantiation-relation to flying' or something similar, and so on. For every added meaning of a concatenation relation, we obtain a new expression whose explicit form contains a new concatenation relation, which calls for another meaning addition, and so on. It thus follows from S^{x} that the meaning of every sentence is infinitely complex. And on the assumption that there are no entities of infinite complexity, S^{x} should be rejected.

For reasons given below, it is unlikely that this is what Davidson had in mind, but the interpretation is nonetheless interesting. In particular, it is noteworthy that what is driving the regress is not any inability of the denoted complex itself to combine, but a requirement for consistency in how S^{x} assigns meanings. This means that the regress (as I have interpreted it), despite Lycan's explicit affirmation to the contrary, is not very closely related to Bradley's regress (which is not framed in terms of the meaning assignments of semantic theories). However, this actually makes the argument stronger, in one sense at least, since good solutions to Bradley's argument (e.g. Maurin 2012) are not applicable to it.

1.2 Interpretation 2: relations do not relate

There is also a more straightforwardly Bradleyan reason to think that S^{x} leads to regress along the following lines: the entity of flying being related by instantiation to Theaetetus, involves a relation relating a person and a property, but in order to relate, all relations need further relations to relate their relata to themselves, and these cannot

relate without further relations, and so on.[1] So a relation cannot really relate anything, and there are no objects whose existence presuppose that anything is being related. So S^x should be rejected, since it presupposes that there are such objects.

This is not convincing as it stands, even if we grant that relations cannot relate. From what we have said so far, relations actually relating anything is not an essential part of S^x. Consider S^s, according to which the meaning of the sentence 'Theaetetus flies' is the mereological sum (or set, or ordered sequence, depending on one's preference) of Theaetetus, flying and instantiation. S^s and S^x are equally attractive from the perspective of distinguishing the meaning of 'Theaetetus flies' from the meaning of 'Theaetetus weeps' and from their account of DPC. Yet, S^s does not presuppose that there are *relating* relations.

It might be objected though, that if it is actually impossible for relations to relate, there are no relations. S^s would then be no better off than S^s since both theories presuppose the existence of things that do not exist. Still, a theory which is like S^s, call it S^e, except that it does not assign concatenation a relation as its meaning, but some other entity, would still have the two aforementioned benefits of S^s and S^x without presupposing that relations relate or even exist.

However, the disunity of the meanings of S^e might be a cause for concern.

[1] This is regrettably crude, but will have to do for present purposes. See Maurin (2012) for a better version (in terms of the inability of a unifier to account for contingent unity within a realist framework) of a Bradleyan argument. See also Matti Eklund's (2019) 'constitution regress' which is very similar to Maurin's rendition.

1.3 Interpretation 3: entities cannot unite propositions

What Davidson (1967) was really after – as indicated by Davidson's 2005 extended discussion of the issue – was that a theory like *S* (and S^x, S^s and S^e) fails to account for the unity of the proposition expressed by an utterance of 'Theaetetus flies'. What generates the regress is that *S* only *assigns another entity* to the concatenation relation as a way to supplement the meanings – which are also entities – of 'Theaetetus' and 'flies'. But no mere collection of entities is going to account for how an utterance of a sentence can express anything, or how it can be true or false. New entities can be piled on top of the previous ones without end without ever delivering the desired account.

This seems to be a very different argument from the Bradleyan argument (cf. Eklund 2019), although Davidson (2005: 105-106) seems to conflate the two. The desired kind of unity might obtain even if the conclusion of the Bradleyan argument is accepted, or perhaps even more plausibly, might fail to obtain even if this conclusion is rejected. What is needed here, according to Davidson, is not the repudiation of metaphysical unbelievables but the incorporation of a special semantic role for predicates into one's semantic theory.

2. Davidson's regress argument as a compositionality argument

In the quoted paragraph, Davidson is engaged with rejecting a simple semantic theory, based on its inability to explain how the meaning of a sentence is generated from the meanings of its parts. The argument is thus an instance of the more general strategy whereby semantic theories are evaluated by trying to establish whether or not

they are compatible with the principle of compositionality. This strategy is perhaps most famously associated with Jerry Fodor and Ernest Lepore (1991, 1996, 2002), who use it to show, among other things, that meanings (/concepts) cannot be prototypes – since the prototype for a complex expression, e.g. 'pet fish', is not derivable from the prototypes corresponding to 'pet' and 'fish'. The strategy can also be seen at work in Frege's (1892) famous rejection of the idea that meaning is reference, when he observes that sentential meanings can differ (e.g. 'The Morningstar is the Evening star' and 'The Morningstar is the Morningstar' are not synonymous) even if there is lexical referential identity. And in Osherson and Smith's (1982) ingenious argument against there being composition rules for fuzzy sets (e.g. rules that can calculate the typicality scores for 'round block' and 'round ball' from the typicality scores for 'round', 'block', and 'ball'). Although the general strategy has been criticized (cf. Jönsson 2008; Szabo 2012) it is interesting to see how the regress arguments described above, compares to compositionality arguments like the aforementioned.

First, it can be noted that the two first interpretations (1.1 and 1.2) reject semantic theories based on the kind of entities that they presuppose exist; infinitely complex objects, and complex objects, respectively. The compositionality arguments rest instead on assumptions about how certain entities – e.g. prototypes, referents or fuzzy sets – distribute over expressions, and try to demonstrate that the entities associated with complex expressions cannot be obtained from the corresponding entities associated with the parts of these expressions. The regress arguments are at an advantage in this regard since the compositionality arguments tend to presuppose specific, sometimes questionable, compositionality formulations that play no role in the regress arguments (cf. Jönsson 2008: Chapter 6). The second interpretation does rest on a very questionable metaphysical

assumption (see Maurin 2012) though, so in total it is not much better off.

Second, the first regress argument depends on a methodological consistency requirement, which one can opt out of without abandoning the core idea of S^x, e.g. by maintaining that not all concatenation relations are meaningful, or that we should accept that an infinite number of concatenations are meaningful, but that all of these concatenations have exactly the same meaning, so the meaning of the sentence is not infinitely complex after all – there is just an infinitude of concatenation relations that all mean the same thing. The latter rebuttal is similar to Horwich's (2005) suggestion that we can resist compositionality arguments by denying a uniformity assumption according to which the meanings of words and complex expressions are similarly constituted. Once dropped, it seems that compositionality loses a lot of its bite (but see Jönsson 2008: 205). Dropping the consistency requirement in S^x seems even less concessive which is a weakness of the first interpretation of the argument as compared to the compositionality arguments.

Third, the Bradleyan argument somewhat overshoots its target when seen from the perspective of a compositionality argument. Since it establishes that there cannot be *any* complex wholes, there cannot be any sentences, and there is thus nothing that *S* fails to account for. The Bradleyan argument makes *DPC* redundant, and compositionality vacuously true.

Fourth, from what has been remarked so far, only the third regress argument shows any promise as a compositionality argument. However, if we reject the Bradleyan argument (e.g. along the lines described by Maurin 2012), it is unclear why something like S^x should fail to account for why an utterance of a sentence can have a truth-value or express something (or relatedly, why e.g. Montague's sup-

posed insight that 'function-argument structure [is] the basic semantic glue by which meanings are combined (Partee 1997:27) is undermined). When there is unity in the entities denoted by sentences, what else remains to be accounted for?

Although this must be understood as preliminary conclusion, due to the brevity of the investigation, it seems that the regress arguments do not give us any information about how to best think about meaning. Since I have argued elsewhere (Jönsson 2008) that neither does the compositionality arguments, these and the regress arguments are on a par in this respect.

References

Bradley, F. (1893) *Appearance and Reality*, Oxford: Oxford University Press.

Davidson, D. (1967) Truth and Meaning, *Synthese* 17(1), 304-323.

Davidson, D. (2005) *Truth and Predication*, London, England: Harvard University Press.

Eklund, M. (2019) Regress, Unity, Facts, and Propositions, *Synthese* 196(4), 1225-1247.

Fodor, J. and Lepore, E. (1991) Why Meaning (Probably) Isn't Conceptual Role, *Mind and Language* 6(4), 329-343.

Fodor, J. and Lepore, E. (1996) The Red Herring and The Pet Fish: Why Concepts Still Can't be Prototypes, *Cognition* 58(2), 243-276.

Fodor, J. and Lepore, E. (2002) *The Compositionality Papers*, Oxford: Clarenden Press.

Frege, G. (1892) Über Sinn and Bedeutung, *Zeitschrift für Philosophische Kritik* 100, 25-50.

Horwich, P. (2005) *Reflections on Meaning* Oxford: Oxford University Press.

Jönsson, M. (2008) *On Compositionality: Doubts about the Structural Path to Meaning*, Ph.d.-thesis, Lund University.

Lycan, W. (2019) *The Philosophy of Language.* 3rd edition, New York: Routledge.

Maurin, A-S. (2002) *If Tropes*, Dordrecht: Kluwer Academic Publishers.

Maurin, A-S. (2010) Trope Theory and the Bradley Regress, *Synthese* 175(3), 311-326.

Maurin, A-S. (2012) Bradley's Regress. *Philosophy Compass* 7(11), 794-807.

Osherson, D. and Smith, E. (1982) Gradedness and Conceptual Combination, *Cognition* 12, 299-322.

Partee, B. (1997) Montague Grammar. In ter Meulen, A. and van Benthem, J. (eds.) *Handbook of Logic & Language* (pp. 5-91) Cambridge, Massachusetts: The Mit Press.

Szabò, Z. (2012) The Case for Compositionality. In Hinzen, W., Machery, E., and Werning, M. (eds.) *The Oxford Handbook of Compositionality* (pp. 64-80) Oxford: Oxford University Press.

Too Many Omissions, Too Much Causation?

Björn Petersson

1

Johan did not feed your guppies when you were away, as he had promised. It does not seem strange to say that Johan's omission caused their death. But if Johan's not feeding the guppies caused their death, then arguably so did Melania Trump's not feeding them, Zlatan Ibrahimovic's not feeding them, and so on. So, if negative acts are causes, there is "far more causation in the world than common sense says there is" (McGrath 2005: 125).[1]

To begin with, I do not find that implication too disturbing. It is part of common sense that on a scientific view of the world, there

[1] The other standard worry about omissions as causes, which I do not discuss here, is that "[a]n omission is not like an act, a real event" (19th century thinker quoted in Bennett 1994: 87). Causation is supposed to be a relation between occurring events, or at the very least a relation between real entities. Judith Thomson suggests that negative acts are *states of affairs*, and that causation can be a relation on states of affairs (Thomson 2003: 86). Jonathan Bennett thinks that not doings are *facts,* and that causation is ultimately a relation on facts (Bennett 1994: 87). Bruce Vermazen argues, more plausibly I think, that negative acts are behavioural *events* under negative descriptions (Vermazen 1985: 95).

are far more causes than the ones we can refer to in the explanations and predictions we make in daily life, and even that there are more causal influences than possibly can be included in scientific explanations and predictions. There is no inconsistency in claiming that acknowledging limitations of common sense, especially in this area, is part of common sense itself.

Still, there is an intuitive difference between Johan's and Melania's inactions. Some attempt to show that the difference is metaphysical; Unlike Johan's omission, Melania's negative act is *not real* (Bennett 1994; Vermazen 1985) or had *no effects* with respect to your fish (Thomson 2003; McGrath 2005). Others, like James Woodward, think that *pragmatic* considerations explain our reluctance to bring up Melania's negative act (Woodward 2003: 86-91, 226-233). I will argue that the first type of strategy fails: if Johan's omission caused your fish's death, then so did Melania's. The pragmatic approach is more plausible. We need pragmatics anyhow to explain how we bring up positively described acts in terms of causation, as well as how we talk about positive as well as negative acts in terms of counterfactual dependence.

2

According to Bruce Vermazen, "unintentional omissions, failures, and neglectings-to-do do not exist" (1985: 93). His main argument is an example in which a watchman unintentionally fails to lock a door twice: first because he is absorbed in reading: then because he is asleep. The watchman's behaviour while reading is intentional under the description 'reading a book'. The watchman's behaviour while asleep is not intentional under any description, so it is not an

act. But since the watchman's behaviours with respect to the door are exactly similar in both cases, if his not locking is not an act in the second case, it was not an act in the first case either.

However, the watchman's sleeping and his reading are occurring behavioural events. Nothing stops us from referring to both events in negative terms, i.e. as his not locking the door. How could these events be brought out of existence by the mere fact that one of them is not intentional under any description? So, I think that the mere fact that Melania's not feeding your fish was unintentional cannot not make it less real than Johan's not feeding them, regardless of whether we call her behaviour an 'act'. [2]

If unintentional negative acts exist, "the agent will be doing far too many negative acts", according to Vermazen (1985: 96). But note that on the Davidsonian notion of acting that Vermazen employs, an agent performs as many positive acts as there are true positive descriptions of her behaviour, if there is at least one true description under which that behaviour is intentional. At the moment I am touching a keyboard, writing a sentence in English, compiling an open-ended list, attempting to honour Anna-Sofia Maurin on her 50th birthday, and so on. That things are too many to be listed does not mean that they are too many to be real.

Jonathan Bennett's attempt to narrow down the class of negative acts proceeds from the assumption that a negative proposition about someone's behaviour can

[2] In some places Vermazen says more cautiously that unintentional negative acts "should at least not be counted as 'acts'" (1985: 93). There is not much at stake in accepting that claim. 'Negative act' is a technical term and we could invent new labels for whatever it is we refer to when we speak about unintentional not doings. That terminological manoeuvre will not help us to cut down the number of them, though.

> be represented by a region that covers nearly the whole space of possibilities for him at that time. That is, a proposition and its complement are positive and negative respectively if they divide the Agent's behaviour space /…/ extremely unevenly. (Bennett 1994: 90)

An agent's not doing something that lies outside her space of possibilities will not count as a negative act. 'Melania did not feed your guppies at t' covers *all* possibilities within her behaviour space at t. A consequence is that this phrase would not express a negative proposition about Melania's conduct — it would not state a negative fact in Bennett's sense. That statement and its positive complement do not divide her behaviour space at all. Since Bennett regards causation as a relation between facts, this is a causally relevant difference between Melania's not doing and Johan's.

In most contexts, the pragmatics of speech dictates that "she did not ϕ" is uttered on the background assumption the she could have ϕ-ed. But that does not make "she did not ϕ" false or meaningless in case we find out that ϕ-ing actually was not a possibility for her. As far as I can see, "she did not levitate" refers to the same kind of conduct regardless of whether levitation was in her behaviour space or not. This way of reasoning will not help us distinguish between Melania's and Johan's negative acts.

Another line of thought is that there is something about the concept of a *cause* that makes a difference in kind between Melania's and Johan's negative acts (Thomson 2003; McGrath 2005). Thomson draws the distinction in terms of being "at fault". Johan's negative act but not Melania's caused the guppies' death in virtue of the principle "If x is at fault for y, then x causes y." The idea is not merely that x being at fault for y *presupposes* that x causally contributed to y, but that "there need not be anything that independently makes it

true that x caused the outcome: it can be that x's causing the outcome follows only from x's being at fault for it" (2003: 100). According to Sarah McGrath, Johan's feeding your fish, unlike Melania's, is according to the standards of promising a normal "would-be preventer" of their death, so his omission caused it. McGrath adds that "we should not be surprised to see normativity implicated elsewhere in causation" (2005: 146). McGrath and Thomson employ this strategy in non-moral contexts as well. The valve's not closing can be at fault for the flooding of the basement, and some gland's not secreting a certain enzyme can be at fault for a person's falling ill, etc.

Bennett presents an obvious objection to one way of understanding the difference between Johan and Melania:

> [Some philosophers] count behaviour as an 'act of omission' only if it involves the agent *notably* not ϕing — usually because he ought to have ϕed. That rescues us from having to credit people with performing countless acts of omissions all at once; but it makes 'act of omission' worthless in our present inquiry into the *bases* for our moral judgments on behaviour. (Bennett 1994: 88)

The obvious objection to counting a negative act as a 'cause' only if it breaks norms is that it makes the concept of a cause worthless in normative reasoning about what the standards should be, and what counts as a fault.

This applies to non-moral cases as well. Suppose we believe that the valve should close when the pressure becomes high. So, the valve's not closing is at fault for the flooding of the cellar. Now the plumber tells us that the valve's not closing is meant to prevent leaks that would cause greater damage. Then, what he tells us according to the normative view of causation is not only that our standards

were wrong, but that the valve's not closing did not cause the flooding of the cellar. Moreover, the correct standards for the valve are based upon expectations concerning the effects of the valve's not closing. That sort of justification would collapse if the standards determined what should be counted as effects to begin with.

3

Johan's feeding your fish was in his behaviour space, his neglection was intentional, and he ought to have fed them. The absence of these elements would not, I have claimed, make a causal difference. Admittedly, it sounds weirder to say that Melania Trump's inaction caused their death than that Johan's did so. However, it also sounds weird to state the uncontroversial fact that if Melania had fed the fish, they would not have died. Regardless of what we believe about the conceptual relation between causation and counterfactual dependence, this indicates that weirdness has little evidentiary value in this context. We must appeal to conversational and moral considerations to explain why certain counterfactual dependencies, and also why certain positively described causes, are rarely brought up. So, if there is causation by omission, we already have a strategy that can be employed to explain why it sounds weirder to mention some negative acts than others.[3]

[3] David Lewis makes a similar point concerning the weirdness of saying "his birth caused his death" (Lewis 2000: 196). In a related manner, Woodward argues that even when counterfactuals concerning nonserious possibilities are straightforwardly *true*, "to the extent that the possibilities that figure in them are nonserious, they do not guide our causal judgments" and his notion of a 'serious possibility' is broadly pragmatic (2003: 90, 227).

