

Wasserman, R., *Paradoxes of Time Travel*.
Oxford: Oxford University Press, 2017, pp. ii + 224.

Wasserman's book is a comprehensive overview and survey of the literature on time travel. Most of, if not all, the different models of time travel are carefully introduced and discussed to assess their pros and cons. Moreover, the philosophical tools needed to understand the debate are introduced before the problems are addressed. Thus, the book will prove to be useful both for people willing to approach this topic for the first time and for more advanced scholars that already work on specific subtopics of time travel but maybe lack the general picture or the (recent) history of the problem. The book offers expansions on the existing literature as well. For instance, the hyper-time model of time travel, which Wasserman seems to favour, is here better developed than it was in his previous work with Hudson (2010).¹ Throughout all the book, the arguments are always well laid out, and each assumption is clearly highlighted. Moreover, the book comes with a lot of figures that help the reader understand what is going on with the example discussed.

The book's main question is whether or not time travel is possible. As it is well known, there are several different kinds of possibilities. For instance, something can be logically possible (in accordance with the laws of logic), physically possible (in accordance with the laws of nature of a given world), technologically possible (permitted by the current level of technology), and many more. The kind of possibility Wasserman is interested in is the metaphysical one. He adopts a straightforward account of metaphysical possibility, according to which something is metaphysically possible only if it is permitted by the laws of metaphysics. What is a law of metaphysics is then something that needs clarification. According to Wasserman, a law of metaphysics is a suitably general generalization about *what grounds what*. The "what" ranges over everything, and the relevant relation at play is the non-causal grounding relation, which imposes a hierarchical structure on reality. Typically, a law of metaphysics will provide a universal generalization in terms of necessary and sufficient conditions, specifying also the direction of the grounding relation. An example of a law of metaphysics candidate might help here. For instance, one might believe that an event C causes an event E if and only if, *and in that case because*, events similar to C are always followed in time by events similar to E. If so, that person would believe in a metaphysical law about causation. The law says something general about causation, and it explains why something causes something else; causal relations are grounded on the relation of precedence of events of similar kind. This account of metaphysical possibility gives Wasserman a way to homogeneously treat the several paradoxes time travel raises. That is, typically the arguments discussed in the book are those *against* the possibility of time travel, whose form is the following. If time travel were metaphysically possible, then something else would be possible. But that something else is metaphysically *impossible* (forbidden by the laws of metaphysics), hence time travel is metaphysically impossible. For instance, backward time travel would make backward causation possible, i.e. an effect

¹ Hudson, H. and Wasserman, R. 2010, "Van Inwagen on Time Travel and Changing the Past", in Zimmerman, D. (ed.), *Oxford Studies in Metaphysics*, 5, Oxford: Oxford University Press, 41-49.

preceding its cause in time. The time traveller pushes her time machine's button, and earlier in time the machine appears out of thin air at the time traveller's desired destination. Pushing the button clearly causes the event of the machine appearing out of thin air in the past, and the latter event comes earlier in time than the former one. But one might think that backward causation is metaphysically impossible. For instance, the law of metaphysics about causation mentioned above rules out backward causation because according to it, causes always come earlier than their effects. If we had conclusive reasons to think that that law of metaphysics is correct, then backward time travel would also be ruled out by the realm of metaphysical possibilities. Of course, arguments of this form can be resisted in at least two ways. One might adopt different laws of metaphysics, such that they allow for what comes together with time travel, and independently motivate them. Or it can be said that what is allowed by time travel is not, even if it might seem so, in contradiction with the laws of metaphysics. Refuting these arguments in ways such as these against the metaphysical possibility of time travel is what Wasserman does for most of the book, drawing from the existing literature and expanding on it, eventually concluding that time travel *is* metaphysically possible.

