Extensional Realism: Interesting and Uninteresting Truths

MSc Thesis *(Afstudeerscriptie)*

written by

Andrew Fish
(born 31.12.1997 in The United Kingdom of Great Britain and Northern Ireland)

under the supervision of **Dr. Sebastian De Haro Ollé**, and submitted to the Examinations Board in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam.*

Date of the public defense:
09.02.2023

Members of the Thesis Committee:
Dr. Katia Shutova
Dr. Dingmar van Eck
Dr. Sebastian De Haro Ollé
Prof.Dr. Sonja Smets

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

Abstract

Sebastian de Haro offers a new kind of scientific realism, *extensional realism*, which allows us to be realists, if not about the intensions of scientific terms, then at least their extensions. After examining the view and how it fits into the wider debate about scientific realism, it is argued that the notion of approximate truth employed threatens the novelty of the position, as well as commits the extensional realist to realism only about scientific claims of no interest. By adapting Karl Popper's verisimilitude in a particular way, extensional realism is equipped to solve this problem by identifying which theories are the best according to the usefulness of the truths they espouse.

Acknowledgements

Chiefly, I wish to thank my supervisor, Sebastian, for your constant guidance and patience throughout the thesis writing process. Before your instruction, I had only the vaguest inkling of the goings on in the philosophy of science, but your instruction ignited in me such a passion for the subject that I nonetheless felt confident in writing this thesis.

Likewise, I thank the members of my committee, who ensured the defence of my thesis was among the highlights of my academic career and whose questions and criticisms prompted much useful reflection.

It is no exaggeration to say that without the support of Tanja, this thesis would surely have been of a poorer quality achieved over a much longer period, if, indeed, it would have been finished at all.

I thank also my parents, Kevin and Rosalind, who, despite displaying no interest in formal logic, are more invested in my studies than any parent has an obligation to be, for which I count myself exceedingly fortunate.

Finally, I thank the flatmates whom I have inflicted myself upon throughout the Master of Logic programme. Three in particular suffered me whilst writing the bulk of my thesis and deserve special thanks (and probably apologies, but these are not forthcoming). Fiona, with you I could always set the world to rights, excepting in matters of food, where you maintained a bizarre insistence that it should have nutritional value and be in anyway authentic. Tibo, with you I could always set the world to wrong. I came to rely on our many heated philosophical debates about things of vanishingly small consequence. If you only had not been constantly correct, I might have admitted to being wrong. Marie, with you I could always share my most profound thoughts– e.g. that there exist only two kinds of sausage, fat and thin, and that the pinnacle of human invention is the existence of cat photos on the internet. Though you often laughed at these genius insights, I could tell, deep down, you agreed.

Contents

Chapter 1

Introduction

Scientific realism, the position, broadly speaking, that scientific theories are approximately true, or that science aims for approximate truth, is a debate at the heart of the philosophy of science. It has had many formulations and many variants– from *structural realism* (realism about the structure of scientific theories), to *causal realism* (realism about those aspects of a theory essential to explain the causal role of those entities the theory invokes). Likewise, there have been many arguments lobbied for and against it– including historical arguments about the development of science and semantic arguments about the references of scientific terms.

With all that said, it might be thought that there is little left of substance to be added to the debate, beyond the clarifying and slight alteration of already existing positions and arguments. Sebastian De Haro, however, has shown that that this is not the case, with his development of *extensional realism* (2020), a new kind of realism, invoking a new notion of approximate truth, that permits us to be realists about the extensions of terms, though not necessarily their intensions. This is framed as applying classical referential semantics to scientific theories, a resource taken to be underutilized in much of the discourse surrounding scientific realism.