How would this affect practical deliberation? Melania's deliberation is a matter of deciding between options that are open to her, so she should not worry about your guppies. What about moral responsibility? Standard conditions for responsibility would require that the inaction was intentional, that the agent could have done otherwise, that the effect was foreseeable, etc. Melania would not fulfil these conditions.

To conclude, if we allow Johan's omission to be a cause of your guppies' death, I suspect that we cannot avoid allowing Melania's negative act to be a cause of their death. This would not lead to practical difficulties that differ in any significant way from the ones we face when attempting to justify our pick of causes among actions positively described, or our pick of counterfactual dependencies. In both cases, we will have to appeal to pragmatic and moral considerations. Letting conversational and moral standards determine the *truth* of causal statements would, however, have implications that square more blatantly against common sense than the entailments that this strategy is supposed to save us from.

Acknowledgements

A long version of this text was presented at the Higher Seminar in Practical Philosophy in Lund, and I am grateful for comments from participants. Thanks, also, for comments from Robin Stenwall and Tobias Hansson Wahlberg.

References

Bennett, J. (1994) *The Act Itself*, Oxford: Clarendon Press.

Lewis, D. (2000) Causation as Influence, *Journal of Philosophy* XCVII.

McGrath, S. (2005) Causation by Omission: A Dilemma, *Philosophical Studies* 123, 125-148.

Thomson, J. (2003) Causation: Omissions, *Philosophy and Phenomenological Research* 66(1), 81-103.

Vermazen, B. (1985) Negative Acts, in Vermazen & Hintikka (eds.) *Essays on Davidson. Actions and Events*, Oxford: Oxford U P.

Woodward, J. (2003) *Making Things Happen: A Theory of Causal Explanation,* Oxford: Oxford U P.

The Metaphysician as Legal Expert

Lena Wahlberg

1. Introduction: philosophers as expert witnesses

In McLean v. Arkansas Board of Education 529 F. Supp. 1255 (E.D. Ark 1982), professor of history and philosophy Michael Ruse was called as expert witness on behalf of the plaintiffs. The case concerned the constitutionality of Act 590, which required Arkansas public schools to "give balanced treatment to creation-science and to evolution-science" (section 1), and Ruse was asked to describe to the court his "understanding of what science is today". Influenced by Ruse's testimony, the court found that creation science was "simply not science", and that Act 590 violated the constitution. From a philosophical point of view, the case is interesting not only because of the science-theoretical problems it raised (see e.g. Laudan 1982) but perhaps even more because it demonstrates that the expertise of a philosopher of science can make a real difference in the courtroom.

However, it is extremely uncommon that a philosopher of science is called to serve as expert witness in court. It is probably even less common that a court or one of the parties appoints a *metaphysician*

as expert witness. But what if the *status quo* is a result of a fatal neglect of the legal relevance of philosophical insights in general and metaphysical insights in particular? Inspired by McLean v. Arkansas, I will devote this paper to a discussion of the metaphysician's chances of having a lucrative career as an expert witness. I will do this by examining the virtues of an argument from metaphysics that is sometimes made by legal scholars in their attempts to identify valid law.

2. The argument from metaphysics (AFM)

The average practising lawyer is not known for paying much attention to metaphysical debates. On the contrary, "law and metaphysics" has been described as a "subject of limited appeal, which gains attention only during periods of defensiveness and self-doubt within the legal profession" (Yablon 1987: 616). This is so despite the fact that the most fundamental question for jurisprudence – *What is law?* – seems to be a metaphysical one, with immediate bearing on questions regarding what is a valid legal argument (Simmonds 2018). Nevertheless, practising lawyers are normally more concerned with whether their arguments will convince the court, than with whether they are consistent with the true conception of law.

Legal *scholars*, however, make metaphysical arguments now and then, some of which are meant to have practical implications. In this paper, I will discuss arguments from metaphysics that have the following schematic form (where *c* is a particular conceptualisation, for example the current interpretation of a legal statute or term):

Premise 1	*if c does not reflect the true nature of reality, then c should not be used in law.*
Premise 2	*c does not reflect the true nature of reality.*
Conclusion	*c should not be used in law.*

Below, I will refer to arguments from metaphysics that have this schematic form as AFMs. If an AFM's first premise (below referred to as the "major premise") is convincing, it would be reasonable for legal scholars and practitioners alike to ask metaphysicians for help with assessing the truth of the argument's second premise (below referred to as the "minor premise"), which in turn would pave the way for the metaphysician as legal expert.

3. An example from Swedish criminal law

In this section, I will present an example of an AFM that I recently came across in the Swedish criminal law literature. Before presenting the argument, I will give a brief overview of the legal context in which the argument was made.

In Swedish law, it is a notoriously difficult question to determine under what conditions a perpetrator's erroneous belief shall free the perpetrator from criminal responsibility. In the legal literature, this question has been partly answered by the aid of a distinction between mistakes about *the law* and mistakes about *factual circumstances*. According to the principle *ignorantia juris non excusat*, mis-

takes about the law do normally *not* free the perpetrator from criminal responsibility.[1] For example, a perpetrator who beats up her neighbour would be convicted for assault even if she did not know that what she did was a crime. If, in contrast, the perpetrator was mistaken about the factual circumstances – she thought that she was beating a mat and not her neighbour – she would not be convicted for assault, since she would lack intent to assault.

As so often, however, the legal situation is more complicated than it first appears. It is generally held that mistakes about *certain parts* of the law shall free the perpetrator from liability, and hence be treated as mistakes about factual circumstances. For example, it is often claimed that a perpetrator cannot be held responsible for the crime "unlawful deprivation of liberty" (Penal Code (1962:700), chapter 4, section 2) if she erroneously believed that the law entitled her to deprive the victim of liberty (see e.g. Asp, Ulväng and Jareborg 2013; Martinsson 2016; cf. my 2019). Legal scholars have tried to explain what, more precisely, marks the difference between relevant and irrelevant mistakes about the law. In older Swedish criminal law literature, it was common to distinguish between "legal mistakes *improper*" [oegentlig rättsvillfarelse] (which free the perpetrator from responsibility for intentional offence, and hence are treated like factual mistakes) and "legal mistakes *proper*" [egentlig rättsvillfarelse] (which do not free the perpetrator from criminal responsibility) to explain the difference. However, the borderline between legal mistakes improper and proper turned out to be hard to define, and moreover did not quite match the perceived borderline between relevant and irrelevant mistakes.

[1] The Swedish Penal Code makes an exception for mistakes that are due to an error in the proclamation of the criminal provision or are manifestly excusable for some other reason (Penal Code (1962:700) chapter 24, section 9).

The distinction between factual mistakes, legal mistakes proper and legal mistakes improper received its coup de grace from Nils Jareborg, a Swedish professor of criminal law. In his text book *Allmän kriminalrätt* [General criminal law] (2001: 355), Jareborg argues that the distinction should be abandoned. One of the reasons he presents for not using the distinction is the following:

> [T]he distinction between "facts" and "what is prescribed by law" is odd. Different kinds of "institutional facts", namely facts constituted by rules, are "real" facts too. It is as much a fact that someone is a prime minister as him being close-shaven.

Although somewhat elliptical, Jareborg's argument can be interpreted as an AFM, with a minor premise saying that the distinction in question does not reflect the true nature of reality, a conclusion saying that the distinction should not be used in law, and an implicit major premise saying that if the distinction does not reflect the true nature of reality, then it should not be used in law.[2] In the next section, I will discuss whether Jareborgs' argument and other AFMs can be sustained. I will not examine the truth of Jareborg's minor premise – I know of several metaphysicians who are much more suited than I to do this. Instead, I will discuss what it takes for the major premise in these arguments to be acceptable, and for metaphysical insights to be relevant in the first place.

[2] In other texts, Jareborg has claimed that there is not one single true nature of reality (see e.g. Jareborg 2004). However, I fail to see how his argument from 2001 could be a meaningful criticism of the distinction if not interpreted as an AFM. In any case, I will make this interpretation for the sake of the argument of the current paper.

4. When is an AFM's major premise acceptable?

I can think of two kinds of argument that could back an AFM's major premise that a particular conceptualisation *c* should not be used in law, if it does not reflect the true nature of reality:

Backing argument A:

Premise A1	*if c does not reflect the true nature of reality, then c will not work in practice.*
Premise A2	*if c will not work in practice, then c should not be used in law.*
Conclusion	*if c does not reflect the true nature of reality, then c should not be used in law.*

Backing argument B:

Premise B1	*if c does not reflect the true nature of reality, then c is legally irrelevant.*
Premise B2	*if c is legally irrelevant, then c should not be used in law.*
Conclusion	*if c does not reflect the true nature of reality, then c should not be used in law.*

I will now examine the soundness of arguments A and B.

Backing argument A.

Is it the case that if a conceptualisation does not reflect the true nature of reality, then it will not work in practice? Not necessarily. As illustrated by, say, the continued use of Newton's laws of motion, a conceptualisation can work sufficiently well under some conditions and for some purposes, despite not being a perfect representation of reality. The fact that a conceptualisation does not reflect the true

nature of reality can hence at most increase the risk that the conceptualisation will not work in practice, but few would say that this risk by itself means that the conceptualisation should not be used in law. Compare this with the pragmatic observation that that the conceptualisation as a matter of fact does not work satisfactorily, which would provide a very strong reason for not using it in law, but at the same time would make metaphysical information about how the conceptualisation relates to the true nature of reality superfluous.

Backing argument B.

Is it the case that if a conceptualisation does not reflect the true nature of reality, then it is legally irrelevant? It could certainly be argued that *some* conceptualisations need to reflect the true nature of things in order to be legally relevant. For example, Michael Moore has argued that "cause" in the law "refers to a natural relation that holds between states and events. Because moral responsibility is tied to such a natural relation, and because the law is tied to morality, the law also is tied to this natural relation." (Moore 2009: 5). If we agree with this reasoning, we could perhaps discard conceptualisations of causation that do not reflect the true nature of causation as legally irrelevant. But this does not mean that *all* conceptualisations that do not reflect the true nature of reality would have to be discarded as legally irrelevant. For premise B1 and argument B to be convincing with respect to a particular conceptualisation, *it must be explained why true nature plays a crucial role for the legal relevance of the conceptualisation in question.* Moore's argument about there being a link between law, morality and true causation is an example of an explanation which if accepted [3] would support premise B1 with respect to the conceptualisation of causation in the law. However, I

[3] I regret to say that I personally am not convinced by Moore's argument.

have great difficulties seeing the merits of premise B1 and of argument B in the absence of a legally convincing explanation for why a particular something's true nature matters for legal relevance.

5. Concluding remarks: the metaphysician as legal expert!

If the reasoning in this paper is correct, an AFM's major premise is convincing only when accompanied by an explanation for why something's true nature is legally relevant. But once an explanation to this effect and the AFM's major premise are accepted, metaphysical insights become highly relevant to assess the truth of the AFM's minor premise. And when this happens, we have very good reasons to invite the lawyer's kindred spirit, the metaphysician, as legal expert.

Acknowledgement

This work is supported by a grant from the Swedish Foundation for Humanities and Social Sciences, grant number M14-0138:1.

References

Act 590 of the Acts of Arkansas of 1981.

Asp, P., Ulväng, M. & Jareborg, N. (2013). *Kriminalrättens grunder*. Uppsala: Iustus.

Jareborg, N. (2001). *Allmän kriminalrätt*. Uppsala: Iustus.

Jareborg, N. (2004). Rättsdogmatik som vetenskap. *Svensk Juristtidning*, 1-10.

McLean v. Arkansas Board of Education 529 F. Supp. 1255 (E.D. Ark 1982).

Laudan, L. (1982). Commentary: Science at the Bar – Causes for Concern. *Science, Technology and Human Values* 7 (41), 16-19.

Martinsson, D. (2016). *Om straffrättsvillfarelse*. Stockholm: Wolters Kluwer.

Moore, M. (2009). *Causation and Responsibility: An Essay in Law, Morals, and Metaphysics*. Oxford: Oxford University Press.

Simmonds, N. (2018). *Central Issues in Jurisprudence: Justice, Law and Rights*. London: Thomas Reuters.

Swedish Penal Code (1962:700). Brottsbalken (1962:700).

Wahlberg, L. (2019). Kommentar till Högsta domstolens avgörande den 18 december 2018 i mål B 5050-17 (NJA 2018 s. 1051). Forthcoming in *Nordisk socialrättslig tidskrift*.

Yablon, C. M. (1987). Law and Metaphysics. *Yale Law Journal* 96 (3), 613-636.

Are there any Institutional Facts?

Tobias Hansson Wahlberg

Following John Searle, social ontologists often distinguish between brute and institutional facts. The fact that there is snow on Mount Everest is a brute fact; the fact that Donald Trump is president of the United States is an institutional fact. Brute facts exist independently of any institutions; institutional facts require institutions to exist (Searle 1995: 2, 27; 2010: 10).

There is an ambiguity in the notion of fact that needs to be resolved here. Are Searle and his followers talking about facts as worldly states of affairs (objects having properties, objects standing in relations to one another) or facts as truths, i.e. facts as true statements or propositions (abstract entities)? Searle's disciples are often silent on this issue. I think it is clear, however, that Searle himself thinks of facts – institutional ones included – as worldly states of affairs. He says that "facts function causally in a way that true statements do not" (1995: 206) and that "the whole point of having the notion of 'fact' is to have a notion for that which stands outside the statement but which makes it true, or in virtue of which it is true, if it is true" – "[facts] are conditions in the world that satisfy the truth conditions expressed by statements" (1995: 211). He summarizes his approach to social ontology: "one […] method in philosophy is to analyze the structure of the facts that make our statements true.

In earlier chapters I have attempted to do that with the structure of […] institutional facts" (1995: 221).

But can there be worldly states of affairs that require institutions for their existence? How are such facts brought into being? Searle maintains that institutional facts are created by collectively accepted "Status Function Declarations", typically of the form "We make it the case by Declaration that object X now has the status function Y in C" (2010: 99).[1] Institutions are systems of Status Function Declarations (2010: 10, 13). Such declarations "change the world by declaring that a state of affairs exist [X's being Y] and thus bringing that state of affairs into existence" (2010: 12).[2] The worldly state of affairs is created simultaneously with the performance of the declaration – hence the use of the present tense and indexicals like "now" in declarations (Searle 1989: 556-557).

How could a declaration synchronically bring a worldly state of affairs into being? By way of causation (as defended e.g. by Elder-Vass 2012: 62-65)? That option seems to be ruled out by the special theory of relativity, according to which causal processes always propagate at a finite velocity (precluding instantaneous causation). By way of grounding (as defended by Schaffer forthcoming)? But syn-

[1] A status function, according to Searle, is a function (i.e., "a cause that serves a purpose" (Searle 2010: 59)) with two special but interrelated features: "First, […] they require collective intentionality, both for their initial creation and for their continued existence. And second, they are functions that a person or other entity has, not in virtue of physical structure […], but in virtue of collective imposition and recognition of a status." (Searle 2010: 59; see also 1995: 40-51; 2010: 7, 102-104)

[2] As Searle puts it in his (1975/1979: 16): "It is the defining characteristic of this class [of speech acts] that the successful performance of one of its members brings about the correspondence between the propositional content and reality, successful performance guarantees that the propositional content corresponds to the world."

chronic grounding in one frame of reference will, in relativistic Minkowski spacetime, involve backward grounding in a frame of reference moving at high velocity relative to (and towards) the reference frame in which the grounding relation, holding between the spatially separated relata, is synchronic.[3] Perhaps there are instances of backward grounding, but backward grounding of worldly institutional states of affairs is simply incredible. How could a person acquire a worldly property (or "status function") of being president, a convicted criminal or a promoted professor before the relevant declaration has occurred (in a certain reference frame)?

Disregard relativistic considerations: synchronic creation of worldly institutional states of affairs by declarations is problematic even within a Newtonian framework, where simultaneity is absolute and action at a distance is possible in principle. What is the exact *mechanism* (of causation or grounding or …) that generates the worldly institutional state of affairs, consisting of an object with an institutional property, located at a distance from the declaration? It is hard to see how there could be such a mechanism. The idea seems to involve magic. As Nikk Effingham puts it, discussing declarations: "I think it is strange that merely speaking and intoning certain phrases could cause anything to exist (except, of course, for the words and intonations themselves). […] The thought is that only wizards and warlocks can bring things into existence by merely uttering a few phrases." (Effingham 2009: 253) Indeed, Searle never explains how worldly institutional states of affairs could be brought into existence by declarations. He is content proclaiming that we simply can bring such states of affairs into being: "We ordinary humans do not have the ability to [successfully] perform supernatural

[3] This is illustrated in detail in my (ms.). Many of the issues mentioned briefly here are discussed in depth in that paper.

declarations [e.g., to create light by uttering "Let there be light!"], but we do have a quasi-magical power nonetheless of bringing about changes in the world through our utterances. (Searle 1989: 549) We can create boundaries, kings, and corporations by saying something equivalent to 'Let this be a boundary!' 'Let the oldest son be the king!' 'Let there be a corporation!'" (2010: 100)[4]

I think it is much more plausible to hold that nothing worldly is created by declarations (except for the words and intonations themselves): no worldly institutional property or object, and hence no worldly institutional state of affairs, is created by a "Status Function Declaration". But that is not to say that collectively accepted declarations do not "create" institutional facts understood as *true propositions* partly made true by declarations. The truth-makers for propositions such as <Donald Trump is President of the Unites States>, <N.N. is a convicted criminal> and <Anna-Sofia Maurin is a professor> do plausibly involve declarational utterances or inscribings (or more fundamentally, utterance acts (Searle 1969: 24), or Austinian phatic acts (Austin 1962: 95)), semantic rules, legal regulation (i.e. further declarational utterances and inscriptions), and people's attitudes – much the way characterized by Searle in his books and articles (Searle 1969; 1975/1979; 1989; 1995; 2010). Only, these truth-makers seem to be brute (as Searle himself seems to acknowledge at

[4] Admittedly, Searle says that institutional facts are "ontologically subjective" (1995: 8; 2010: 18), and thus it might be that he holds that they do not really (or objectively) exist (see also Searle 2010: 100, 120). But if institutional states of affairs do not really exist, how can they serve as truth-makers? How can they be causal? How can they be part of ontology?

various places, e.g. 1995: 12; 2010: 110-115). If this is correct, institutional truths[5] can be taken to be made true by *brute* truth-makers. We need not postulate institutional state of affairs as truth-makers for institutional truths (*pace* Searle 1995: 221).