Into the book structure: Chapter 1 provides some examples of time travel and a definition of it. Chapter 2 offers a brief overview of the main metaphysical theories of time and shows how they behave when it comes to time travel. Chapter 3 mostly has to do with the paradoxes of past-alterations. In short, what is paradoxical is that it seems possible for time travellers to perform actions such that, if performed, would bring about changes in the past, whereas there are strong reasons to think that changing the past is impossible. Wasserman thinks that two-dimensional models of time travel, models that feature hyper-time as a second dimension of time, offer the means to coherently overcome this difficulty. In such models, it is argued, time travellers can change the past by making the same past-time feature different events relative to different hyper-times. Chapter 4 digs more into the paradoxes raised by the time traveller's abilities with respect to possible changes in the past. Traditionally, this debate had a lot to do with issues related to free will because if the past cannot be changed, then the only things time travellers can do are the ones they have already done, leaving no room for the power to do otherwise. However, the same paradoxes can be raised through thought experiments involving machines instead of human time travellers. Wasserman's position on this is that both the case involving humans and the cases involving objects should receive the same treatment. Chapter 5 deals with causal paradoxes that pose a threat to the possibility of time travel. Some are: the causal loop paradox (where a causal loop is a series of events e_1, e_2, \dots, e_n , such that each event causes the next one, and e_n causes e_1), the *ex nihilo* paradox (an object involved in a causal loop may lack an origin), and the restoration paradox (an object in a causal loop has to be somehow restored to its initial condition before it gets to the beginning of the causal loop). The last chapter deals with the Paradoxes of Identity and Self-Visitation.

Overall, Wasserman's book on the paradoxes of time travel makes an excellent read, mostly for how it carefully goes through the literature and thoroughly addresses the arguments in favour and against the metaphysical possibility of time travel. On a final remark, I want to highlight two possible worries about the book content. The first worry has to do with the definition of

time travel Wasserman adopts (Chapter 1). The traditional definition of time travel comes from Lewis (1976).² The Lewisian definition says that we have time travel whenever there is a discrepancy between external time, i.e. time itself, and the time traveller's personal time. Personal time is what orders the time-traveller's stages that are scattered throughout external time. This ordering is made on the basis of the regularities which normally hold in the world the time-traveller inhabits. It is a regularity in our world that, say, we first eat pizza and then digest it. Consider a time-traveller in our world who has a slice of pizza for lunch, time travels 40 years back, and then digests that slice 1 hour later. Given the regularity just mentioned, the time-traveller eating-stage is earlier than the digesting-stage according to her personal time, even though the latter is earlier than the former according to external time. This discrepancy is what qualifies this case as a time travel case. However, Wasserman observes, this definition is wanting when we try to apply it to objects. Wasserman makes us consider an electron that never changes its intrinsic properties. Given that we have a bunch of stages of the electron always intrinsically identical, we are free to assign to those stages several personal times to the electron that preserve the regularities we normally observe, namely the electron never changing its intrinsic properties. Some of those personal times, actually all but one, would generate a mismatch with external time and hence the definition would (wrongly) predict that the electron travels in time. Wasserman tries to overcome this problem by arguing that the correct account of personal time that must enter the Lewisian definition in terms of discrepancy has to be causal. Causal dependence, Wasserman argues, seems to be the relation that correctly orders the stages of a time traveller, whether it be a person, a particle, or everything else (8). However, this proposed definition can be challenged by considering a little variation of the electron case. Say that the intrinsically unchanging electron does time travel. At some point, it enters a wormhole and it appears back in the past. No doubt that in such case the electron time travels. Yet, Wasserman's causal definition does not predict so. In fact, it is plausible to think that causal relations among the stages of an object necessarily bring about some kind of intrinsic change in that object. But the electron never undergoes any kind of intrinsic change, and hence there isn't any causal relation among its stages. Therefore, if the personal time of an object is the ordering provided by the causal relation among its stages, we don't even have personal time for the electron here, and hence no mismatch between its personal time and external time. Notice that we cannot appeal to causes external to the electron to fix this problem, for when we are looking for discrepancies between external time and personal time of an object, what happens externally to that object contributes to making up the ordering of external time.

The second worry has to do with the hyper-time models of time travel Wasserman discusses. I will here focus on the A-theoretical hyper-time model coupled with eternalism (presented in Chapter 3, Section 4.3). According to this model, Reality comprises an eternalist block that stretches out across time. The further temporal dimension, hyper-time, is the temporal dimension across which an objective passage of the Now over the temporal slices of the block takes place. Wasserman argues that this model can provide genuine cases of past

² Lewis, D.K. 1976, "The Paradoxes of Time Travel", *American Philosophical Quarterly*, 13, 2, 145-52.