To help extensional realism breathe fresh life into the realist debate, it should be shown that the position is genuinely distinct from those forms of realism that already exist. Through the consideration of *Ramsey sentences*, sentences which have had their theoretical terms eliminated through existential quantification, the worry arises that this is not the case. On the assumption that, between two competing theories, the better one is the more approximately true, it follows from the extensional realist definition of approximate truth that the best theories ought to be Ramsey sentences. For different reasons, this is the position taken up by certain structural realists, and is thus a problem for the novelty of De Haro's approach. Worse still, if we carry on with the assumption that more approximately true theories are simply better, we get the result that the best theories are the most nebulous.

To resolve this, a criteria for theory choice is needed that balances the approximate truth of a theory against its content. The best theories are not only more true than mot, but they have more to say about how the world works, and what they have to say is of more use. Work has already been done to develop such a criteria, but, as shall be shown, not all of it can be adopted by the extensional realist. A variation of Karl Popper's *verisimilitude* (1963, 1975) is proposed that adopts the extensional realist framework in such a way as to defeat the slide into structural realism and nebulousness via Ramsey sentences. It does so first by requiring that the worth of a theory be domain-specific, hence a given theory that might have utility in one area of inquiry lacks value in another. Second, whilst the extensional realist gives a semantic theory of approximate truth, the theory of verisimilitude give is syntactic. Ramsey sentences, it shall be shown, do not take priority over their competitors, under this picture, in domains when both are approximately true.

The aim of this thesis, therefore, is to give an account of extensional realism that truly distinguishes the view from its competitors with this new criteria of theory choice.

1.1 Structure

Chapter two shall give a more precise explanation of scientific realism. This is important to establish the significance of extensional realism in the realist debates, as well as highlight the new moves De Haro is making. The chief argument against scientific realism, the pessimistic meta-induction, as well as the chief argument in its favour, the miracle argument, are also outlined. Extensional realism, it shall be shown, proports itself properly to these two arguments, escaping the problems of the firts whilst satisfying the demands of the second.

Chapter three explains extensional realism itself. It begins with an overview of the theory and then gets down into specifics. De Haro's position is motivated by debates in the philosophy of language regarding reference, and how they are applied to the meaning of scientific language. Discussion of these debates and their relevance is thus required. Next, the technicalities of how classical referential semantics generate a new kind of approximate truth is provided, which serves to meet the demands for a realist view laid out in chapter two.

Chapter four is where the problem of novelty is introduced. Structural realism as a general position is first defined. Next is given the problematic structural realism of Grover Maxwell that invokes Ramsey sentences. It shall be shown that the Ramsey sentences Maxwell accepts are more approximately true, under the extensional realist framework of chapter three, than the commonly accepted claims of today. Worse still, a hypothetical variant of structural realism is thought up that endorses only the most nebulous of Ramsey sentences. These are, by far, the most approximately true, and apparently commit the extensional realist to saying that the only true claims of science say almost nothing of actual importance. It is concluded that a new criteria for theory choice is required, one that takes approximate truth into account but does not rely on it solely.

Chapter five is where this solution is offered. Popper's verisimilitude is first examined. De Haro also considered this solution, but rejects it. It shall be shown why, as De Haro, conceived verisimilitude applying to extensional realism, this was a good decision. Next shall be considered the semantic notion of truthlikeness. On the face of it, this seems more coherent with extensional realist semantics, allowing one to easily adapt the use of possible worlds into a neat framework for selecting the best theories that excludes Ramsey sentences where needed. However, this makes use of a graded metric of how informative a proposition is that is entirely at odds with the extensional realist's view of approximate truth. If employed, the extensional realist would be committed to saying that the most true theories are the least informative. The solution to the problem laid out in chapter four, it shall be argued, does in fact lie in Popper's verisimilitude, only applied in a much different way than De Haro imagined. A syntactic variant of verisimilitude, one which relatives which scientific theories are the best to particular domains of scientific inquiry, solves the problems laid out in previous sections.

Chapter six concludes the thesis. Extensional realism needs more than just its semantic notion of approximate truth, but also a syntactic explanation of verisimilitude. With these two together, the theory is a useful and innovative addition to a debate that otherwise looked worn down.