To wrap up: are there any institutional facts? If understood as worldly institutional *states of affairs* (created by declarations), arguably no; if understood as institutional *truths* (partly made true by declarations), arguably yes.

Consequence: institutional facts cannot be relata of causal relations. *Qua* states of affairs, they are non-existent; *qua* true propositions, they are too abstract. *Mental representations* of institutional facts may however influence our reasoning, decision making and behaviour and thereby give rise to the complex patterns of behaviour and social interactions we see in modern societies. (These conclusions are supported further and elaborated in my forthcoming and ms.)

Acknowledgements

Thanks to Robin Stenwall and Lena Wahlberg for helpful comments on an earlier draft. The writing of this paper was financed by *Riksbankens Jubileumsfond* (grant number: P15-0744:1).

[5] Note that we may very well allow that such truths express so-called "abundant properties" (see my ms. for discussion).

References

Austin, J. L. (1962) *How to Do Things with Words*, Cambridge, MA: Harvard University Press.

Effingham, N. (2010) The Metaphysics of Groups, *Philosophical Studies* 149(2), 251-267.

Elder-Vass, D. (2012) *The Reality of Social Construction*, Cambridge: Cambridge University Press.

Hansson Wahlberg, T. (forthcoming) Causal Powers and Social Ontology, *Synthese*, doi.org/10.1007/s11229-018-1763-2

Hansson Wahlberg, T. (ms.) The Creation of Institutional Reality, Special Theory of Relativity, and Mere Cambridge Change.

Schaffer, J. (forthcoming) Anchoring as Grounding: On Epstein's The Ant Trap, *Philosophy and Phenomenological Research*.

Searle, J. R. (1969) *Speech Acts – An Essay in the Philosophy of Language*, Cambridge: Cambridge University press.

Searle, J. R. (1975/1979) A Taxonomy of Illocutionary Acts. In *Expression and Meaning – Studies in the Theory of Speech Acts* (pp. 1-29), Cambridge: Cambridge University Press.

Searle, J. R. (1989) How Performatives Work, *Linguistics and Philosophy* 12(5), 535-558.

Searle, J. R. (1995) *The Construction of Social Reality*, London: Penguin Books.

Searle, J. R. (2010) *Making the Social World: The Structure of Human Civilization*, Oxford: Oxford University Press.

Towards a Nominalist Understanding of Institutions

Johan Brännmark

A very common understanding of institutions is that they are rules of some kind. For instance, an institutional economist like North (1990: 3) suggests that '[i]nstitutions are the rules of the game in a society' and a social ontologist such as Gilbert (2018: 30) characterizes an institution as 'a system of rules that is a blueprint for human behavior.' In political theory, Rawls (1999: 47-48) takes the stance that an institution is 'a public system of rules which defines offices and positions with their right and duties, powers and immunities, and the like.' What this means is that if I hold a particular status, such as the right to perform a specific action, the fact that I, as a concrete and particular individual, hold this status is explained in terms of a certain rule being established in my community or society. Typically, that rule being established is then explained by it being collectively accepted in some relevant way (and where explicating the more exact relevant way is a main focus for many social ontologists). Ultimately, such collectives do of course consist in a number of concrete and particular individuals, but one notable feature of the standard account is that in order to explain particular institutional facts, an abstract entity is being postulated: a rule. This standard

model stands in contrast, however, to another possible type of account which instead understands institutions in terms of regularities, with Hume's (1978: 490) account of property being an early example, and Lewis' (1969) account of conventions being another. Contemporary versions of this kind of approach are however more common in economics than in social ontology.

At least given a broadly nominalist starting-point, the standard model is *prima facie* unattractive. While it starts with (A) particular individuals holding certain attitudes and ends with (B) particular individuals holding certain deontic statuses, it goes from (A) to (B) via an abstract and general type of entity, rules, whose ontological status is often highly unclear. In this paper, it will be argued that we can get from (A) to (B) without this detour into the abstract and the general. We can do without the level of rules in accounting for institutions or institutional facts. Of course, one might not generally think that we should take a broadly nominalist approach to metaphysical matters, and there is no room here for arguing that we should; but at the very least, given that many do find such an approach attractive, there seems to be reason to explore the *possibility* of a broadly nominalist approach to social ontology.

1. Why consider nominalism in social ontology?

While social ontology is arguably a species of metaphysics, the question of nominalism is rarely discussed there. One reason for this is probably that the type of entities studied in social ontology tend to be much more complex than the ones considered in debates over nominalism. For instance, one important form of nominalism is trope theory (see Maurin 2002). Tropes are simple, particular, and

abstract. This then means that the kinds of objects and properties that we encounter in daily life are arguably not directly characterizable in terms of tropes – at the very least, they are not simple, but have a variety of constitutive parts. Accordingly, when turning from fundamental metaphysics to a domain-specific metaphysics like social ontology, there is little reason to expect, even given that we believe that tropes are the ultimate building-blocks of reality, that the notion of tropes will be useful in identifying the building-blocks of social reality – the latter entities are bound to be highly complex complexes already to begin with. Whether we reject the existence of universals or not would accordingly seem to have little bearing on social ontology.

There is however another main understanding of nominalism as well (Rodriguez-Pereyra 2015), namely as involving the rejection of abstract objects, and where these are typically understood as non-spatiotemporal and causally inert objects. This sense of nominalism is at least at first sight more directly relevant, since the rules in question seem to be precisely such abstract objects. While a specific community, collectively accepting or recognizing certain rules, would certainly be spatiotemporally located, the rules themselves are not – rather they make up an institution that could be collectively accepted at different times and in different places. The rules do not *cause* me to hold a particular set of rights and duties – they play a constitutive role instead. Here it should however be noted that social ontologists do not tend to explicate the more exact ontological status that the relevant rules are supposed to have. They could merely be an explanatory device that we use, presumably for pragmatic reasons, to facilitate our understanding of concrete institutions, but where the idea is not that those rules are ultimately part of what makes it true that I hold certain rights and duties. Given that type of ap-

proach, perhaps social ontologists could even be quietists about abstract objects, leaving it open whether the rules which are appealed to in order to make sense of concrete institutions are ultimately more than pragmatically convenient explanatory devices.

However, at least given that one is drawn to a broadly nominalist approach, whether in terms of rejecting universals or rejecting abstract objects or both, this kind of quietism is arguably unsatisfactory. Even if fundamental metaphysics and social ontology are in many ways different enterprises, they are still both parts of a complete account of reality, and it seems reasonable that we would prefer such a complete account to exhibit what might be called *meta-theoretical congruence*. For instance, unless there is any positive specific reason for why the list of theoretical virtues will differ between different species of metaphysics, the working assumption would be that it is the same list. Additionally, if there are certain type of explanatory models that we think make sense in one case, they should *ceteris paribus* make sense more generally as well – indeed, the extent to which a certain explanatory model makes sense in one case should partly depend on the extent to which it makes sense in other cases as well. One thing that characterizes a broadly nominalist approach is a commitment to what might called *the primacy of the particular* in how one seeks to explain things, i.e., even if we speak in abstract and general terms about certain matters, or sometimes understand particular events in term of our knowledge about certain regularities, the direction of explanation (in the metaphysical sense) ultimately runs from the level of the particular. Unless there are specific reasons for not doing so, the default type of explanation would be of this kind, not just in fundamental metaphysics, but also in areas like social ontology. To the extent that we can achieve such meta-theoret-

ical congruence, it will arguably have bearing on fundamental metaphysics as well, further strengthening the case for pursuing that direction of explanation there.

2. Doing without rules

A main reason for nominalist approaches tends to be Ockhamite, at least if we are already working with concrete objects: why postulate abstract objects unless we really have to? In the case of the standard model of institutions, it seems clear that the ontology in question already includes concrete objects, namely human individuals like you and me. We presumably relate to each other in a variety of ways and the question, then, is if we have to postulate certain abstract objects, rules, in order to make sense of the statuses we hold. Although there will not be room here to develop the argument in detail (for more, see Brännmark 2019), this does not seem to be the case.

To begin with, it needs to be noted that in terms of deontic statuses, we are talking here about rights and duties that have a real existence, not rights and duties that we should have, but ones we do have as a matter of social fact. Let us start with a simple example of how social rights and duties can be established in terms of how we develop certain expectations: You walk across my lawn while I am standing on it. I do not protest, I just say hello. You do it again. I do not protest, I just say hello. And then this continues. Eventually we are at a point where it makes sense to say that you have an informal *right*, mutually recognized, to walk across my lawn. If after several years, I suddenly would protest, it would be I rather than you who needed to explain my behavior. The reason is that by then certain stable expectations have formed. We could of course say that

what has happened is that a certain rule has become established (*you are allowed to walk across my lawn*), but in order to understand the situation, saying this appears explanatorily redundant. All we need to say is that because of certain regularities being in place, certain expectations have been formed and gradually solidified. In fact, postulating the existence of a rule that we both accept just seems to add to our explanatory burden: Exactly when did we get to the point where the rule was established? In exactly what sense do we accept it? What is the exact content of the rule?

Although a precise tracking of such mechanisms of synchronizing and coordinating social expectations will be more difficult on a societal level, there seems to be no principled reason against thinking in terms of this kind of explanatory model on that level as well. By looking at how we gradually adjust our expectations on each other, and how we come to synchronize and coordinate these expectations so that enough people hold similar expectations, certain regularities might be established. For instance, an institution like the family clearly involves a range of expectations on what individuals occupying certain roles should or should not do, and where by a massive amount of interactions stretching far back in time, we as individuals have gradually modified our expectations into a state where we are, at least on the whole, relatively synchronized and coordinated even on a societal level. Now, if we understand our statuses as flowing from certain rules we would presumably think that, say, *being a father* or *being a mother* comes with certain rights and duties and that these are then, *ceteris paribus*, the same for all fathers and all mothers. On the alternative nominalist approach, however, the main relation obtaining between the statuses that we hold as individuals will be similarity or resemblance, not sameness. It can make sense to speak of what is typically involved in *being a father* or *being a mother*, but

ultimately different fathers and mothers will have somewhat different rights and duties, in terms of the social expectations that they face, depending on their more exact concrete circumstances. The deontic statuses that they hold come from below rather than from above, so to speak.

One consequence of adopting this kind of approach would be that it becomes natural to understand institutions as temporally extended complex concrete objects. Identity over time will not be about how the same set of rules are instantiated at a series of different timepoints, but about continuity and connectedness throughout this temporally extended object. In fact, to the extent that we were to generalize rules from the patterns of expectations that obtain at different time points, these rules could, and probably would, be different at different times. Such differences would however then be about that particular institution being differently shaped at different times, not about a series of different institutions (*qua* systems of rules) consecutively being in place. Exploring this consequence in more detail, however, is something that will have to be left for another occasion.

References

Brännmark, J. (2019) Institutions, Ideology, and Nonideal Social Ontology, *Philosophy of the Social Sciences* 49, 137-159.

Gilbert, M. (2018) *Rights and Demands*, New York: Oxford University Press.

Hume, D. (1978) *A Treatise of Human Nature*, ed. L. A. Selby-Brigge, rev. ed. P. H. Nidditch, Oxford: Clarendon Press.

Lewis, D. (1969) *Convention*, Cambridge, Mass.: Harvard University Press.

Maurin, A-S. (2002) *If Tropes*. Dordrecht: Kluwer.

North, D.C. (1990) *Institutions, Institutional Change and Economic Performance*, Cambridge: Cambridge University Press.

Rawls, J. (1999) *A Theory of Justice, rev. ed*, Cambridge, Mass: Harvard University Press.

Rodriguez-Pereyra, G. (2016) Nominalism in Metaphysics, *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2016/entries/nominalism-metaphysics/>.

Will Science and Proven Experience Converge or Diverge? The Ontological Considerations.

Johannes Persson

Hans Larsson (1862-1944) is perhaps our most well-known Lund philosopher. He was a prolific writer, and the author of philosophical monographs, essays and also novels. He became professor of theoretical philosophy in Lund in 1901 and member of the Swedish Academy in 1925. Two of his research interests make him a natural starting point for this essay. First, he believed that there is reason and logic in experience and intuition, and that this gives experience and intuition a role not only in decision-making in ordinary life, but also in science and philosophy (Larsson 1899). Second, he became interested in the balance between the convergence and divergence of philosophical ideas (Larsson 1924). The philosopher as an individual often seeks divergence, he thought, but philosophical ideas tend to converge as they are worked out in depth and detail (Larsson 1944: 84-87). He supported his convergence thesis with a number of examples from the history of philosophy but he also accepted the general hypothesis that the convergence of principles follows from the ambition to see them universally applied (Larsson 1924: 217).

Larsson's claim was not that fully worked out philosophical ideas converge completely – a small but important "netto divergence" might remain and prolong their difference:

> Den handlar för resten om 'konvergens och divergens' och ensamt det säger tydligt nog ifrån att jag icke önskar divergensen förbisedd. Den är enligt min uppfattning tvilling till konvergensen. Jag finner denna då jag söker göra mig reda för vad som verkligen skiljer, divergensnettot, så att säga. Innan man kommer fram till det, får mycket av den gängse uppfattningens brutto dragas ifrån. Men denna sista divergens får aldrig utplånas. Mellan Mills välfärdslära och Kants pliktlära är bruttoskillnaden väldig. Men när man ser att Mill räknar rättvisan som en av de viktigaste välfärdsposterna, kommer välfärdsberäkningen oftast (jag tror alltid) att till resultatet sammanfalla med pliktbudet […] (Larsson 1944: 85)

Larsson's convergence idea is intuitive. It belongs to a family of such ideas, of which C. S. Peirce's version might be the most well-known:

> But human opinion tends in the long run to a definite form, which is the truth. […] There is, then, to every question a true answer, a final conclusion, to which the opinion of every man is constantly gravitating. (Peirce 1871/1992: 89)

It may be that the motivations behind some of the convergence ideas presented in the literature differ, and that this is true of Peirce and Larsson. In Larsson's case, there is ample room, it seems, for different kinds of negotiating processes of the kind that can be used by advocates of two or more philosophical or (indeed) political ideas. The mechanism behind convergence and divergence might be competition, or it may be more peaceful interaction between viewpoints aiming at explaining as many as possible of the appearances:

> Idéerna, dem skall man aldrig försona! Säger Oswald Spengler; de skola föra sin kamp till slut, så världen gå trött och bruten ut ur vår tids blodströmmar. Men då är det nog falska idéer det är fråga om; icke genomtänkta och mognade, alltför partiella för att med rätta kallas idéer (Spengler tänker på den tyska och engelska anden), ännu ej inställda under det universella förnuftskravet. (Larsson 1924: 208)

Here I propose to stick with Larsson's convergence-with-a-possible-netto-divergence idea, and to apply it to two systems of ideas or experiences—systems distinguished by their form, or perhaps the knowledge generating mechanisms they rely on, rather than the content of the ideas or experiences they contain.

In Sweden, the notion of 'science and proven experience' has featured in the regulation of healthcare for more than a century. In 1890, the Swedish king, Oscar II, issued a Royal Decree explicitly obliging a physician to "deliver such counsel, and, as far as circumstances permit, to extend such therapeutic endeavours, to every patient under his care as are necessitated by the patient's condition and as are consonant with science and proven experience" (Pontin 1891). Today, the Patient Act (2014:821, ch. 1, p. 7) states that patients shall be given medical care that is consonant with science and proven experience, and the Patient Safety Act (2010:659, ch. 6, pp.1-2) makes it clear that healthcare workers have a personal duty to perform their work in accordance with this standard.

Most of us would say that we know, roughly at least, what one of the conjuncts in *science and proven experience* is. The nature of scientific knowledge has of course been debated, and in these debates different features of what we count as science have been in focus at different times. Deductive proof used to be regarded as a hallmark of science, and so did certainty. At other times, scientific method and fallibility have been seen as more salient. Nowadays, the concept

of systematicity seems especially important (e.g. see ALLEA 2017). But at one point, at least, the idea that science is grounded in observation was equally influential.

However, the nature of the other conjunct—*proven experience*—is something we are less certain about. Clearly, proven experience has something to do with experience. Without experience it is impossible. We can add three further observations. (1) Someone can act in accordance with proven experience with no experience of his or her own of that of which there is proven experience. (2) Proven experience is often a particular kind of experience of a measure or treatment—namely, that it works. (3) Proven experience is generally well tried, in the sense that the belief or practice it validates is often put to the test repeatedly.

Science and proven experience are sources of evidence, and are treated as such (Persson et al. 2017). Sometimes only one of the two is present to a significant degree. Small-scale farming, in particular—one of the most common forms of employment in the world—is still based largely on knowledge acquired through practical experience (Altieri 2004; Akullo et al. 2007), and some of the most sustainable farming systems in the world are entirely based on knowledge and practices acquired through the practical experience of generations of farmers, the so-called Globally Important Agricultural Heritage Systems (GIAHS). Sometimes one of them is marginalized, politically and ideologically (Hountondji 2002).

Similarly, thrombus removal following ischaemic stroke involves removing the clot mechanically. The usual procedure is to try to dissolve the clot using drugs, but if the clot is a big one this is not always successful. In advance of scientific support, or clinical trials, specialized centres have tried to remove larger clots mechanically. The results have been good, and consensus as to the effectiveness of the technique has emerged. Here we have a case where medical decisions

are based on proven experience for a considerable time, until science catches up and corroborates the experience (e.g. Persson et al. 2018; Wallin et al. forthcoming).