alterations by means of time travel. Here is how. Say that at the moment of hyper-time ht1 the eternalist block features Tim, a would-be time traveller pressing the button of his time machine at t100. At ht1, Tim is unsuccessful because the block does not feature his desired appearance out of thin air at t50. But when the Now shines over t100, say that it happens at ht100, the block changes and it hyper-now features Tim's relocation in the past. The past year t50 changes from not containing Tim relative to ht99 to containing him relative to ht100. Given that in hyper-time models time has hyper-temporal parts, Wasserman argues that this is a genuine case of past-alteration because Tim really arrives to the past time t50, a time at which he has never been around before. However, there is still room for arguing that this is not a case of past alteration. For once hyper-time enters the picture, a location in time is no longer individuated by a single time-coordinate. Rather, it is individuated by a pair of time/hyper-time coordinates. And if so, no time-location so construed can ever change because a pair time/hyper-time can never change from containing an event to not containing it, or else we would have a contradiction. What happens is that the time-location t50/ht99 eternally does not feature Tim, whereas the time-location t50/ht100 eternally features him.³

*University of Tyumen, School of Advanced Studies
(Russian Federation)*

GIACOMO ANDREOLETTI⁴

³ See Baron, S. 2015, "Back to the Unchanging Past", *Pacific Philosophical Quarterly*, 98, 1, 129-47 for considerations along these lines.

⁴ I would like to thank Samuele Iaquinto, Giuliano Torrengo and an anonymous referee for comments on an earlier version of this review.

Turbanti, G., *Robert Brandom's Normative Inferentialism*.
Amsterdam: John Benjamins, 2017, pp. v + 245.

Inferentialism is a house with many mansions, and usually the books devoted to it just try to sketch the landscape in order to focus on more specific subjects. Turbanti's book defines a different, more ambitious, project: the author attempts a wide reconstruction of Robert Brandom's inferentialism that takes advantage of a more mature phase of reception and discussion. In particular, Turbanti tried to figure out how the main Brandomian commitments hang together, something scholars in the field judged premature until now, and for a time to come. This means that Brandom's inferentialism is here understood in its wider connection with the recent project called 'analytic pragmatism' (hereafter AP), developed in the book *Between Saying and Doing* (thereafter BSD), and also with Brandom's still unpublished work on Hegel.¹ Even though Turbanti, in his introduction, describes the scope of his book as 'narrow' and low profile—because of its focus *just* on inferentialism—as a matter of fact, it is not a narrow scope at all. So, this is not the typical book about inferentialism for many reasons. In particular, it is noteworthy and important for a number of rather unique features. But let us first

¹ Brandom, R.B. 2008, *Between Saying and Doing*, Oxford: Oxford University Press; Brandom's reading of Hegel's *Phenomenology of Spirit* can be found in *A Spirit of Trust*, online: http://www.pitt.edu/~brandom/spirit_of_trust_2014.html

take a look at the structure of the book where such features emerge as properly put in context.

Chapter 1 introduces in general the main themes presented, the claims advanced, and the challenges undertaken by this book. A first important feature comes from the recent work by Jaroslav Peregrin—with its distinction of inferentialism, as a fundamentally ‘normative’ approach to conceptual content and discursive practice—from the views that come under the heading *inferential role semantics*, understood as rather ‘causal’ accounts.² Inferentialism, according to this distinction, is an account that focuses on which inferences a speaker *ought* to draw to participate in a discursive practice where the performances of speakers are liable to be assessed by other speakers; according to (causal) inferential role semantics, the inferences that determine the content of a linguistic expression (or of a thought) are those that speakers are disposed/caused to draw (6). From this point of view, the distinction is fundamental in clarifying how deeply Brandom’s project differs from a number of accounts proposing explanatory views based on inferential role. This is particularly useful when it comes to reconstructing the historical and argumentative genealogy of normative inferentialism. This presentation sharply distinguishes an argumentative path towards normative inferentialism that goes from Frege to Brandom, and rules out those figures which are fundamental in shaping the causal accounts, but whose contribution is not directly significant for the Brandomian project. It helps also to distinguish and isolate the problems which are genuine for normative inferentialism from those that, coming from the *causal* field, can make the dialectics spurious and lead to certain conflation. In this very context, a similar point is made to better distinguish normative inferentialism from its relatives in the field of proof-theoretic semantics (7).

Chapter 2 introduces Brandom’s normative pragmatics, the conception of discursive practice as governed by “the game of giving and asking for reasons”. Turbanti here starts with an explicit account of what Brandom calls *sapience*: the idea that human cognition and intentionality depend essentially on the use of concepts and on the participation in normative practices.³ This account of sapience is also a nice introduction to Sellars’ criticism to the Myth of the Given: accordingly, perceptual episodes alone are not sufficient to ground perceptual knowledge, thus acknowledging a crucial difference between ‘responsiveness’ (that characterizes such episodes) and ‘contentfulness’ (characterizing knowledge). Furthermore, the chapter explores the main pillars of Brandom’s pragmatics: the normative notions of ‘commitment’ and ‘entitlement’, and the basic understanding of discursive practitioners as deontic scorekeepers. Moreover, the chapter deals with ‘normative phenomenalism’—the idea that normative statuses of speakers depend on their normative attitudes—here introduced in great depth, and this presentation, that introduces and faces the main challenges in the debate, is arguably the best in the literature. Again, another aspect of interest is the pragmatics-semantics interface, that is presented with a detailed analysis. Here, the reader can also appreciate the direct contrast between Brandom’s pragmatics and mainstream literature in cognitive pragmatics. In particular, Turbanti does a good job in emphasizing how

² Peregrin, J. 2014, *Inferentialism. Why Rules Matter*, Basingstoke: Palgrave Macmillan.

³ This idea of beginning with sapience is shared with Jeremy Wanderer’s book. Cf. Wanderer, J. 2008, *Robert Brandom*, Stokesfield: Acumen.

Brandom's perspective is capable of putting serious pressure to the very foundations of the cognitive approach. For example, Brandom's insights here are relevant in questioning the explanatory role that the notion of 'speaker's intention' plays in theories like Grice's and its developments: namely, these accounts just presuppose the contentfulness of intentional states (48). Finally, the chapter presents and discusses the challenge of the so-called 'declarative fallacy' with which Rebecca Kukla and Mark Lance, although from a rather sympathetic perspective, addressed Brandom's pragmatics: the idea of an unjustified prominence of assertive speech acts over other types (49-59).⁴

Chapter 3 deals with the inferentialist account of conceptual content, a "semantic theory that represents linguistic contents in terms of inferences" (61). Here the reconstruction is wide, and the contributions of the 'founding fathers' of inferentialism (Frege, Sellars, Dummett) are presented with great detail. A noteworthy feature is the *negative narrative* that Turbanti employs in order to present and discuss the typical lessons of inferentialism. The author skilfully introduces inferentialist insights and solutions in the philosophy of language by presenting in detail the problems of alternative accounts as the basic motivations for introducing typical Brandomian (and Sellarsian) points. These narratives often make Turbanti's presentations and discussions of these insights wider and more complex than those proposed by Brandom himself; they not only deal with theoretical details and dialectics, but are often enriched with historical perspective. For example, the book presents inferentialism by starting with a taxonomy of the problems of the nominalist conceptions of meaning—the idea that all linguistic expressions work like names. In this context, Turbanti's negative narrative is at its best, especially presenting Gilbert Ryle's *'Fido'-Fido* objection against Millian nominalism (62), and Dummett's criticism of Frege's assimilation of sentences to complex names (63). This extension of the dialectics, as readers may easily appreciate, is particularly interesting not only for the contribution of Frege and Sellars, but also of Carnap, as for example the first source, even without a personal endorsement, of the very idea of 'material inferences': "if ever Carnap was close to inferentialism, the idea that good inferences go before logical forms is where he certainly gets off the train" (71). The chapter smoothly goes on to present the well-known characteristics of Brandom's semantics: its holistic shape; the putative difficulties in explaining the compositionality of meaning; the substitutional account of sub-sentential expressions; and the anaphoric conception of semantic vocabulary. Here, Turbanti adds to this reconstruction a final section (3.2) devoted to introducing Brandom's expressivism, and the 'meaning-use analysis' taken from BSD, in order to complement inferentialism with a preliminary grip on the conceptual toolbox of AP. In particular, Turbanti's presentation of logical expressivism—the idea that logical vocabulary plays the fundamental expressive role of 'elucidating', or 'making explicit', conceptual contents and relations—is the most complete and exhaustive in the literature (with an interesting reconstruction of its Fregean roots).

⁴ Kukla, R. and Lance, M. 2009, *'Yo!' and 'Lo!': The Pragmatic Topography of the Space of Reasons*, Cambridge (MA): Harvard University Press.