Many of our interventions in the public sector—whether in health care, or social work, or primary or secondary education—are required to be consonant with both science and proven experience. But is this requirement only instrumentally, and perhaps temporarily, motivated by the fact that sometimes our most reliable knowledge comes in the form of proven experience while sometimes it emerges as scientific knowledge. Larsson's assumption would presumably be that eventually, provided enough work is put into science and proven experience, the two will become similar—partially identical, perhaps. Would they overlap completely? Or could there be a substantial netto divergence?

A straightforward—and too plain?—objection to Larsson's and Peirce's convergence thinking is that we may never reach convergence for the simple reason that we start asking different questions. If person A starts asking question Q1 and person B starts asking Q2, where is the guarantee that they will end up with the same beliefs or experiences?

There is no such guarantee, of course. The multitude of academic subjects, some of them with a very long history, testifies to this fact. Topic incommensurability (Hacking 1983) might be the result. This objection applies in a more interesting way to Larsson than it does to Pierce. Peirce, as we have seen, sometimes frames the convergence claim relative to a specific question, effectively excluding scenarios where A and B ask different questions. Larsson comes at the issues from a different angle, but on the other hand, he seems to restrict his discussion to philosophical system builders—who, it might be assumed, should deliver answers to every significant philosophical question.

There may be room for a connected complaint. The order of questions and answers can affect the result, or epistemic destination, considerably (e.g. Farber 2005). If we start with Q1, we are likely to understand it, and answer it, differently from the way we would do so if we were to approach it after dealing with Q2. Popper's P1>TT>EE>P2 schema is relevant here. Popper (1963) observes that, as a rule, error-elimination (EE) applied to a tentative theory (TT) leads to the emergence of a new problem (P2 as compared with P1).

These subtleties aside, it is certainly possible that questions raised from the scientific and the proven experience perspective are bound to develop along different trajectories.

Consider Larsson's conviction that the ambition of philosophical ideas is to be universally applicable. Not everyone would wholly agree. Baruch Fischhoff (2018) has argued that the philosopher's ambition is moderate in this respect, while the sciences are more ambitious. However this theoretical difference is settled, Fischhoff provides support, in effect, for the view that science is likely to develop its questions in accordance with the ambition that the tentative theories should be universally applied.

Things appear to be different when it comes to proven experience. Even if the ambition might in one sense be the same, namely to have proven experience of everything relevant, there is little to be said for the idea that proven experience should be formulated in such a way that it is always the same proven experience we rely on.

The unsought effect of such a strategy might even be that a third category–local knowledge–relating the abstract and universally warranted experience to the real cases, at the different hospitals, would have to be introduced and put alongside science and proven experience.

There is perhaps an ontological formulation of what has just been said. A theme has emerged in the VBE group working with Nils-Eric Sahlin:

Let us assume that proven experience is experience. Experience is in many ways similar to belief. It, too, has representational content. The idea of proven experience signals that it can be more or less uncertain, more or less robust. In addition, however, experience has a certain character. And it might be that proven experience has an additional character. It is, we may assume, unclear what these characters are. They might, for instance, be of the kind discussed in Leon (1987).

Proven experience can be shared. As mentioned earlier those without experience of their own of, say, the success of treatment X, can still reliably act on that experience. Given this, we cannot assume that the character of proven experience is always manifest as a physical token in each individual sharing it. But the token might still exist somewhere. Perhaps that is a condition of the proven experience's existence.

Something similar could have been accepted as true of scientific knowledge, especially if those who argued that scientific claims were only shorthand for more complicated claims about observations had been right. But it seems that they were not.

Hence there might be an ontological difference explaining why there will always be a netto divergence between science and proven experience.

References

Akullo, D., R. Kanzikwera, P. Birungi, W. Alum, L. Aliguma, and M. Barwogeza (2007). Indigenous knowledge in agriculture: a case study of the challenges in sharing knowledge of past generations in a globalized context in Uganda. World Library and Information Congress: 73rd IFLA General Conference and Council, Durban, South Africa.

ALLEA (2017) The European Code of Conduct for Research Integrity. Revised Edition. ALLEA - All European Academies, Berlin.

Altieri, M. A. (2004) Linking ecologists and traditional farmers in the search for sustainable agriculture. *Frontiers in Ecology and the Environment* 2:35-42. http://dx.doi.org/10.1890/1540-9295(2004) 002[0035:LEATFI]2.0.CO;2

Farber, I. (2005) Peirce on reality, truth, and the convergence of inquiry in the limit. *Transactions of the Charles S. Peirce Society* XLI(3), 541-566.

Fischhoff, B. (2018) The art of proven experience. *Science and Proven Experience – Johannes*. Lund: Mediatryck.

Hountondji, P. J. (2002) Knowledge appropriation in a post- colonial context. Pages 23-38 in C. A. Odora Hoppers, editor. *Indigenous knowledge and the integration of knowledge systems: towards a philosophy of articulation*. New Africa Books, Claremont, South Africa.

Hacking, I. (1983) *Representing and Intervening*. Cambridge University Press.

Larsson, H. (1899) *Poesiens logik*. Lund: Gleerups.

Larsson, H. (1924) *Filosofiska uppsatser*. Stockholm: Bonniers.

Larsson, H. (1944) *Post-scriptum*. Stockholm: Bonniers.

Leon, M. (1987) Character, content, and the ontology of experience. *Australasian Journal of Philosophy* 65(4), 377-399.

Peirce, C.S. (1871/1992) 'Review of Fraser's *The Works of George Berkeley*. In *The Essential Peirce* Vol. 1. Eds. Houser, N. and Kloesel, C. Bloomington: Indiana University Press.

Persson, J., Anttila, S. & Sahlin, N.-E. (2018) Hur förstå 'och' i 'vetenskap och beprövad erfarenhet'?, *Filosofisk tidskrift* 39(1), 44-54.

Persson, J., Vareman, N., Wallin, A., Wahlberg, L. & Sahlin, N-E. (2017) Science and proven experience: a Swedish variety of evidence based medicine and a way to better risk analysis?, *Journal of Risk Research*. https://doi.org/10.1080/13669877.2017.1409251

Persson, J., Johansson, E. & Olsson, L. (2018) Harnessing local knowledge for scientific knowledge production: challenges and pitfalls within evidence-based sustainability studies, *Ecology and Society* 23(4), 38. https://doi.org/10.5751/ES-10608-230438

Popper, K. (1963) *Conjectures and Refutations: The Growth of Scientific Knowledge.* London: Routledge & Kegan Paul.

Pontin, D. M. (1891) *Författningar m.m. angående medicinalväsendet i Sverige, omfattande år 1890.* Stockholm: P.A. Norstedt & Söner.

Wallin, A., Wahlberg, L., Persson, J., & Dewitt, B. (forthcoming). How should the epistemology of medicine inform the regulation of healthcare? Theoretical and empirical perspectives from Sweden.

The Ontology of Decision-Making

Nils-Eric Sahlin

We make many decisions each day. Some of them are rational, some irrational. Some we are entertained by, some we wish we'd never made … Most of them we will never remember, still less how we came to make them. We need help – or don't we? But shouldn't there be one and only one guide to rational action? A cacophony of competing theories may be slightly bewildering. How do we choose between them?

In his seminal paper "Subjective expected utility: A review of normative theories" Fishburn (1981) discusses the "primitives, axioms and representation-uniqueness theorems" of miscellaneous theories of rational decision-making. One reason for the large number of theories is what Fishburn describes as the interplay between realistic decision structures and structural axioms that help us derive the desired theorems. Another reason is of course the "primitives". The structure of the different types of entity varies from one theory to the next, and so do the relations among these entities. In other words, the ontological assumptions we make matter.

Asked what a decision is, most of us would say that it is a choice of one alternative over a set of other, possible alternatives, and that what determines the choice is decision maker's beliefs and desires.

More often than not we do not have control over the factors determining the consequences of our choices. Those consequences are determined by the way the world unfolds, by what turns out to be the true state of the world. Already we have something that looks like an embryonic ontology. We have acts, consequences and worlds. Fumbling onwards, we are outlining an ontology of decision-making. A well-designed ontology of decision-making will, we hope, make it easier to structure real-life decision problems while also making it evident how and when to apply the decision tools available. A rough-hewn or excessively proof-driven design, on the other hand, will be very likely to lead to something more or less inapplicable.

Ramsey's quintessential decision ontology contains consequences – worlds, he calls them – and propositions. Acts are constructed out of worlds and propositions. Ramsey's theory is, as Fishburn so elegantly puts it, "extremely rich, insightful, and very carefully reasoned". Among the primitives there is, of course, also a preference relation over worlds and acts. Ontologically, this relation is far more important than one might think. If not only consequences (worlds), but also propositions and acts consort with values, things quickly become messy. Ramsey saw this. His first axiom is an ontological axiom stating that there is an ethically neutral proposition believed to degree a half – a device designed to avoid the problem of proposition-dependent preferences.

Savage's ontology is in many respects similar to Ramsey's. His (1972: 9) decision ontology consists of the world, i.e. "the object about which the person [decision maker] is concerned"; states of the world, each being "a description of the world, leaving no relevant aspect undescribed", and true states, each of which is "the state that does in fact obtain, i.e., the true description of the world". An event is simply a set of states. An innovation in Savage's ontology is that acts are identified with their consequences. An act, as he puts it, is

"a function attaching consequences to each state of the world" (Savage 1972: 14) and it is assumed that there is a preference relation on this set of acts.

As already said, there are many theories of decision-making under uncertainty, all differing in various ways from the others. They assume different types of entity and postulate the existence of different types of relation between entities. Entities are not to be multiplied beyond necessity, we are told. Ontological parsimony is a virtue: an ontology inhabited by just tropes, facts or particulars is preferable to more abundant world views.

Why, for example, spend time constructing facts out of universals and particulars, if everything turns out better if we do it the other way around? Why assume entities and recombine them if they were never separated in the first place? Why get into a jam looking for a glue, or endless number of different glues, if such adhesives were not needed in the first place? If tropes are all we need, things will in many ways be far less complicated. (See, e.g. Maurin and Sahlin 2005)

Decision theoreticians, too, have tried to streamline their ontologies. Jeffrey (1965) is one of them. Do we really need worlds, propositions, events, consequences and acts? Can propositions do the job on their own? The answer is Yes, but that a thin ontology comes at a price. The ontology of Jeffrey's theory is so meagre that it is hard, if not impossible, to apply it, or to recognise in it a real-world decision problem. The very idea of what a decision is may be lost if the ontology is too streamlined. Ramsey's theory has an applied ontology that is carefully thought through, and therefore it does the work it should do. However, the question is whether Ramsey's decision ontology is also undernourished.

Years ago (Sahlin 1991), I criticised decision scientists for being too empirical. Empirical research can be too empirical. My complaint was that Rational Man (theories of rational decision-making) was not the best point of departure if we want to model, or understand, human decision-making. What I didn't appreciate at the time was that even if you take-off in the wrong direction and with a completely unrealistic map in your hand, you may after all end up in the right place. Today we can clearly see that some of the research done on human decision-making has given us invaluable knowledge.

Psychology, a good friend of mine says, is the study of abnormal behaviour. Aberrant behaviour illuminates the normal. Of course, we also have a desire to help people with psychological and psychiatric problems. But if the norm is Rational Man, we will all deviate from the norm – we are all abnormal. I stick to my old guns and maintain that theories of human decision-making need well-designed applied ontologies. This was not how I expressed myself at the time, but today I see that it is what it was all about. For example, clearly, we have to say what a belief is – what a value is. We must design an ontology that is rich enough, hangs together and applies to the problems we want to solve. The classical normative theories do not give us such an ontology. The "primitives" are important, propositions, events, acts, consequences, worlds … So are the relations among these entities. But somewhere there must be an agent – a *Homo economicus* or *sapiens*, with beliefs and values, information and preferences – taking a decision. What is a belief? What is a value? However we answer these two questions we have to design, or redesign, our decision ontology accordingly.

If we lack theories of belief and value, it will be difficult to explain human decision-making in depth. To make the case for this, I made use of three distinct conceptions of what it is to believe (1991). First, a mentalistic theory on which beliefs are mental acts. In Hume's

words: a belief "may be most accurately defin'd, a lively idea related to or associated with a present impression" (*A Treatise of Human Nature*, Book I, Part II, Section VII). Second, a dispositional theory of belief on which "I believe that p" means that I entertain p and have a disposition to act as if p were true. Third, beliefs as mental states, or as Ramsey puts it: "a belief ... is a map of neighbouring space by which we steer" (Ramsey 1929/1990: 46). We believe p if p is on our mental map. These views on what a belief is limit the options, the ways in which decision behaviour can be interpreted (explained and understood). If it is only by establishing a lively idea in a decision maker's mind that we can understand and explain decisions, we have to design our experiments accordingly. Equating belief with behavioural consequences, on the other hand, we risk being unable to explain people's behaviour through their beliefs and desires. And a map view of beliefs delivers explanatory power but at the same time requires empirical studies to be designed and carried out in the framework of that particular theory, which might be very difficult.

The point here is simple: the theory of belief we adopt – representationalism, dispositionalism, interpretationism, functionalism, eliminativism or instrumentalism (see Schwitzgebel 2015) – inevitably influences our interpretation of empirical results. I now see that this criticism can, with equal force, be directed against normative theories and their ontologies. However, this time it is a question not of explainability but applicability that is the focus. The theories of rational decision-making, however mathematically beautiful, are so meagre from an ontological point of view that it is hard to apply them in any real-life situation.

Schwitzgebel (2015) puts his finger on a particular but very important problem when he writes: "However, the phrase 'degree of

belief' may be misleading, because the relationship between confidence, betting behavior, and belief is not straightforward. The dispositionalist or interpretationist may regard exhibitions of confidence and attitudes toward risk as only part of the overall pattern underwriting belief ascription. Similarly, the representationalist may hold that readiness to deploy a representation in belief-like ways need not line up perfectly with betting behavior." With a thin ontology, the type of ontology a mathematician likes, we do not even notice that this is a problem. We have what we need to prove the desired representation theorems, and that is what we wanted to do in the first place. But with a thick, more applied type of ontology, the problems suddenly appear, and we note that they are both interesting and intricate.

Shaffer (2009) has argued that Bayesian "decision-theoretic characterizations of decision situations fail to adequately account for knowledge concerning the causal connections between acts, states, and outcomes in decision situations". This, he says, makes them incomplete, and the problem is fundamentally an ontological one. This is a much-discussed issue, and one that shows how important it is to take ontological questions seriously before indulging in mathematics. In a forthcoming paper Fischhoff (2019) discusses the interplay between bounded rationality, satisficing and risk-taking. He is not upfront discussing any ontological questions, but what is said in the paper still pinpoints the need for good applied ontologies in the area of risk-taking. A good example of the way in which re-thinking the ontology of decision-making and risk can produce a more capable theory is to be found in Weirich's forthcoming book *Rational Responses to Risks*. But there is far more to be done.

Ontologists have taught us how important applied ontologies are – e.g. ontologies of medicine, and of institutions, and of law, war and crime. I hope these scattered thoughts of mine may inspire

someone, ideally an eminent metaphysician, to study the ontology of decision-making with greater care. In the end, that will provide us with better tools for wise decision-making.

References

Maurin, A.-S. and Sahlin, N.-E. (2005) Some Ontological Speculations: Ramsey on Universals, Particulars and Facts, *Metaphysica* 3(3), 7-28.

Fishburn, P. C. (1981) Subjective Expected Utility: A Review of Normative Theories, *Theory and Decision* 13, 139-199.

Fischhoff, B. (2019) The Microbiomes of Gut-Level Decisions, forthcoming in *Critical Quarterly*.

Jeffrey, R. C. (1965) *The Logic of Decision*, Chicago: University of Chicago Press.

Ramsey, F. P. (1929/1990) Truth and Probability. In Mellor, D. H. (ed.), *Philosophical Papers*, Cambridge: Cambridge University Press.

Ramsey, F. P. (1929/1990) General Propositions and Causality. In Mellor, D. H. (ed.), *Philosophical Papers*, Cambridge: Cambridge University Press.

Sahlin, N.-E. (1991) Baconian Inductivism in Research on Human Decision Making, *Theory & Psychology* 4(1), 431-450.

Savage, L. J. (1972) *The Foundations of Statistics*, second revised edition, New York: Dover Publications.

Schwitzgebel, E. (Summer 2015 Edition) Belief. In Zalta E. N. (ed.), *The Stanford Encyclopedia of Philosophy*, URL = <https://plato.stanford.edu/archives/sum2015/entries/belief/>.

Schaffer, M. J. (2009), Decision Theory, Intelligent Planning and Counterfactuals, *Minds & Machines* 19, 61–92.

Weirich, P. (forthcoming) *Rational Responses to Risks*, Oxford: Oxford University Press.

How Cellular Automata Dissolve the Fine-Tuning Argument

Einar Duenger Bohn

I have always found the so-called fine-tuning argument to be among the more rationally respectable and interesting arguments in favor of the existence of a god, or more precisely, in favor of the existence of a supernatural intention behind the universe.[1] I once mentioned this to Anna-Sofia, that I was somewhat impressed by the fine-tuning argument. She gave me the incredulous stare. She was *not* impressed, neither by the argument nor by me.

Her incredulous stare made me think again. I now suspect she was right. So, in her honor, I will in what follows show how work in computer science on so-called *cellular automata* in fact dissolves the fine-tuning argument.

[1] See Bohn (2016).