Chapter 4 presents ‘incompatibility semantics’ (hereafter IS), the formal semantics developed in BSD as a pragmatic meta-vocabulary⁵ for logical and modal vocabularies. This enterprise depends on the general possibility of using such meta-vocabulary to make explicit “the relations between practices and vocabularies” (97). In this sense, it is both a pragmatist and expressivist attempt. The chapter begins with a fruitful exploration of the reasons that may lead to develop a formal semantics in a general context influenced by Sellars’ distinction between formal and philosophical semantics: a framework that recommends the latter view as the default option. Turbanti clarifies how such misunderstandings rest on dubious representationalist conceptions of formal semantics; it is a tool for representing meanings, not a representational account of meanings (110-11). The presentation goes on to explore both the formal aspects and the underlying theoretical motivations of this apparatus, contributing to a more robust understanding of Brandom’s overall expressivism. Of particular interest is the semantic interpretation of IS (118-25), then the way in which logical vocabulary is defined (126-40), and the highlighting of some of its problems. Furthermore, the chapter introduces some noteworthy formal properties of the system, especially *conservativity*, as warranting the semantic recursiveness of IS (140-43) instead of the usual accounts based on compositionality. Conservativity permits the meanings expressed within the system to be fully recursive, even though these are *holistic*, and therefore *non-compositional*. And this property is of special importance for a holistic semantics like inferentialism, that *prima facie* would entail serious difficulties in explaining compositionality. As the last point shows, one master feature of this chapter (and of this book) is the use of the technical apparatus of AP, developed by Brandom in BSD, as a main tool in order to better clarify the wide project of inferentialism put forward with the monumental *Making it Explicit* (hereafter MIE).⁶ More generally, Turbanti manages to use effectively AP to provide a global account of the rational expressivism of which MIE is species of a genus.

Chapter 5 explores the possibility of extending the expressive power of IS in other directions and with slightly different philosophical motivations. These explorations provide interesting philosophical insights, especially dealing with open problems for both IS and its connection with inferentialism. Turbanti extends the formal framework in order to further develop the expressive power of this language. This chapter presents the most original sections of the book, and since Turbanti is a logician—like others who devoted special attention to this framework⁷—the main results are formal in character. He first uses IS to frame a Kripkean ‘possible worlds semantics’ with the effect of vindicating “incompatibility as a serious ground for modal vocabulary” (145). Then, he tries to use IS to develop a non-monotonic type of logical entailment (purported to

⁵ According to BSD, pragmatic meta-vocabularies are those sufficient in *specifying* the practices required to count as using certain *target* vocabularies: e.g., one may use “non-indexical vocabulary” as “sufficient to specify the practice required to count as using indexical vocabulary” (99).

⁶ Brandom, R.B. 1994, *Making it Explicit. Reasoning, Representing, and Discursive Commitment*, Cambridge (MA): Harvard University Press.

⁷ Especially the already mentioned Jaroslav Peregrin and Christian Fermüller; cfr. Fermüller, C. 2010, “Some Critical Remarks on Incompatibility Semantics”, in *The Logica Yearbook 2009*, London: College Publications, 81-95.

match the defeasible character of material inference). Both exercises are revealing: first Turbanti shows that certain fundamental results proved by Brandom about IS, e.g., the fact that it is a holistic semantics which is fully recursive without being compositional, can be proved as well in the modified Kripkean framework (152); second, after a nice summary of the connection between modality and defeasibility of material inferences, Turbanti explores the chances of developing a tenable notion of non-monotonic entailment suitable for Brandom's purposes, by developing IS in the direction of *Preferential Calculus*, even though this attempt still presents some open problems (172).

Chapter 6 relocates inferentialism in the wider context of two great philosophical traditions: its connections with the legacy of American Pragmatism and with German Idealism (especially the dialectics that goes from Kant to Hegel). Such relocation deals with the main open problems for normative inferentialism, that is, realism and the objectivity of conceptual norms in a context of subjective/perspectival discursive commitments. After summarizing Brandom's main views on inferentialism and realism, section 6.2 tries to identify Brandom's debts and connections—together with a number of divergencies—with the pragmatist tradition. Then Turbanti tackles Kant and Hegel: first, with the problems of Kant's normative theory of judgment; and then with *conceptual realism*, the 'Hegelian' solution to the problem dealing with the objectivity of conceptual norms. In particular, Turbanti explores Brandom's 'semantic' reading of Hegel's *Phenomenology of Spirit*, dealing with conceptual realism. This is the idea that reality is conceptually structured. Turbanti devotes some efforts to clarify how this Hegelian route is in line with Brandom's overall pragmatism, and also tries to present the main challenges for conceptual realism, such as the sharp remarks advanced by Jürgen Habermas.⁸ In 6.4, another important aspect deals with the Brandomian reading of Hegel's notion of 'determinate negation' in terms of material incompatibilities between commitments undertaken by means of assertoric judgments.⁹ Since conceptual contents become 'progressively determined' by ruling out other contents that show up as *incompatible* with them, this very practice amounts to a progressive *updating* of commitments and beliefs undertaken by speakers. This suggests a parallel with theories of belief-revision. The chapter ends with an attempt to emphasize the open problems of this Hegelian enterprise, well summarized also in the conclusion.