The fine-tuning argument: where it ends up

Roughly, here is what I used to consider a plausible version of the fine-tuning argument (FTA). The fundamental physical laws, constants and ratios are extremely fine-tuned for physical life (or, perhaps even better, for consciousness) in the sense that each one of them has one out of extremely many possible values, and pretty much the one value each one of them in fact has is the only value among extremely many possible values they could have had in order for there to be physical life as we know it, anywhere in the universe.[2]

The question is: Why is the one and only possible universe with life also the one and only actual universe shining bright with existence? In other words, why isn't the actual universe without life, given that it is much more likely? This cries out for an explanation, and there are three leading contenders. First, it is due to chance; shit happens, and life happens. Call this the chance hypothesis. Second, it is due to the fact that all possible universes are actual universes; and no wonder we could not be in a universe we could not be in. Call this the multiverse hypothesis. Third, it is due to there being an intention behind it all; there is a god after all. Call this the god hypothesis. By a common principle of evidential favoring employed not only in the heart of the hard sciences, but throughout all of rational reasoning, whichever one of the three hypotheses makes the fine-tuning most likely is also the hypothesis most supported by the evidence.

As an explanation of the fine-tuning, each one of the three hypotheses has its own pros and cons, but for present purposes we need

[2] Strictly speaking, there is probably a small range of values compatible with life, but the range is so small compared to all possible values that for all practical purposes we can think of it as a single value for each law, constant and ratio.

not even choose between them because the argument is blocked before we even get that far. We need to ask: What is it that cries out for an explanation? Obviously, it is not the mere fact that the physical laws, constants and ratios have the values they do that cries out for an explanation, but the fact that those actually existing values are the only ones that can support life. Ultimately, it is the extreme unlikelihood, but still actuality of life that cries out for an explanation. In other words, ultimately it is the fact that there is actually *life* in contrast to non-life that really cries out for an explanation. So, ultimately, the question becomes: why does life as such cry out for an explanation?

If life cries out for an explanation, neither the chance hypothesis nor the multiverse hypothesis will explain it; if life doesn't cry out for an explanation, the god hypothesis is neither here nor there. Also, none of us knows what it really means for something to cry out for an explanation. Why do some things cry out for an explanation, but others not? This is where I think the debate over the fine-tuning argument ends up without much more to say to move it forward other than simply that life is an amazing and mind-blowing thing that we feel we should explain. The problem is that some feel it, but others don't. In short, some feel life needs an explanation, but others don't. None of us knows why we have that feeling, nor what could justify such a feeling.

Cellular automata: how work in computer science dissolves the argument

First, consider a plane of white squares, infinite in all directions. Second, pick an arbitrary square on the plane, paint it black, and move

to the row below it. Third, consider various very simple algorithms (if-then rules) telling you which squares to paint black on a row, given how the black and white squares are arranged on the row above it. Fourth, follow the various simple algorithms where they lead, at least for several thousand steps, preferably millions, or even billions.

Such constructions are known as *cellular automata*.[3] The better our computers get, the greater such cellular automata we can realize and study the result of, which has been done in computer science for many years already. What shocks everyone who starts studying cellular automata is that, on the basis of one black square (with the rest being white), very simple algorithms with very small differences between them in fact leads to very complex patterns with very big differences between them. What's more, some of the patterns look completely random and chaotic, others boringly repetitive, but others beautifully organized throughout, while yet others look organized in some of its areas but not throughout. What's even more, if you zoom in on a pattern, it might look nothing special, but if you zoom out, you'll see the most beautiful or artsy pattern you have ever seen.[4]

And that is just with two colors: black and white. If you add grey, not to mention all other colors, it will most definitely blow your mind. And we can only begin to imagine the results of cellular automata in four or eleven dimensions, even on the basis of a simple handful of properties.

What's most important for present purposes though is that for some of these patterns, if you saw it on a rock you found on the beach, you would think nothing special of it, it would be like en-

[3] See e.g. Wolfram (2002) for an ambitious, punchy and fun introduction.

[4] See e.g. Adamatzky & Martinez (2016) for some artsy patterns.

countering any other rock on the beach, but for others of these patterns, you would think there was something special about it, seeing it on a rock would be like encountering a watch on the beach. You would think it cried out for an explanation: Whence that beautiful and amazing pattern?! It cannot be random!

But the sober fact we have learned from work on cellular automata in computer science over the last couple of decades is that it all arises from simple and boring algorithms on e.g. the basis of a black square among a bunch of white squares. Such a simple and boring algorithm gives you the pattern you think nothing special of, but an almost identical simple and boring algorithm also gives you the pattern you think something special of. So, you believe there is an important difference, but you're wrong. *As a matter of fact*, there is no important difference between the two patterns, for any sense of 'important' you might here care about.

But then at least I start suspecting that the exact same thing might be going on in the case of life (or consciousness). We think nothing special of the patterns of universes with no life, but some of us think something special of the patterns of universes with life. Some thus believe there is an important difference. But given what we now know from work on cellular automata in computer science, we should suspect, in fact hypothesize that those who believe there is an important difference here are just wrong. We should hypothesize that, *as a matter of fact*, there is ultimately no important difference between the two patterns. They both are the result of the same kind of simple and boring algorithms, based on the same simple and boring handful of dimensions and properties. At least we should no longer trust our own sense of which patterns are special, or in need of an explanation, because we are, as a matter of fact, often wrong elsewhere. Work on cellular automata thus *empirically show* that the distinction between what we think cries out for an explanation and

what we think don't cry out for an explanation is often all in our heads. Given that I know this is the case with respect to cellular automata, I cannot conclude that it is not the case with respect to life. For all I know, it is all just patterns arising from a handful of properties and a completely uninteresting algorithmic process. At least that is the simplest and most empirically informed hypothesis.

Conclusion

The fine-tuning argument thus leads to neither the chance hypothesis, nor the multiverse hypothesis, nor the god hypothesis. Rather, the very heart of the fine-tuning argument dissolves before we should conclude anything, and it dissolves from considering cellular automata: they empirically show us that there is no good reason to believe that life (or consciousness) is special, even though some of us feel it is. Life's algorithm might very well be as simple and boring as any other, even though its result *seem to us* very exciting, in need of an explanation. The study of cellular automata show that we cannot in any way trust our own sense of some patterns being special, or in need of an explanation.

So I now think that Anna-Sofia's incredulous stare was justified. The fine-tuning argument gives us *no good reason* to conclude from the fact that *we* find life (or consciousness) in need of an explanation that it in fact needs an explanation compared to non-life (or non-consciousness). The fine-tuning argument might tell us something interesting about our psychology, but it does not take us closer to the existence of a supernatural intention behind our universe.

Acknowledgement

Thanks to Atle Ottesen Søvik for comments.

References

Adamatzky, A. & Martinez, G.J. (eds.) (2016). *Designing Beauty: The Art of Cellular Automata*, Springer.

Bohn, E.D. (2016). Why this Universe? In Asle Eikrem & Atle Ottesen Søvik (eds.) *Thinking Seriously about God*, LIT Verlag, 2016.

Wolfram, S. (2002). *A New Kind of Science*. Wolfram Media Inc.

What is Field's Epistemological Objection to Platonism?

Ylwa Sjölin Wirling

1. Introduction

This paper concerns an epistemological objection against mathematical platonism, due to Hartry Field (1989). Platonism is a view according to which mathematical truths are about mind- and language-independent, abstract, and acausal entities. The argument poses an explanatory challenge – *the challenge to explain the reliability* of our mathematical beliefs – which the platonist, it's argued, cannot meet. Is the objection compelling? Philosophers disagree, but they also disagree on (and are sometimes very unclear about) how the objection should be understood. Here I distinguish some options, and highlight some gaps that need to be filled in on the potentially most compelling version of the argument.

2. Two characteristics

Here are two characteristics (C1 and C2) that any good reading of Field's objection should have.

It's commonly agreed that Field's argument constitutes an improvement on Paul Benacerraf's (1973) influential dilemma for platonists, primarily because it doesn't rely on any particular theory of knowledge. Benacerraf's dilemma (BD) can be stated as follows: platonism is our best theory of mathematical truth, seeing that it preserves our strong conviction that mathematical truths are objective and mind-independent, but platonism is incompatible with the claim that we have mathematical knowledge, seeing that our best theory of knowledge has it that knowledge requires a causal dependence of beliefs on the relevant facts. Read as an epistemological objection against platonism, the upshot is that assuming platonism, mathematical knowledge is impossible. This relies on the outdated causal theory of knowledge, which makes it uncompelling. But BD does point to something interesting, according to Field, namely the puzzling question of how our mathematical beliefs can be so accurate if they are about platonic objects (1989: 25-26). Field shows that this can be posed as a problem for the platonist without appeal to any particular theory of knowledge. He proceeds from the assumption that our mathematical beliefs are largely true. That is, our methods for supporting mathematical beliefs are reliable with respect to finding out about the mathematical facts.[1] Such reliability must be explained. This explanatory challenge is at the heart of Field's objection (FO), and the point is that assuming platonism, it cannot be

[1] Schechter (2010: 441-443) has pointed to the importance of focusing on the reliability of the method rather than the reliability of the beliefs.

met.[2] Thus, FO should be independent of any particular theory of mathematical knowledge (C1).

Second, FO is distinct from two standard nominalist arguments against platonism (C2). Platonist often claim that mathematical objects are indispensable to our best scientific explanations, and therefore we are justified in postulating their existence (see e.g. Colyvan 2001; Baker 2005, 2009). One common nominalist argument seeks to establish that this indispensability claim is false. It's well-known that Field gives such an argument elsewhere (1980), but FO is supposed to be a distinct objection against platonism. Another common nominalist strategy attacks the indispensability project more generally, i.e. questions the Quinean idea that we are justified in postulating the existence of an entity *x* if *x* is indispensable (see e.g. Maddy 1992; Melia 2000; Leng 2002; and Finn 2017). But Field clearly accepts indispensability as ontological justification, so FO should be compatible with that general idea.

In sum, FO should not rely on the correctness of any particular theory of mathematical knowledge, and the argument should be relevant even assuming that there is an indispensability argument supporting the claim that mathematical objects exist. Is there a compelling version of such an argument?[3]

[2] Notably, the problem is not that platonists haven't *actually* provided an explanation, but that it seems they couldn't, even in principle.

[3] I want to stress that my aim here is not exegetical, i.e. not to pin down the argument Field actually had in mind, but to find the most compelling version of the argument given certain constraints that arguably distinguish Field's general approach here from other arguments in the same ballpark.

3. Two readings, in light of C1

The conclusion of FO is that the reliability of our mathematical beliefs, like the belief that 2+2=4, cannot be explained, given platonism. There is no doubt that this is supposed to "lower philosophers' confidence" in platonism, as Liggins (2018) puts it. But *why* is it bad for platonism if this reliability is unexplainable? There are at least two possible answers here, which can be presented as two readings of the following passage by Field:

> The idea is that *if it appears in principle impossible to explain this* [the reliability of our mathematical beliefs], then that tends to *undermine* the belief in mathematical entities, whatever reason we might have for believing in them (1989: 25-26, emphasis in original).

First, we may read it as saying that if reliability is unexplainable then our mathematical justification, for beliefs like 2+2=4, is undermined. This reading gives us what I call the sceptical version of Field's objection, SFO for short. It can be constructed as follows:

a If it seems in principle impossible to explain the reliability of our mathematical beliefs, then any *prima facie* justification we have for our mathematical beliefs is undermined.

b Assuming platonism, it seems in principle impossible to explain the reliability of our mathematical beliefs.

c Assuming platonism, our mathematical beliefs are not justified. (*a, b*)

In a nutshell, unexplainable reliability implies mathematical scepticism, and *that's* why we should reject platonism if SFO is sound.

The sceptical reading is endorsed by e.g. Baras (2017); Burgess and Rosen (2005); Clarke-Doane (2017); Pust (2004); and Rosen (2001).

On another, less popular but more literal, reading, what is undermined ("the belief in mathematical entities") is *an ontological belief,* about the existence of the entities postulated by the platonist. This makes a difference: not being justified in claiming that 2+2=4 is one thing, not being justified in claiming that platonic numbers exist is quite another. We may thus construct the objection as follows:

d If it seems in principle impossible to explain the reliability of our mathematical beliefs given some theory *T*, then the ontological postulates of *T* should be rejected.

e Assuming platonism, it appears in principle impossible to explain the reliability of our mathematical beliefs.

f Platonic mathematical objects should be rejected (*d, e)*.

I'll call this the ontological version of Field's objection, OFO for short. In a nutshell, we should reject platonism because its core ontological claim – that there are platonic mathematical objects – should be rejected. Something like this reading is endorsed by Liggins (2006; 2010; and 2018).

The two readings differ importantly, in light of C1. Both readings are independent of any particular theory of mathematical *knowledge,* but SFO very centrally rests on a claim about the necessary conditions of justified mathematical belief, since it ties the badness of rendering reliability unexplainable to the undermining of justification for mathematical beliefs. Clarke-Doane (2017: 20), willingly acknowledges this:

> According to Field, if one's beliefs from a domain *F* are justified then it does not appear to her in principle impossible to explain the reliability of her *F*-beliefs.

This is expressed in the argument by *a*. But insofar as justification is a necessary condition for knowledge, this also involves a substantial assumption about the conditions of mathematical knowledge. To be sure, it's not considered as implausible an assumption as the causal theory of knowledge assumed by BD. Nevertheless, there is about as little in the way of consensus when it comes to theories of justification as there is when it comes to theories of knowledge. OFO, in contrast, does not make any assumptions about what is required for knowledge *or* justification of mathematical beliefs. Nor does it make any assumptions about whether our mathematical beliefs actually amount to either knowledge or justified belief. It only has it that they are reliably true (whatever else that might entail, epistemically speaking) and that this must be explained.

OFO thus seems preferable if the aim is to have an objection independent of theories of knowledge. It's potentially more compelling than SFO, since the platonist won't be able to reject it on the basis of not subscribing to any particular epistemological theory. Now, it might well seem an odd ambition for an epistemological objection to be independent of any epistemologically substantial assumptions, so C1 itself could perhaps be questioned. But even then, SFO has problems. It has been argued at length by e.g. Baras (2017), Burgess and Rosen (2005) and Clarke-Doane (2017), who all assume SFO to be the correct reading, that the epistemological argument against platonism is uncompelling because premise *a* is false. For the moment then, OFO seems to be the most promising version.

4. Defeating ontological justification

But for OFO to present a compelling case against platonism, some gaps must be filled in. Most obviously perhaps: what does unexplainable reliability have to do with ontological justification (or lack thereof)? Differently put, why should one accept *d*? There is no good answer in the literature to date. Liggins just says that the lack of reliability explanation is an embarrassment for platonism because the reliability of our mathematical beliefs is "the sort of phenomenon which demands explanation". Field himself similarly stresses that it's bad because it forces the platonist to regard this reliability of ours as a *brute fact* (1989: 238), which is highly unpalatable. Granted, it's perfectly legitimate to regard *some* facts as brute, but this reliability-fact just isn't one of them. But how, or why, should we take this to impact the question of ontological justification?

Now, recall that the most common, and most forceful, reason to believe in the existence of platonic objects comes from an indispensability argument. And by C2, the point of FO can't be that platonic objects are *not* indispensable. The point must be that even if they are, this does not justify the claim that they exist. Platonism's commitment to the bruteness of a fact which appears to demand an explanation, is thus supposedly some form of *defeater* for the justification from indispensability.

It's common to distinguish between *rebutting* and *undercutting* defeaters for some claim that *p*. A rebutting defeater for *p*, is a reason to believe not-*p*. An undercutting defeater for *p* is a reason to think that one's original justification for *p* is not sufficiently indicative of the truth of *p*. Either the commitment to a brute fact can be pitched as a rebutting defeater, i.e. as a reason to think that platonic objects do *not* exist. Or, it can be pitched as an undermining defeater, i.e. as a reason to think that the apparent justification for the claim that

platonic objects exist isn't sufficiently indicative of the truth of this claim. In particular, OFO either gives a reason to reject the existence of platonic objects *to be weighed against* the reason to affirm their existence, afforded by an indispensability argument. Or, it gives us a reason to think that their indispensability isn't, after all, sufficiently indicative of their existence – it breaks the link between the conclusion of a valid indispensability argument and the platonic existence claim.

Both of these options are interesting, but badly in need of further elaboration. Consider first the rebutting strategy, according to which the principled absence of a reliability explanation is a reason to reject the existence of platonic objects. Is this plausible? Is it the job of an entity to enable us to explain how we come to have reliable beliefs about it? Undoubtedly, it's intellectually frustrating if our reliability turns out to be a brute fact, and in some sense unintuitive. But is *that* really a reason to doubt the existence of the entities that seemingly put us in this situation? Moreover, for OFO to be a real threat to platonism, it must be a pretty strong reason, if it is to compete head-to-head with the positive reason afforded by an indispensability argument. Because remember that in light of C2, the objection should be independent of whether mathematical objects are indispensable.

Consider next the undercutting strategy. It needs to be specified how commitment to brute reliability can break the link between indispensability and existence. Since, again by C2, the point with FO isn't a wholesale attack on indispensability arguments, the wielder of this strategy would presumably be saying something like this: indispensability of *x*s is only a reason to assume the existence of *x*s under certain circumstances, and the commitment to a brute reliability fact suggests that these circumstances are not at hand in the case of platonic objects vis-à-vis our best scientific explanations. But what are

the relevant circumstance, and why should we doubt that they obtain when reliability turns out to be a brute fact?

In sum, on either way to pitch OFO there are a number of assumptions about ontological justification and how certain explanatory tasks play into such issues, which must be spelled out and then assessed, before it can be decided whether there is a compelling version of Field's epistemological objection to platonism.