Despite original and innovative ways to introduce and discuss normative inferentialism, this book is more in line with other works in identifying the main axes of Brandom's theory: a normative pragmatics that understands the game of giving and asking for reasons as the core of discursive practice; an inferential account of the conceptual contents mongered by discursive practitioners; and an expressivist conception of logic, language, and rationality. But this presentation offers some interesting and original features. In particular, this reconstruction presents a sophisticated understanding of Brandom's expressivism, and rightly stresses the centrality of it for the overall inferentialist enterprise:

⁸ Habermas, J. 2003, *Truth and Justification*, Cambridge (MA): The MIT Press.

⁹ Being committed to a certain content *C* precludes entitlement to the contents that are *incompatible* with *C*.

Brandom's *rational* expressivism is the thesis that the application of concepts is essentially a process of expression, consisting in making explicit what is implicit, in the sense of turning something that can only be *done* into something that can also be *said* (8).

Furthermore, this understanding greatly profits from the vantage point provided by Brandom in *BSD*, where meta-vocabularies and vocabularies are analysed with a special focus on their expressive power and in connection with social practices. The way in which expressivism is worked out here provides more refined tools, also in order to look back at MIE. Turbanti claims that the two books are connected in a tighter way than Brandom himself believes. He claims that AP is "necessarily required in order to appreciate Brandom's later work in the philosophy of language" (10). From this point of view, Turbanti's combined presentation is a substantial improvement. In fact, according to this reading expressivism can be seen as a "unitary perspective" from which *BSD* and MIE "can be seen as part of the very same philosophical enterprise" (10). This focus on expressivism is also very important since it is in general, and despite its relevance, the less understood and appreciated part of Brandom's proposal, and this nice presentation may surely be of help to the reader.

University of Cagliari

PIETRO SALIS

Button, T., *The Limits of Realism*.
Oxford: Oxford University Press, 2013, pp. 278.

Tim Button's *The Limits of Realism* offers an in-depth discussion of Hilary Putnam's internal realism. The centrepiece of the book is a discussion of the model theoretic argument and the vat argument against external realism, the position that Putnam saw as diametrically opposed to internal realism. Button gives a compelling defence of both arguments, and takes them to refute external realism. However, he does not endorse internal realism, concluding instead that we must be 'something vaguely in-between' the two poles of external and internal realism (3). Hence the title of the book: we are to be realists within limits. Vague limits, but limits nonetheless.

The book deserves recognition for its spirited defence of the model theoretic argument and the vat argument. The vast majority of the literature on these arguments is negative, and this is a shame because, as Button shows, both arguments are far richer and more compelling than they are usually taken to be. Hopefully his book will help to rectify this situation. Certainly, it does a good job of presenting the arguments and their implications in a clear style, working through the details where necessary, whilst always keeping the broader picture in view.

In particular, I am sympathetic to Button's defence of the vat argument. However, even if Button succeeds in showing that the vat argument is sound and cogent, I do not think that it can move us away from metaphysical realism as he claims. Thus, I do not agree with Button that the vat argument shows that we must be 'something vaguely in-between' the two poles of external and internal realism. (I do not have space here to consider the model-theoretic argument. However, I think that similar considerations apply to it.)

Button starts out by considering the three claims that Putnam attributes to the external realist. The first is:

The Independence Principle: The world is (largely) made up of objects that are mind-, language-, and theory-independent (8).

The second is:

The Correspondence Principle: ‘Truth involves some sort of correspondence relation between words or thought-signs and external things and sets of things’ (8).¹

Taken together, these two claims might be thought to entail:

The Cartesian Principle: Even an ideal theory might be radically false (10).

An ‘ideal’ theory is one that fits all of the data we can gather, is simple, elegant, etc.

Without getting into details, one can appreciate the intuitive line of thought. The first principle says that the world is independent of the mind, and the second principle says that truth is a matter of producing representations that correspond with that world. So it would seem to be possible that even an ideal theory might be false due to a non-cooperative world, as the Cartesian Principle says.

Button points out that there are ways of reading the Independence Principle and the Correspondence Principle on which they present obviously true claims which, moreover, an internal realist would not deny. It is the Cartesian Principle that seems straightforwardly particular to external realism. The Independence Principle and the Correspondence Principle are only distinctive of external realism insofar as they are understood in such a way that they entail the Cartesian Principle. It is not obvious exactly how the Independence Principle and the Correspondence Principle are to be read so that they imply the Cartesian Principle, but Button does not press the external realist on this issue. Rather, his strategy is to wage a ‘war by proxy’ on the external realist versions of the Independence Principle and the Correspondence Principle by showing that the Cartesian Principle that they entail is false (71).

Now the issue arises: how are we to read the Cartesian Principle? Specifically, how are we to read ‘might’ in the claim that an ideal theory might be radically false? The claim that the Cartesian Principle is entailed by external realist versions of the Correspondence Principle and the Independence Principle suggests that it is to be read as an expression of a metaphysical or conceptual possibility. That is the sort of consequence that one would expect a metaphysical claim such as the Independence Principle and a semantic claim such as the Correspondence Principle to have, if any.

There is another reading, suggested by the label ‘*Cartesian Principle*’. This is to read the principle as a statement of epistemic possibility. The claim would then be that for all we know (and, presumably, ever could know) even an ideal theory could be false. Of course, a natural thought is that the metaphysical version of the Cartesian Principle could play a key role in an argument for the

¹ Button is here quoting Putnam, H. 1980, “How to Be an Internal Realist and a Transcendental Idealist (at the Same Time)”. In Haller R. and Grassl W. (eds.), *Language, Logic, and Philosophy*, Vienna: Holder-Pichler-Tempsky, 100-108.

epistemic version of the Cartesian Principle. Many sceptical arguments start out with the claim that a sceptical scenario is a metaphysical possibility, and then move to the claim that it is also a for-all-we-know possibility. Perhaps this is why Button does not sharply distinguish the two readings of the principle, taking it that both versions of the principle stand or fall together. However, we should recognise the possibility that the Cartesian Principle will turn out to be false on an epistemic reading, despite being true on a metaphysical reading. If this should happen, it would not seem to be problematic for an external realist who endorses versions of the Independence Principle and the Correspondence Principle that jointly entail the metaphysical version of the Cartesian Principle.

In fact, I think that if Button's defence of the vat argument succeeds then this is exactly the situation in which we find ourselves. The Cartesian Principle, on both its metaphysical and its epistemic reading, seems to be true because we might be in a sceptical scenario. For example, we might (metaphysically, epistemically) all be brains in vats, fed sensory experiences by a supercomputer. If this were the case, then it seems that even the best theory we could ever come up with would be false. It would say, for example, that I have hands, that I live on Earth, that it is sunny. However, I would not have hands, would not live on Earth, and it would not be sunny. So to establish that the Cartesian Principle is false we need an argument that shows that we are not all permanently envatted brains (and, ultimately, that we are not in any sort of radical sceptical scenario, but we will not be concerned with this further step here; let us grant Button's convincingly argued claim that if the vat argument works on one radical sceptical scenario it works on them all).

The vat argument looks like it might do the job. Button puts it like this. Let us write 'x is a BIV' as shorthand for 'x is an eternally envatted brain and so is everyone else'. We can then argue as follows:

- (1) A BIV's word 'brain' does not refer to brains.
- (2) My word 'brain' refers to brains.
- (3) So: I am not a BIV (118).

Premise (1) is defended by an appeal to semantic externalism. A BIV would not have had the kind of causal contact with brains necessary in order to refer to them. After all, it would never have had perceptual contact with a brain, and it cannot have talked to anyone who has, for, by hypothesis, if something is a BIV then everyone else is also an eternally envatted brain. Premise (2) is defended by appeal to the claim that I can disquote to specify the referents of words in my own language. Thus, when I say 'brains' I refer to brains. Moreover, if I am to take myself to be entertaining the sceptical worry in the first place I am committed to the claim that I can refer to brains: if I am to worry that I am a *brain* in a vat I must be able to refer to brains.

Button concludes that the vat argument shows that I am not a BIV, and so the Cartesian Principle is false (remember that we are granting that if the vat argument succeeds in showing that I am not a BIV, it succeeds in showing that I am not in any radical sceptical scenario). However, he notes that the vat argument cannot be used to rule out some *less* radical sceptical scenarios, that is, scenarios that falsify fewer of our beliefs about the external world. For example, suppose we read 'x is a BIV' as shorthand for 'x was envatted via some undetectable process yesterday, although up to that point x lived a normal life of

the sort that I normally take myself to be living'. In that case, the hypothesis that I am a BIV will falsify many of my beliefs, but fewer than the hypothesis that I am eternally envatted would falsify. If we read 'I am a BIV' in this way then premise (1) will be false. No plausible semantic externalism will entail that a subject who has spent most of its life interacting with brains in the way I ordinarily take myself to cannot refer to brains. So the vat argument cannot be used to rule out the possibility that I am in this less radical sceptical scenario.