5. Conclusions

Is there a compelling version of Field's epistemological objection to platonism? Assuming we want an argument free of substantial epistemological assumptions, the most compelling reading of the objection has it targeting the justification for the existence claim at the heart of platonism. However, challenges lie in wait for someone wishing to pursue this version of the argument. I distinguish between two ways to pitch it: as a rebutting defeater and as an undercutting defeater for the justification afforded by an indispensability argument. In either guise, the objection raises multiple questions concerning ontological justification, especially in light of how Field's objection is supposed to relate to other arguments in the literature. Assuming the objection is uncompelling on the sceptical reading (seeing that it rests on substantial assumptions about mathematical justification, the tenability of which have been questioned elsewhere), whether there *is* a compelling version of Field's epistemological objection thus depends on the tenability of certain assumptions about ontological justification that are yet to be spelled out. Doing so is further work awaiting anyone wishing to argue that unexplainable reliability defeats the platonist's ontological justification.

References

Baker, A. (2005) Are there Genuine Mathematical Explanations of Physical Phenomena? *Mind 114* (454), 223-238.

Baker, A. (2009) Mathematical Explanation in Science. *British Journal for the Philosophy of Science 60*(3), 611-633.

Baras, D. (2017) Our Reliability is in Principle Explainable. *Episteme 14*(2), 197-211.

Benacerraf, P. (1973) Mathematical Truth. *Journal of Philosophy 70*(19), 661-679.

Burgess, J. P., & Rosen, G. (2005). Nominalism Reconsidered. In Shapiro, S. (Ed.), *The Oxford Handbook of Philosophy of Mathematics* (pp. 515-535). New York: Oxford University Press.

Clarke-Doane, J. (2017) What is the Benacerraf Problem? In Pataut, F. (Ed.), *New Perspectives on the Philosophy of Paul Benacerraf: Truth, Objects, Infinity* (pp. 17-43). Switzerland: Springer International.

Colyvan, M. (2001) *The Indispensability of Mathematics*. Oxford: Oxford University Press.

Field, H. (1980) *Science Without Numbers* (Vol. 17): Princeton University Press.

Field, H. (1989) *Realism, Mathematics & Modality*. Oxford: Basil Blackwell.

Finn, S. (2017) The Role of Existential Quantification in Scientific Realism. *Philosophy, 92*(3), 351-367.

Leng, M. (2002) What's Wrong with Indispensability? *Synthese 131* (3), 395 - 417.

Liggins, D. (2006) Is there a Good Epistemological Argument Against Platonism? *Analysis 66*(290), 135–141.

Liggins, D. (2010) Epistemological Objections to Platonism. *Philosophy Compass 5*(1), 67-77.

Liggins, D. (2018) The Reality of Field's Epistemological Challenge to Platonism. *Erkenntnis, 83*(5), 1027-1031.

Maddy, P. (1992) Indispensability and Practice. *Journal of Philosophy 89*(6), 275-289.

Melia, J. (2000) Weaseling Away the Indispensability Argument. *Mind 109*(435), 455-480.

Pust, J. (2004) On Explaining Knowledge of Necessity. *dialectica 58*(1), 71-87.

Rosen, G. (2001) Nominalism, Naturalism, Epistemic Relativism. *Noûs 35*, 69 - 91.

Schechter, J. (2010) The Reliability Challenge and the Epistemology of Logic. *Philosophical Perspectives 24*(1), 437-464.

Reflections on Metaphysical Explanation

Rögnvaldur D. Ingthorsson

The nature of metaphysical explanation is a question that should be constantly on every metaphysician's mind, and yet it is rare to see explicit statements about the methodological approach that writers take. We tend to just enter the flow of ideas and words in a particular 'discourse' and see where it leads us. It is easier that way but can lead us astray. I can't claim to be a role-model in this respect. I have offered a comment here, a remark there, but plenty room for improvement. However, I have come across quite a few confusions that can be traced to failed understanding of method/approach, and one or two really interesting statements of method. Here I share one such confusion about method, and one interesting view about method.

1. A confusion about Hobbes

The necessity of causal connections is usually associated with causal realism. Say, Aristotle's claim that "whenever the potential active and potentially affected items are associated in conditions propitious

to the potentiality, the former must necessarily act and the latter must of necessity be affected" (*Metaphysics*: Bk. 9, §5). For Aristotle, this is not the result of an *a priori* argument outlining a logical/conceptual necessity, but a consequence of a certain view of material reality; objects with powers interact to provoke changes in each other. I have always supposed that this was the view Hume targeted (*Treatise*: Bk. I, Part 3).

However, in Elizabeth Anscombe's famous discussion of causal necessity, we see Hobbes being represented as Hume's target and as someone who 'evidently' sees causal necessity as a "logical connection of some sort" (Anscombe 1971: 89). To be sure, Hobbes does incriminate himself in the following passage:

> […] an entire cause, is the aggregate of all the accidents both of the agents how many soever they be, and of the patient, put together; which when they are all supposed to be present, it cannot be understood but that the effect is produced at the same instant; and if any one of them be wanting, it cannot be understood but that the effect is not produced (Hobbes 1656: Ch. X, §3)

However, I think Anscombe is wrong to infer from Hobbes' use of the phrase 'cannot be understood' that he is developing a purely logical argument, at least not of the *a priori* kind.

Hobbes' epistemology is empiricist, albeit with clear rationalist overtones. He thinks the *senses* provide us with *empirical knowledge* about the powers of material bodies, and thus knowledge of causes, because powers are causes (Hobbes 1656: Ch. I, §4). From our empirical knowledge of these causes we can then rationally calculate the effects they produce, and, *vice versa*, the causes from the effects. In light of this, I think it would be more charitable to interpret Hobbes as making claims about what can/cannot be thought, *given* the na-

ture of the external world *as it is empirically known to us* (or assumed). This could offer a basis for a deductively valid argument, yes, but using premises with material (non-logical) content which is empirically true; hence the conclusion is *a posteriori* of experience. If we find fault with his claims about what can or cannot be understood, we should find fault with his empirical knowledge and not with his logic.

Consider the same reasoning in the early 18th century, from someone with knowledge of Newton's *Principia*. That person could argue like Hobbes that—*if* Newtonian mechanics is accepted as a true description of the world—it cannot be understood, on pain of contradicting Newton's mechanics, but that if an object is acted upon by an external force it will change its state of motion in proportion to the force applied. It is a logically valid argument, but moving from premises based on empirical research, and so the conclusion is *a posteriori*. The approach suits a naturalist approach to metaphysical explanation.

Indeed, Naomi Thompson (2019) outlines something very similar. On her view, metaphysical explanations are subject to epistemic constraints imposed by the context in which a question is asked; they are not explanations of what must be the case without regards to any concerns except what can or cannot be conceptualised.

Hume may well have had Hobbes in mind. But he didn't just attack his logic. Hume first had to deny Hobbes' premise that the senses give us knowledge of the nature of external objects, and thus turned the question of causation into a mere conceivability issue; one unrestrained by epistemic concerns (*Treatise*: Bk. I, Part 3).

Now, it is difficult to assess today in what way exactly it can have mattered that Anscombe misrepresented Hobbes' position. But we can tell that it did matter.

2. McTaggart's approach to metaphysics

McTaggart's *The Nature of Existence* (1921/1927), is well known for containing McTaggart's argument for the unreality of time. It is less well known for containing a long and very interesting elaboration of metaphysical method. Indeed, he devotes the whole of chapter 3, 'Method', to an elucidation of his approach. Contemporary readers may be even more surprised to find out that we can find something useful about metaphysical explanation in the writing of a self-confessed Hegelian idealist.

It is important to note that McTaggart is an idealist in two different ways. First, he is an idealist with respect to the method of metaphysics; a *methodological* idealist. He thinks we can only gain knowledge about the ultimate structure of reality—which is the aim of metaphysics—through an examination of the general content of our ideas; not through a study of the content of our experiences of particular matters of fact. Empirical observations are inadequate to the task because they are particular both with regard to the qualities observed and the identity of the entities observed, while the ultimate structure of reality will have to be perfectly general. We might be able to infer by abstraction from a number of particular observations that they have some general feature in common, but this would fail to show that this general feature belongs to more than just the limited sample we have observed. Instead, McTaggart believes (we may be tempted to downgrade this to *hopes*) that the rational mind is able to directly 'grasp' the general nature of the fundamental features of reality via rational reflection of the ideas we have about them. In other words, his preferred method of finding out about the ultimate structure of reality is by introspective reflection on the general content of our minds.

Today it is popular to ridicule idealism as a silly idea from the past, but in actuality much of contemporary metaphysics cannot really be distinguished from idealism. Whenever anyone focuses solely on the conceptual connections between ideas in their philosophy, which is the essence of *a priori* philosophical reasoning, they are doing pretty much what the idealists thought philosophy is all about. Indeed, anyone who thinks that the theories and findings of empirical science is irrelevant to metaphysics because it can only tell us about the actual world—just one manifestation of all the possible worlds allowed by the ultimate structure of reality—are applying a method I find difficult to distinguish from McTaggart's.

Second, McTaggart is an idealist with respect to the nature of reality; an *ontological* idealist. This is perhaps the aspects of idealism that tends to be ridiculed. He thinks he can show that reality cannot fundamentally be material, and that therefore reality must instead consist of spiritual substance (ideas). His argument for this conclusion is pretty much identical to Berkeley's, notably that we are mistakenly inferring from the fact that we are acquainted with phenomenal properties in experience, that there must be something non-phenomenal that is causing the experience. Now, we may not want to accept this argument as proof of the conclusion that there is no material reality outside the mind. But, I think we should pay attention to the importance McTaggart bestows on subjective experience, as a way to judge the success of metaphysical theories. I think it holds good for idealists and material realists alike.

3. Taking experience seriously

According to McTaggart, the way things appear to be in experience does not give us knowledge about what things are really like. Even an idealist can distinguish what ideal reality appears to be like and what it is really like. He nevertheless thinks that the content of our experience provides an important criterion for the success of metaphysical theories. A metaphysical theory need not portray reality as it appears to be in experience. Indeed, such a theory is often trying to say what things are really like, as opposed to what they appear to be like. However, since the content of experience constitutes a rare type of *certain and indubitable knowledge*—what Russell called 'knowledge by acquaintance', notably of our own inner states—every adequate metaphysical theory must be able to explain how the experiences we actually have can arise. *If* it says the world is not as it appears to be, it must be able to explain how it can appear to be otherwise

To calm the jitters of hardcore realists, let's be clear about the limits of knowledge by acquaintance. It is a form of knowledge that doesn't extend beyond the content of the experience itself. My experience of an apple on a table indubitably constitutes knowledge of the state of my mind, notably that I am having an experience of that kind. But, it does not constitute knowledge of whether there is in reality an apple on the table. However, since the former is indubitable knowledge, our account of reality must take into account the fact that reality contains my experience of an apple on a table.

McTaggart's idea, roughly, is that any theory about what things are really like can be tested by asking whether it offers a conceptual model of reality that explains why things actually appear to us in the way that they do. The model need not resemble the appearance, but if it differs from the appearance it must explain what it is about the

world that allows it to appear so different from what it really is. As long as the model cannot explain the characteristics of experience, the facts of experience constitute an anomaly for the theory. On the other hand, when we have a model that claims the world is different from experience, yet is able to explain how it can appear in the way that it does, the appearance becomes a *phenomenon bene fundatum*; a well-founded datum of experience (McTaggart 1927: §494).

I think we can observe that many of the core disputes in metaphysics revolve around the question of how well theories explain experience. For instance, presentism is meant to have the upper hand on eternalism in explaining why we only ever experience the present. Not just because presentism says the world is pretty much as it appears to be, but more because eternalism cannot adequately explain how we can have a continuous experience of successive states and of ourselves as continuously remaining in the present, when in fact nothing exists continuously through time or moves from one time to another. For all its other faults, McTaggart's philosophy has one piece of good advice; take experience seriously as a criterion of success for metaphysical explanations.

References

Anscombe, G. E. M. (1971). Causality and Determination, in M. Tooley & E. Sosa (eds.), *Causation,* 1998, Oxford: Oxford University Press.

Aristotle, (1924). *Aristotle's Metaphysics*, W.D. Ross (ed.), Oxford: Clarendon Press.

Hobbes, T. (1656). *Elements of Philosophy Concerning Body,* in W. Molesworth (ed.), *The English Works of Thomas Hobbes of Malmesbury*, 1839, London: Bohn.

Hume, D. (1739). *Treatise of Human Nature*, edited by L. A. Selby–Bigge, 1896, Oxford: Clarendon Press.

McTaggart (1921/1927), *The Nature of Existence,* Vol. 1 & 2, Broad, C.D. (ed.), 1988, Cambridge: Cambridge University Press.

Thompson, N. (2019). Questions and Answers: Metaphysical Explanation and the Structure of Reality, *Journal of the American Philosophical Association 5*(1), 98–116.

Ramsey and Alien Structure*

Matti Eklund

We tend to think of the world – or the structurally most basic facts about the world – as involving objects, properties and relations (or, if you like, particulars and universals). When something is the case, that is a matter of how things stand with these objects and properties. There is much disagreement about what objects and properties there are, and there is disagreement about the nature of objects and properties in the first place. But there is widespread agreement regarding the general structure of the world: the world consists of objects, properties, relations, etc. and when something is the case it is a matter of how things stand with these entities the world consists of.

The agreement is widespread, not universal. Some theorists do not believe in such entities as properties or relations. Others don't believe in objects but think that what are normally regarded as objects are merely bundles of properties. However, these renegades all depart from orthodoxy only by *eliminating* something from the orthodox picture.

* Anna-Sofia Maurin is a deeply appreciated philosophical friend and colleague. It is a privilege to honor her through contributing to this *Festschrift*.

But might the world have positively alien structure, with features not found in the familiar picture? Might the world in some sense contain alien elements, not found in the familiar picture? These alien elements would either be alien entities in addition to familiar ones, or modes of combination of entities different from what we are otherwise familiar with. The question I am attempting to raise is admittedly vague, with the vagueness being due in part to vagueness in what counts as familiar, but some of the vagueness is hopefully remedied in the discussion to follow.

Here is an intuitive motivation for exploring the possibility of alien structure seriously. The picture of simple facts as consisting of objects and properties mirrors the subject-predicate structure of simple sentences. And it may be suspected that when we think of the world the way we do, we are simply naively mistaking what is a feature of our languages for a feature of the world. Maybe there are other languages, with different structure, which we could use to describe the world. Or, if we are somehow psychologically constrained to use representational systems with familiar semantic structure: maybe other creatures could use other languages, with different structure, to successfully describe the world. And there may be no reason for thinking that familiar languages better capture the world's structure than do these other possible languages.

Return now to the "familiar". We can think of familiar structure as structure corresponding to the structure of ordinary linguistic representations, and alien structure as corresponding to alien kinds of linguistic representations. This does not do away with unclarity in "familiar", for the explanation relies on a distinction between alien and familiar kinds of representations, but it still helps illustrate what the intended target is. In the below, an *alien language* is a language employing alien kinds of representations: expressions of alien categories, or alien modes of composition.

These issues mentioned are big issues, and the ideas I have just briefly described certainly stand in need of elucidation before they can even be sensibly discussed. My aim in this brief note is much more modest. In his celebrated article "Universals" (Ramsey 1925), F.P. Ramsey famously questioned the particular-universal distinction. What I will consider here are some relations between on the one hand the themes and theses introduced by Ramsey and on the other alien structure. I will simply bring up some considerations that come up in connection with Ramsey's text, and note how they relate or not to the possibility of alien structure in the sense above indicated.[1]

Let me begin by providing some relevant background. One main argument in "Universals" is this. Two sentences with different structure – Ramsey's examples are "Socrates is wise", and "Wisdom is a characteristic of Socrates" – can, in Ramsey's words, "assert the same fact and express the same proposition" (404). In each sentence a subject can be distinguished, but since the sentences have different subject-predicate structure, one cannot infer anything about the subject-predicate structure of the proposition they both express by considering the sentences. Indeed, he goes so far as to say that in a "sufficiently elastic language", any propositions that can be expressed by a sentence of that language can be expressed with "any of its terms" as subject (ibid.). Ramsey concludes, "Hence there is no essential distinction between the subject of a proposition and its predicate, and no fundamental classification of objects can be based upon such a distinction" (ibid.). The kind of distinction Ramsey has in mind is exemplified by, e.g., the idea that particulars are exactly the entities which can only occur as

[1] In the below, all references are to Ramsey (1925).

subjects in propositions, whereas universals can occur both as subjects and as predicates. (In the article, Ramsey associates this specific view with Johnson (403).)

For Ramsey, the most promising way to explicate the particular/universal distinction would be in terms of what entities can play what roles in propositions. He insists that disputes concerning the particular/universal distinction have to do with the "logical nature" of entities (402, 404, 405). But by arguments like that presented, he is skeptical at least about one way of distinguishing between entities on the basis of their logical nature: distinguishing between them on the basis of their ability to occur as subjects and as predicates in propositions.

I say "he is skeptical". Towards the end of the article, Ramsey pretty clearly gives voice to a general agnosticism. He says that "we know and can know nothing about the forms of atomic propositions", and that "[o]f all philosophers Wittgenstein alone has seen through the muddle and declared that about the forms of atomic propositions we can know nothing whatever" (417). However, the conclusion of the argument we have been concerned with is the more definite "there is no essential distinction between the subject of a proposition and its predicate, and no fundamental classification of objects can be based upon such a distinction" (404). There is nothing 'agnostic' or cautious about this. Now, Ramsey doesn't say here that there is no *particular/universal distinction*, but that claim follows if we assume that if there is such a distinction then it must be possible to discern subjects and predicates in at least some propositions. One way of making Ramsey consistent is by interpreting him as making a categorical negative claim – there is no subject/predicate structure in propositions – but at the same time allowing that there is some other way in which entities can play broadly speaking different logical roles in propositions. Maybe entities play different logical roles in propositions, but in a

way that does not correspond to how expressions play different semantic roles in sentences of familiar kinds of languages. If there is any kind of language at all whose sentences mirror the structure of propositions, that language would have to be an alien language. Of course, even before looking at details, one can worry that the strategy behind Ramsey's argument against a subject/predicate generalize to other distinctions between different logical roles in propositions. And if sentences of an alien language express the same propositions as sentences of a more familiar kind of language, then one can make the Ramsey-style point that we cannot determine whether the propositions have familiar or alien structure.

Whatever Ramsey's conclusions are exactly, he is in one way philosophically radical: he challenges the time-honored particular-universal distinction. (Nominalists, denying the existence of universals, in a sense deny it too, by denying the existence of universals. But even the nominalist would agree that all *putative* entities can be classified as particulars or universals.) But even so, there is another way in which he is not radical. He does not challenge the idea that reality consists of the entities we otherwise take it to consist of. He does not challenge the existence of familiar entities like Socrates and wisdom; nor does he introduce alien entities in addition to the familiar ones. He only challenges the idea that the entities there are can be reasonably divided up into particulars and universals, given the way "particular" and "universal" are properly understood. But a more radical skepticism suggested by Ramsey's arguments would be that the entities that are the real constituents of propositions are not familiar entities like Socrates and wisdom and are not in any straightforward way named by any expression of the familiar sentences we use to express them.

In the concluding section of the article, Ramsey says,

> In conclusion let us describe from this new point of view the procedure of the mathematical logician. He [sic] takes any type of objects whatever as the subject of his reasoning, and calls them individuals, meaning by that simply that he has chosen this type to reason about, though he might equally well have chosen any other type and called them individuals. The results of replacing names of these individuals in pro-positions by variables he then calls functions, irrespective of whether the constant part of the function is a name or an incomplete symbol, because this does not make any difference to the class which the function defines. (417)

(Where Ramsey here uses "object" I would use "entity", in order better to emphasize that the remarks do not specifically concern particulars but concern any entities whatsoever.)

Ramsey's "mathematical logician" is agnostic about the logical structure of reality. He holds that there are no distinctions in reality (of a "logical nature") to be concerned about between different broadly speaking logical categories of entities, for we cannot know what these possible distinctions are. All that we need to be concerned about is that there are sentences, some of them containing names - and any type of entity can be named – and the sort of thing that results from replacing a name by a variable, the "incomplete symbols" referred to. The strategy of Ramsey's mathematical logician is reasonable independently of any supposed further structure in reality and the question of whether reality might "really" contain alien structure is moot. The important thing is just that reality lends itself to the kind of representation the mathematical logician prefers.

This general broadly *instrumentalist* view, as we may call it, in turn comes in different flavors. It can be used for conservative purposes. It can in principle be used to criticize attention to alien structure:

the claim would be that anything *can* in this way be described in familiar terms, and hence we need not concern ourselves with alien structure. On the other hand, the friend of alien structure can say in response to this that there is still the question of whether it is also possible to say things using a language with different, alien kinds of expressions; and whether such an alien language might serve useful purposes. Even if any entity whatsoever can be named and picked out as subject of reasoning and even if one can use representations of a familiar kind in the way Ramsey's 'mathematical logician' does, one can raise questions of whether one can also devise alien kinds of representations for various theoretical purposes. If instrumentalism is right, then the usefulness of these representations would not indicate that reality has alien structure – the latter question is rejected – but for the same reason, the usefulness of familiar representations does not indicate that reality has familiar structure.

A possible view slightly different from the agnostic instrumentalist view mentioned is one which positively insists that there are no logical distinctions in reality to track. There are differences between how expressions occur in sentences, and that is that. The same remarks that applied to the instrumentalist view as first described apply also to this variant.

References

Ramsey, F.P. (1925) Universals, *Mind* 34, 401-417.

On Hyperintensional Metaphysics

Darragh Byrne & Naomi Thompson

A certain kind of *realism* about various rather theoretical phenomena has been gaining traction lately in metaphysics. One strand in Anna-Sofia's recent work resists one instance of this (though we note that plenty of other strands in her work pull the other way!). This is her (2019) critique of a certain *unionist realist* conception of metaphysical *grounding*. Advocates of this position (e.g. Raven 2012; and Rosen 2010) hold that grounding is explanatory by nature and thus that grounding inherits its properties from those of explanation, and yet also that grounding is an objective, *worldly* matter. Anna-Sofia argues that this combination of doctrines is unstable, because our familiar conception of explanation is mind-involving.

Daniel Nolan (2014) defends a position suggestive of an especially abstract, general instance of this topical realism: a kind of overarching realism about a whole family of theoretical phenomena. Nolan's focus is *hyperintensionality,* an especially abstract property that is (sometimes somewhat mysteriously) attributed to myriad metaphysical phenomena, e.g. properties, essence, intrinsicality, and again, grounding. Nolan's central contentions are, first, that we need hyperintensional notions to describe and explain important aspects of the world, and second, that this need cannot be accounted for in terms of 'features of our representational systems' (2014: 150). He

does not explicitly describe his position as 'realism about hyperintensionality', but his remarks are supportive of it (and in conversation, he's sympathetic). We worry that this realism about hyperintensionality exhibits an instability similar to that which Anna-Sofia diagnoses in the particular instance of realist unionism about grounding, in that it threatens to conflate a metaphysical feature with the way in which it is represented. Thus in criticising Nolan's argument below and defending the view that hyperintensionality derives from features of our representational system, we take ourselves to be siding with Anna-Sofia, even though our position is much more general than hers.

Hyperintensionality

The reason that attributions of hyperintensionality to metaphysical phenomena sometimes seem mysterious is that in the first instance at least, hyperintensionality is, like intensionality, a feature of representations: its bearers are *positions in sentences*. A position in a sentence is intensional if substituting co-referential/co-extensive occupants of that position can change the truth value of the sentence. In

(1) Alex hopes that Anna-Sofia will write another paper

the position occupied by 'Anna-Sofia' is intensional because it may be possible to change the truth value of (1) by replacing 'Anna-Sofia' in it with a co-referential term, e.g., 'the author of the *Stanford Encyclopaedia* entry on tropes'. The position occupied by 'Anna-Sofia' in (1) is also *hyperintensional* because we can change the truth value of the sentence by replacing the name with e.g. 'Professor Maurin', a term that is not only co-referential, but *necessarily* so in the sense

that 'Anna-Sofia is Professor Maurin' is not only true, but (in contrast to 'Anna-Sofia is the author of the *Stanford Encyclopaedia* entry on tropes') *necessarily* so.

We accept Nolan's first contention: hyperintensional locutions are needed to describe and explain the world. But we defend the doctrine that his second contention rejects: that hyperintensionality derives from features of representations. We defend this doctrine, which we label 'conceptualism', by adopting a broadly Fregean conception of representation.

A Fregean perspective

According to Fregeans, words don't refer *directly* to objects, properties and states-of-affairs: rather, speakers refer by associating words with their conceptions of these subject-matters – the *senses* of words and sentences – and these are what people *express* or *mean* when they speak and think. If Alex associates a certain conception of a person with the name 'Anna-Sofia', but not with the description, 'The author of the *Stanford Encyclopaedia* entry on tropes', then what he says/thinks when he utters/entertains the sentence, 'Anna-Sofia will write another paper' is distinct from what he would say/think were he to utter/entertain 'The author of the *Stanford Encyclopaedia* entry on tropes will write another paper'. We can think of Alex's hopes (and his beliefs etc.) as relations he bears to the senses expressed by sentences like these, so that as expected, (1) may be true even if the sentence we get by replacing the first of these terms in it with the second is false.

Should we regard this conception of the intensionality of (1) as realist or anti-realist? On the one hand, the account trades crucially

on the fact that there are two distinct ways of thinking about a single thing, so it seems to suggest that the intensionality of (1) results from features of ways of representing the world. So the Fregean position is a version of conceptualism – the position whose hyperintensional analogue Nolan opposes. On the other hand we're a little reluctant to label the position 'anti-realist', since on the Fregean view, senses are not ethereal, subjective or psychological. Senses are objective representational perspectives: they do not spring from the representational activities of particular human thinkers.

Next note the ease with which this account of the intensionality of sentences like (1) may be extended to deliver an account of their *hyper*intensionality. Fregeans allow that terms which necessarily co-refer (in the sense explained above) can differ in sense just as terms which only contingently co-refer can, and wherever such differences arise it will be possible to formulate sentences whose truth values are sensitive to suitable substitutions. The key ingredient in this explanation of (1)'s hyperintensionality – and what makes it *conceptualist* – is the same as in the explanation we offered of its intensionality: it's the notion of Fregean sense.

Nolan's argument

Nolan does not consider the conceptualist position offered above, but he would not be impressed by the example in terms of which we explained it, because his key move is to argue that hyperintensional notions are needed to describe and explain *non-representational* features of the world. Intentional states like *Alex's hopes* are representational phenomena so Nolan would say it's no surprise that we can explain *their* hyperintensionality in representational, conceptualist

terms. Nolan offers several examples of non-representational phenomena of whose hyperintensionality he thinks we should not expect a tenable conceptualist account: e.g. *properties, intrinsicality, essence,* and *grounding*.

To articulate a conceptualist account of the hyperintensionality of these notions one could formulate a broadly Fregean semantics of sentences about properties, essences, etc., in the expectation that those semantics would feature operators that generate hyperintensional contexts. We have not worked out the details of Fregean semantics of the locutions needed to express Nolan's non-representational metaphysical phenomena, but we're confident that such semantics could be formulated for most of them at least. (And if there are some of which there is no plausible Fregean treatment, there is dialectical space for the conceptualist to deny that the phenomena at issue are genuinely hyperintensional: we've not questioned any of Nolan's examples, but that's just for argument's sake). Of course, advocates of the Fregean approach incur various substantial commitments, some of which are at odds with other realist theses in contemporary metaphysics. E.g. Fregeans sympathetic to *grounding* probably need to think of it as an explanatory relation holding between senses/thoughts. Neo-Fregeans since McDowell (1977) and Evans (1982) maintain (in a 'Russellian' spirit) that senses can have worldly constituents, but even if that picture is embraced, this Fregean conception of grounding conflicts with the radically realist view that holds grounding to be a relation between worldly entitles which are expressly non-representational. (See e.g. Audi 2012)

Instead of embarking on the semantic project here, we want to take issue with Nolan's overarching reason for thinking it cannot succeed. This is the following conditional – implicit in our presentation above and in his paper:

(N) If a subject-matter is not representational, then we cannot explain the fact that we need hyperintensional idioms to describe and explain it in conceptualist terms.

To see that something may be wrong with (N), consider an analogous conditional about *intensionality:*

(M) If a subject-matter is not representational, then we cannot explain the fact that we need intensional idioms to describe and explain it in conceptualist terms.

To examine (M) we don't need to develop Fregean semantics for terms associated with the 'hyperintensional metaphysics' cherished by Nolan, for a Fregean analysis of the non-representational intensional operator *per excellence* is well-known.

(2) Necessarily, 2 + 3 = 5.

The subject-matter of (2) is not representational, and the position occupied e.g. by '2' in it is intensional because we can change the truth value of (2) by replacing '2' with a term that co-refers, e.g. 'the number of my ears'. (The positon is not hyperintensional because we cannot affect such a change by replacing '2' with a necessarily co-referring term, e.g. '√4'.)

The Fregean account of the intensionality of (2) is much like that of the intensionality of (1). 'Necessarily' (like 'Alex hopes that') is an intensional operator, which is to say that (like 'Anna-Sofia' in (1)) '2' in (2) refers to its sense rather than its referent. And the same is true of any term we substitute for '2': hence by replacing '2' with a co-referring term which has a different sense, we can change (2)'s truth value. What (2) says is that a certain sense/thought has the

property of being necessary, and while this is true of the thought that 2 + 3 = 5, it is not true of the thought that the number of my ears + 3 = 5.

This Fregean account of the intensionality of (2) is every bit as conceptualist as the account we offered in §2 of the intensionality of (1): it depicts intensionality as arising from features of representations (senses). Yet the subject matter of (2) is no more representational than are those of sentences about properties, essences etc. Thus, the Fregean account falsifies (M): the subject-matter of (2) is not representational, but there's a plausible (indeed, mainstream) conceptualist analysis of its intensionality. We submit that this puts a defender of (N) on the back foot. If, as just demonstrated, there's a plausible conceptualist account of an intensional locution about a non-representational phenomenon, and if, as shown in §2, plausible conceptualist accounts of intensional locutions can be extended to deliver plausible conceptualist accounts of hyperintensional ones, why should we expect these extensions to break down when the locutions are about non-representational phenomena? We suggest that the reason the Fregean explanation of the hyperintensionality of sentences about representational phenomena such as (1) is conceptualist is not, as Nolan seems to suspect, because of the representational nature of those sentences' *subject-matters:* rather it's because of the representational nature *describing* and *explaining*.

Acknowledgements

For helpful discussions of earlier versions we're grateful to Daniel Nolan and to other participants at colloquia and conferences at which we presented these ideas, in Hamburg, Bratislava, and Cardiff.

References

Audi, P. (2012) Grounding: Toward a Theory of the In-Virtue-Of Relation, *Journal of Philosophy* 109(12), 685-711.

Evans, G. (1982) *The Varieties of Reference*, Oxford: Clarendon Press.

Nolan, D. (2014) Hyperintensional Metaphysics, *Philosophical Studies* 171, 149-160.

Maurin, A-S. (2019) Grounding and Metaphysical Explanation: It's Complicated, *Philosophical Studies* 176(6), 1573-1594.

McDowell, J. (1977) On the Sense and Reference of a Proper Name, *Mind* 86, 159-85.

Raven, M. (2012) In Defence of Ground, *Australasian Journal of Philosophy* 90(4), 687–701.

Rosen, M. (2010) Metaphysical Dependence: Grounding and Reduction, in Hale, B. & Hoffmann, A. (eds.), *Modality: Metaphysics, Logic, and Epistemology*, Oxford: Oxford University Press.

Maurin on Grounding and Explanation

Alexander Skiles & Kelly Trogdon

In the classical and contemporary literature on grounding, explanatory language is routinely used to communicate what it is and to motivate substantive principles about how it behaves. For an example of the first use, one might say that what explains why a conference is taking place are certain intentional attitudes and activities of its participants that ground it, where the sense of 'explains' operative here is in some sense constitutive rather than causal (Dasgupta 2014: 1). For an example of the second use, one might argue that since a fact cannot explain itself, neither can a fact ground itself (Raven 2013: 193). But what is the exact relationship between grounding and explanation?

Two views have emerged about how to answer this question. Some stipulate that by 'grounding' they mean a distinctive form of determination, what we will call *determination*$_G$, where to determine is, roughly speaking, to *produce* or *bring about* (Audi 2012; Schaffer 2016; and Trogdon 2013). In this case to say, for example, that the brittleness of the bowl is grounded by the covalent bonds of the bowl's constituent atoms is to say that the bonding of the atoms produces or brings about the brittleness of the bowl. Others stipulate

that by 'grounding' they mean a distinctive form of explanation, what we will call *explanation*$_G$ (Dasgupta 2017; Litland 2015; and Rosen 2010). In this case, to say that the brittleness of the bowl is grounded by the ionic bonds of the atoms is to say that the bowl is brittle *because* the bonding of the atoms is ionic.

As is customary, let us call the latter view *Unionism* (grounding is explanation$_G$) and the former view *Separatism* (grounding is determination$_G$). Compatible with Separatism is the idea that there are conditions under which grounding backs or underwrites explanations, and Unionism is compatible with the view that there are conditions under which grounding itself is backed or underwritten by other relations.

In "Grounding and Explanation: It's Complicated" (2019), Anna-Sofia Maurin aims to show that, despite appearances, Unionism and Separatism in fact undermine the use that explanatory language has been put to in elucidating grounding, rather than undergird it. In what follows we will critically assess her interesting argument.

As Maurin notes, and as we indicated before, some structure their theorizing about grounding in part around the following two principles:

> *Explanation:* we have reason to think that explanation has thus-and-so features.
>
> *Inference:* if we have reason to think that explanation has thus-and-so features, we thereby have reason to believe that grounding has those features.

Let *Inheritance* be the conjunction of Explanation and Inference. For the purposes of our discussion, let us table three issues. First, there

is the matter of *which* features acknowledged by Explanation can be legitimately fed through Inference. Presumably not all of them, obviously so if one is a Separatist. Second, there is the matter of *which cases and for what agents* Inference applies. Presumably that question would require an appeal to broader epistemological considerations before it could be fully resolved.[1] Third, there are alternative approaches to theorizing about grounding that do not (or do not obviously) appeal to Inheritance. For instance, Audi's (2012) point of departure concerns features of *determination* (considered as a genus, with determination$_G$ a species) rather than explanation—he claims, for instance, that since determination is non-monotonic, so too is grounding, given Separatism. For another instance, Kovacs (2018) argues that, given either Unionism or Separatism, grounding plays the theoretical roles normally assigned to it only if it has the features that we normally assign to it. Nonetheless, Maurin is surely correct about Inheritance's widespread appeal, so her case against it is of great interest even if some do without it.

In brief, Maurin's central contention is that our theorizing about grounding shouldn't be guided by Inheritance—regardless of whether one endorses Unionism or Separatism. Her argument in a nutshell is this:

1. Either Unionism or Separatism is true.
2. If Unionism is true, then explanation in the relevant sense (explanation$_G$) isn't epistemically constrained.
3. If explanation in the relevant sense isn't epistemically constrained, then Explanation is implausible.

[1] We do, however, later address related issues in considering different ways in which grounding and explanation might be epistemically constrained.

4. If Separatism is true, then, while explanation in the relevant sense is epistemically constrained, grounding in the relevant sense (determination$_G$) isn't.

5. If explanation in the relevant sense is epistemically constrained, while grounding in the relevant sense isn't, Inference is implausible.