Because it does not rule out the possibility that I am in some less radical sceptical scenarios, Button concludes that the vat argument does not support internal realism, which does not countenance the possibility that we might be in *any* sceptical scenario, even of a less radical sort. Rather, Button argues that as we increase the radicalness of sceptical scenarios there is no precise point at which vat style arguments kick in. However, they undoubtedly *do* kick in, and they certainly rule out radical sceptical scenarios. Thus, there is a limit, albeit a vague one, to how radical a sceptical scenario we can worry about. Since the point at which we position ourselves between the poles of external and internal realism is intended to be tied to how radical a sceptical scenario we can worry about, this conclusion leaves us 'vaguely in-between' external realism and internal realism.

I agree with Button that I can use the vat argument to show that I am not in any radical sceptical scenario. I also agree with him that as sceptical scenarios become less radical the point at which vat style arguments stop working is vague. However, I do not think that any of this pushes us away from external realism.

I can use the vat argument to rule out the possibility that I am a BIV (on the original reading: 'x is an eternally envatted brain, and so is everyone else'). That is, the argument shows that it is not an epistemic possibility that I am a BIV, and so (on the assumption that I can also use it to show that I am not in any other sort of radical sceptical scenario) it shows that the *epistemic* version of the Cartesian Principle is false. However, it does not show that it is not metaphysically possible for me to be a BIV. For all that the vat argument says, BIVs are metaphysically possible, and there is a possible world where I am a BIV. So it is unclear why an external realist who endorses versions of the Correspondence Principle and the Independence Principle that entail the *metaphysical* version of the Cartesian Principle should be bothered by any of this.

It is true that the vat argument does entail a metaphysical impossibility. Although the vat argument does not entail that there is no possible world in which I am a BIV, it does entail that there is no possible world in which a BIV can refer to brains. However, that there are possible worlds in which BIVs can refer to brains was not the claim that the external realist versions of the Correspondence Principle and the Independence Principle were said to entail. Nor, intuitively, does it seem that they would be expected to entail it, no matter how robustly we read these claims. The idea that truth is a matter of producing representations that corresponded with a world that is independent of our minds does not seem to carry any implications regarding who can produce certain representations. It takes nothing away from whatever robustness these ideas might be thought to have if we say that a BIV would not be able to refer to brains, and thus would not be able to describe its predicament. If anything, the fact that the BIV cannot describe its predicament only makes that predicament

more terrible. Another reason for the external realist to be relieved that we are not in it.

So I do not think that the external realist need be worried that the vat argument shows that the epistemic version of the Cartesian Principle is false, since it does so in a way that is compatible with the metaphysical version of the Cartesian Principle. Indeed, the external realist may welcome this result, for without it she might have been stuck with an insoluble epistemic problem, namely, the truth of the epistemic version of the Cartesian Principle, which looks more or less like scepticism. Now she can be rid of such scepticism, whilst continuing to claim that the world is in some very strong sense mind independent, and that truth consists of us producing a representation that corresponds with that world. The external realist has a metaphysical and semantic outlook on which it is deeply contingent—lucky even—if we are not in a sceptical scenario in which even our best theory would be false. But the vat argument shows that, luckily, we happen *not* to be in such a sceptical scenario. The external realist will breathe a sigh of relief, for she endorses a picture of the world on which we easily might have been. Nonetheless, she will do so while continuing to endorse that picture.

So, in conclusion, I do not think that the vat argument can push us away from external realism. This is because, if it succeeds, it shows that the epistemic version of the Cartesian Principle is false, but it does not show that the metaphysical version of the Cartesian Principle is false. However, the external realist is only committed to the metaphysical version of the Cartesian Principle, so this will not bother her. Indeed, she may welcome it as the solution to a sceptical problem that her commitment to the metaphysical version of the Cartesian Principle may seem to give rise to.

A more speculative conclusion suggested by these considerations is that the relevance of the vat argument to the realism debate is limited. If the vat argument has consequences, they would seem to be epistemological, allowing us to rule out at least some radical sceptical scenarios. Exactly how far this gets us as a response to scepticism remains an open question, however.²

University of Edinburgh

JOSHUA ROWAN THORPE

² Thanks to Massimo Dell'Utri for helpful comments.