6. Hence, Inheritance is implausible.

The argument only explicitly undermines the *second* of the two uses that explanatory language has been put to in elucidating grounding that we canvassed at the start. But the *first* use seems to be threatened as well. Consider Dasgupta's attempt to communicate what grounding is, mentioned above. Presumably, implicit appeal is made to Inheritance here as well: Explanation is appealed to when it is said that the relation that the conference's taking place stands to certain facts about its participants is a relation of explanation$_G$, and Inference is appealed to when one then infers that this is also a case of grounding. But if Maurin's argument is sound, using Inheritance to glean information about grounding's non-structural features seems no better off than using Inheritance to glean information about its structural features.

Let us grant (1) in Maurin's argument. And clearly, the argument is logically valid. So in the remainder of this paper, we will critically assess Maurin's defense of (2)–(5).

Start with (2)—if Unionism is true, then explanation in the relevant sense (explanation$_G$) isn't epistemically constrained. The rationale of this premise seems to be this. Suppose that Unionism is true. If so, then explanation$_G$ and grounding are one and the same. But that would seem to imply that explanation$_G$ is not epistemically constrained. For if explanation$_G$ were epistemically constrained, then grounding would be too—and that would appear to make it an

at least partially non-objective matter which facts are grounded by which, and this is ruled out by the "inflated" notion of grounding at issue in Maurin's discussion. Here grounding is understood to be "an objective and mind-independently obtaining hyperintensional and non-monotonic strict partial ordering relation" (p. 1574).

In reply, note that there are at least two ways in which grounding may be said to be epistemically constrained. First, say that a relation, *R*, is *strongly* epistemically constrained just in case whether a certain *R*-relationship holds on a particular occasion is constrained by what actual subjects know or don't know, their cognitive capacities, and so forth on that particular occasion. Maurin seems to have this type of epistemic constraint in mind when, for example, she writes that explanation "is a function of the needs, knowledge, and expectations of those to whom the explanation is offered" (p. 1580). In this case, some facts ground another fact only if particular subjects are epistemically related to this collection of facts in the appropriate way.

Second, say that relation, *R*, is *weakly* epistemically constrained just in case whether a certain *R*-relationship holds is constrained by general epistemic facts (roughly, epistemic facts that concern hypothetical, idealized agents). For example, you might think that some facts ground another fact only if questions about why the latter should obtain given that the former obtain lack substantive content for any individual fully informed about the natures of the entities these facts involve (Trogdon 2013). Or you might think that some facts ground another only if there are conditions under which the right sorts of subjects are in a position to understand in the right way why it is that the grounded fact obtains given that its grounds obtain. Maurin seems to have something closer to this type of epistemic constraint in mind when she speaks of how explanation must bear on "the understanding of its (potential) receiver," implying that it need not be an actual one (p. 1580).

Suppose it is granted that grounding being epistemically constrained in the first, stronger way is incompatible with grounding being objective.[2] Nonetheless, grounding being epistemically constrained in the second, weaker way is compatible with its objectivity. So if grounding is weakly yet not strongly epistemically constrained, then it seems that the rationale for (2) is off the mark.

Maurin might respond to our criticism of (2) by claiming that it is not only a *necessary condition* on explanation in general—and thus explanation$_G$ in particular—that certain epistemic constrains be met; it is also *part of the essential nature* of explanation itself that this be so. On some views, this would be enough for explanation$_G$ to be less-than-fully objective (cf. Jenkins 2005). But if unionism is true, grounding just *is* explanation$_G$; thus it would appear to follow that grounding would be essentially constrained by epistemic factors, and thus less-than-fully objective too.

We have two replies to this type of response. The first appeals to the distinction between what is part of the nature of a plurality of things vs. what is part of the nature of any individual amongst this plurality. As an example, although it is plausible to think that it is essential to Socrates and Plato *taken together* that they be distinct, it is implausible to think that it is essential to either Socrates or Plato *alone* that they be distinct, given that it is implausible to think that there are any essential truths about the one that concern the other (cf. Fine 1994: 54).

With this distinction in tow, one might respond as follows. It is *not* part of the nature of explanation *alone* that, say, the right sorts

[2] Generally speaking, though, it is not. Let [p] be the fact that Kelly knows that 2+2=4, and let [q] be the fact that Alex knows that 5+7=12. Now, [p] and [q] together ground [p & q]. Yet surely one can hold that this grounding relationship is fully objective, even if its holding requires the obtaining of certain facts about what particular subjects know.

of subjects must be in a position to understand in the right way why the explanandum obtains given that the explanans obtains. Rather, this is an essential truth about explanation and understanding *taken together*. We see no reason why one cannot make the second claim without making the first. But if so, that alone is no threat to explanation, and thus grounding, being fully objective. After all, it is also an essential truth about understanding taken together with whatever fully objective phenomenon one chooses—say, photosynthesis—that they are distinct. But the fact that understanding figures into an essential truth about understanding and photosynthesis is obviously no reason for believing that photosynthesis is less-than-fully objective. Similarly, we say, in the case at hand.

Our second response instead appeals to the thought that the broadly Aristotelian notion of essence at issue here is itself a notion of explanation: what is part of the nature of something in some sense helps to explain what it is to be that very thing (cf. Fine 2015). Suppose that thought is correct. Then even if it is conceded that explanation$_G$ is essentially linked to certain epistemic factors, it does not follow that grounding is too, *even if* grounding just is explanation$_G$. The reason is that explanatory language is generally agreed to be the sort that generates opaque contexts, since whether an explanation statement is true is sensitive to the way in which it represents the explanans and explanandum (cf. Ruben 1990: 219). Yet the view under consideration is precisely that the expression "… is essentially linked to such-and-such epistemic factors" is sensitive to the sorts of factors that render explanatory contexts opaque in general. Hence, if the "…" position in this expression is opaque, one cannot validly infer that grounding is essentially linked to such-and-such epistemic factors even if explanation$_G$ is and Unionism is true. To do so would be to invalidly apply Leibniz's Law within an opaque context.

Let us now turn to (3)—if explanation in the relevant sense isn't epistemically constrained, then Explanation is implausible. (To say that Explanation in this context is implausible is to say that any interesting instance of Explanation, e.g. explanation is irreflexive, is unmotivated.) According to Maurin, explanation in the ordinary sense is epistemically constrained. Maurin writes, "…when we let the properties of explanation guide us in our characterization of grounding, our [judgements] arguably derive from intuitions formed based on our encounters with 'normal' explanation," where normal explanation is epistemically constrained (p. 1581). So the thought is that when we say that we have reason to think that explanation has thus-and-so features (i.e. when we say that Explanation is true), we're appealing to the ordinary sense of explanation, a notion that we have a good grip on. By contrast, any conception of explanation according to which it isn't epistemically constrained is obscure. Here Maurin agrees with Thompson, who claims that when we strip away the epistemic features of explanation, we "lose our grasp" of what explanation is (2016: 397).

In reply, it is true that, given Unionism, the objectivity of grounding ensures that explanation in the relevant sense isn't constrained by what particular subjects already know or don't know, their cognitive capacities, and so on. But, given our discussion above, explanation in the relevant sense being tied to general epistemic facts is compatible with explanation being objective.

Now, perhaps Maurin thinks that if there is nothing more to explanation in the relevant sense being epistemically constrained than it being tied to general epistemic facts like those described above, then the relevant notion of explanation, unlike the ordinary notion, is obscure. In that case we don't really know what features explanation in the relevant sense has—Explanation is implausible.

We're willing to grant for the sake of argument that the conception of explanation at issue here might not capture the ordinary notion of explanation. This is compatible, however, with this conception of explanation figuring in ordinary thinking. And, even if it doesn't, we don't find the notion obscure. Compare: while what is *conceivable* is arguably epistemically constrained, a conception of conceivability that understands this dimension of conceivability solely in terms of general epistemic facts (e.g. ideal rational reflection not detecting contradictions) isn't obscure (Chalmers 2002). The moral: once we get clear on what Unionism rules out with respect to the potential epistemic dimension of explanation and what it doesn't, for all we have been told Explanation is in good standing.

It's worth noting that in considering Strevens' (2008, ch. 3) discussion of the connection between explanation and understanding, Maurin briefly considers an approach to explanation that resembles in certain ways the view that explanation is weakly epistemically constrained (i.e. explanatory relationships are tied to general epistemic facts). Maurin goes on to claim that, given Separatism, explanation so understood is sufficiently different from grounding so as to undermine Inference. We will critically assess this claim below. But note that in the present context this claim is neither here nor there, as we are considering Unionism and the idea that grounding is identical to a form of weakly epistemically constrained explanation.

Let us next turn to (4)—if Separatism is true, then, while explanation in the relevant sense is epistemically constrained, grounding in the relevant sense (determination$_G$) isn't. We have already considered Maurin's rationale for the claim that grounding isn't epistemically constrained—it's having this feature is incompatible with the "inflated" conception of grounding according to which grounding is objective. And moreover, Maurin notes that, while it's possible for

the Separatist to say that explanation in the relevant sense isn't epistemically constrained, the package of views consisting of the inflated notion of grounding combined with an epistemically unconstrained notion of explanation is a "raw deal," and it is "…unclear why one would want to hold this view" (p. 1578).

At this juncture, what was said above with regard to (2) applies to (4) as well. We have already seen that the inflated notion of grounding is compatible with grounding being only weakly epistemically constrained, i.e. only by general epistemic facts. And we have also already seen that we can maintain that explanation is weakly epistemically constrained without it being strongly constrained, i.e. by what particular subjects know or don't know, their cognitive capacities, and so on.

We would like, however, to focus here on a different idea. We think that there are viable approaches to explanation according to which explanation isn't epistemically constrained, even in the sense of being tied to general epistemic facts. One such approach combines a version of the so-called ontic view of explanation with a more complex view of *good* explanation: while epistemic matters aren't relevant to what makes something an explanation, they are relevant to what makes something a *good* explanation. More specifically, the view we have in mind is this: while explanations must satisfy certain ontic constraints—roughly, they need to represent the right stuff in the world—good explanations must satisfy certain ontic *and* epistemic constraints—roughly, they need to represent the right stuff in the world in the right way. Compatible with this proposal is the idea that one of these constraints is more fundamental than the other regarding what it takes to be good explanation—one might think,

for example, that the epistemic constraint is met only if the ontic constraint is met but not vice versa.[3]

Let us finally turn to (5)—if explanation in the relevant sense is epistemically constrained while grounding in the relevant sense isn't, Inference is implausible. (To say that Inference is implausible in this context is to say that the sort of inferences at issue with Inference, e.g. since explanation is irreflexive, so too is grounding, are unlicensed.) What is the rationale behind this claim? Maurin writes:

> … as part of what it is to *be* an explanation is to be this mind-dependent and epistemic thing, why think that explanation having the properties it does, justifies our thinking that those are the properties had by worldly and mind-independent grounding? No good reason comes to mind. (pp. 1578–9, emphasis in original).

Maurin's rationale seems to be that if grounding is fully objective, yet explanation$_G$ is not, then it is implausible to infer that grounding has certain features given that explanation$_G$ has certain features.

We have two replies, the first of which is to reiterate again what we said before. If explanation$_G$ is merely *weakly* epistemically constrained—and no reason has been provided for believing it to be otherwise—then explanation$_G$ may well be fully objective. And if that is so, then there is no obvious reason why one cannot be justified in inferring that a fully objective relation of grounding has certain features given the fact that explanation$_G$ does.

Our second reply is to challenge Maurin's contention that it is implausible to infer that a fully objective phenomenon has certain

[3] See Illari (2013) for a defense of the claim that good mechanistic explanations in particular must satisfy both ontic and epistemic constraints, as well as discussion of the priority issue. And see Krämer & Roski (2017) for discussion of good explanation that appeal to grounding, which they understand as having an ontic constraint that concerns difference-making in particular.

features from the fact that some less-than-fully objective phenomenon does. Suppose that Naomi is having an experience as of a cube (however that is spelt out: something looks like a cube in Naomi's environment, Naomi is being appeared to cube-ly, etc.). The existence and nature of this experience (i.e. that Naomi is having this experience and what it's like for her to have it) is, of course, at least a partially mind-dependent matter. Yet it seems that having this experience gives Naomi at least some reason to believe something in her environment is cubical. (Which is compatible with that reason being *defeasible*, and with her belief being *false*.) Yet whether there is something cubical in Naomi's environment is presumably a fully mind-independent matter. Hence, assuming that a general skepticism about perceptual experience can be set aside, we see no immediate reason to accept Maurin's general prohibition against making inductive inferences from features of less-than-fully objective phenomena to features of fully objective phenomena.

Let's wrap up. Maurin has argued that regardless of what one thinks about the relationships between grounding and explanation, one can learn little if anything about the features of the former by appealing to the latter. We have replied by attempting to clarify the ways in which grounding and explanation are related to various other notions of perennial metaphysical interest, chief among them the notions of objectivity and essence. We tentatively side with the status quo on the theoretical usefulness of Inheritance.

We expect that Maurin's paper will play an important role in shaping future work in this area (we know that it will in our own case). The paper, like her other contributions to metaphysics, is insightful, written with verve and humor, and rewarding to engage with.

References

Audi, P. (2012) Grounding: Toward a Theory of the In-Virtue-Of Relation, *Journal of Philosophy* 109, 685–711.

Chalmers, D. (2002) Does Conceivability Entail Possibility? In T. Gendler and J. Hawthorne (eds.), *Conceivability and Possibility*. Oxford: Oxford University Press.

Dasgupta, S. (2014) The Possibility of Physicalism, *The Journal of Philosophy* 111, 557– 592.

—— (2017) Constitutive Explanation, *Philosophical Issues* 27, 74–97.

Fine, K. (1994) Senses of Essence. In W. Sinnott-Armstrong, D. Raffman, & N. Asher (eds.), *Modality, Morality, and Belief: Essays in Honor of Ruth Barcan Marcus*. Cambridge: Cambridge University Press.

—— (2015) Unified Foundations for Essence and Ground, *Journal of the American Philosophical Association* 1, 296-311.

Illari, P. (2013) Mechanistic Explanation: Integrating the Ontic and Epistemic, *Erkenntnis* 78, 237–255.

Jenkins, C. S. (2005) Realism and Independence, *American Philosophical Quarterly* 42, 199–209.

Kovacs, D. (2018) What is Wrong with Self-Grounding? *Erkenntnis* 83, 1157–1180

Krämer, S. & S. Roski. (2017) Difference-Making Grounds, *Philosophical Studies* 174, 1191–1215.

Litland, J. (2015) Grounding, Explanation, and the Limit of Internality, *Philosophical Review* 124, 481–532.

Maurin, A.-S. (2019) Grounding and Explanation: It's Complicated, *Philosophical Studies* 176, 1573–1594.

Raven, M. (2013) Is Ground a Strict Partial Order?, *American Philosophical Quarterly* 50, 193-201

Rosen, G. (2010) Metaphysical Dependence: Grounding and Reduction. In R. Hale and A. Hoffman (eds.) *Modality: Metaphysics, Logic, and Epistemology*. Oxford: Oxford University Press.

Ruben, D.-H. (1990) *Explaining Explanation*. London: Routledge.

Schaffer, J. (2016) Grounding in the Image of Causation, *Philosophical Studies* 173, 49–100.

Strevens, M. (2008) *Depth: An Account of Scientific Explanation*. Cambridge, MA: Harvard University Press.

Thompson, N. (2016) Grounding and Metaphysical Explanation, *Proceedings of the Aristotelian Society* 116, 395-402.

Trogdon, K. (2013) An Introduction to Grounding. In B. Schnieder, M. Hoeltje, & A. Steinberg (eds.), *Varieties of Dependence*. Munich: Philosophia Verlag.

About the Authors

Johan Brännmark, Associate Professor in Ethics and Political Philosophy, Malmö University, Sweden

Darragh Byrne, Lecturer in Philosophy, University of Birmingham, UK

Einar Duenger Bohn, Professor in Philosophy, University of Agder, Norway

Matti Eklund, Professor in Theoretical Philosophy, Uppsala University, Sweden

Tobias Hansson Wahlberg, Associate Professor in Theoretical Philosophy, Lund University, Sweden

Rögnvaldur D. Ingthorsson, Researcher, Theoretical Philosophy, Lund University, Sweden

Ingvar Johansson, Professor Emeritus, Department of Historical, Philosophical and Religious Studies, Umeå University, Sweden

Martin L. Jönsson, Associate Professor in Theoretical Philosophy, Lund University, Sweden

Johannes Persson, Professor in Theoretical Philosophy, Lund University, Sweden

Björn Petersson, Associate Professor in Practical Philosophy, Lund University, Sweden

Nils-Eric Sahlin, Professor in Medical Ethics, Lund University, Sweden

Peter Simons, Professor Emeritus, Trinity College Dublin, Republic of Ireland

Ylwa Sjölin Wirling, Postdoctoral Researcher, Royal Institute of Technology in Stockholm, Sweden

Alexander Skiles, Postdoctoral Researcher, Massachusetts Institute of Technology, USA, and University of Gothenburg, Sweden

Jeroen Smid, Postdoctoral Researcher, University of Manchester, UK, and Lund University, Sweden

Robin Stenwall, Senior Lecturer in Theoretical Philosophy, Lund University, Sweden, and Researcher, University of Gothenburg, Sweden

Fredrik Stjernberg, Professor in Philosophy, Linköping University, Sweden

Naomi Thompson, Lecturer in Philosophy, University of Southampton, UK, and Researcher, University of Gothenburg, Sweden

Kelly Trogdon, Associate Professor, Virginia Tech, USA

Lena Wahlberg, Associate Professor in Jurisprudence, Lund University, Sweden

Festschrift

This book is in honour of Professor Anna-Sofia Maurin on her 50th birthday. It consists of eighteen essays on metaphysical issues written by Swedish and international scholars.

LUND
UNIVERSITY