Chapter 2

Scientific Realism

2.1 What is Scientific Realism?

Scientific realism has had many formulations. Very broadly, it is the claim that the theories of science are, or purport to be true or approximately true. Bas Van Fraassen points out the danger of defining it as the theoires of science merely *being* true, as it opens a philosophical thesis to scientific refutation. Rather, he defines scientific realism as the position that science *aims* at truth:

> [such formulations] are too naive; [they attribute] to the scientific realist the belief that today's theories are correct. It would mean that the philosophical position of an earlier scientific realist such as C. S. Peirce had been refuted by empirical findings. I do not suppose that scientific realists wish to be committed, as such, even to the claim that science will arrive in due time at theories true in all respects—for the growth of science might be an endless self-correction; or worse, Armageddon might occur too soon (1988, p. 1065

This is not an uncontroversial point, however. Darrell P. Rowbottom, for example, rejects any kind of talk ragrading the "aims of science" as it pertains to the realism debate, since the proper way to account for the aims of science is not philosophical analysis of methodology, but instead to simply ask scientists, since the aims of science must be the product of the aims of those who practice it:

> Now if the aim of science were understood to be determinable by looking to the aims of individual scientists in doing science—what they hoped to achieve, be it fame, fortune, finding the truth, or finding empirically adequate theories—then the difference of opinion would be easily settled by sociological study. The appropriate procedure is simple. State what has to be true of the aims of individual scientists in order for the overarching 'aim of science' to be

x, and then study the aims of individual scientists (2014, p. 1213).

The question of whether realism is to be defined in terms of the goals of science shall not be answered here. By authorial preference, realist theories shall be taken to denote the view that scientific theories *are*, at least in some respects, (approximately) true. Nothing of consequence to the claims of this work is lost because of this choice, but, if its reader takes upset at this, it is allowed that phrases of the kind "...is (approximately) true" be read as though "...is aimed at (approximate) truth".

More specific formulations of scientific realism are prone to their own objections. The famous (amongst philosophers of science) account given by Larry Laudan(1981), for example, has drawn much criticism that realists and anti-realists alike may wish to avoid:

> R1) Scientific theories (at least in the 'mature' sciences) are typically approximately true and more recent theories are closer to the truth than older theories in the same domain;
>
> R2) The observational and theoretical terms within the theories of a mature science genuinely refer (roughly, there are substances in the world that correspond to the ontologies presumed by our best theories);
>
> R3) Successive theories in any mature science will be such that they 'preserve' the theoretical relations and the apparent referents of earlier theories (i.e., earlier theories will be 'limiting cases' of later theories).
>
> R4) Acceptable new theories do and should explain why their predecessors were successful insofar as they were successful[...]
>
> R5) Theses (R1)-(R4) entail that ('mature') scientific theories should be successful; indeed, these theses constitute the best, if not the only, explanation for the success of science. The empirical success of science (in the sense of giving detailed explanations and accurate predictions) accordingly provides striking empirical confirmation for realism (pp. 20-21).

Though many of these points are adopted to some extent by many realists, they have not been without criticism. (R2), for instance, invokes "observational terms". There is much debate as to whether there is any proper distinction between observable and non-observable entities. If there is none, then "observable terms" is empty and realists need not hold to it (if, indeed, they did in the first place).[1] (R1) is rejected by David Papineau, who, though still calling himself a realist, grants that the terms of scientific theories routinely fail to refer. To

[1] The paradigm instance of this debate occurrs between Grover Maxwell (1962) and Van Fraassen (1998). According to the former, the distinction is vague– there being a sorites series

remedy this, he uses a descriptivist notion of approximate truth according to which this is no worry for realism (2010).

Though these points cannot, then, be taken to provide a neutral definition of scientific realism that all can accept, they provide a flavour of the kind of points of contention in the debate. We may thus characterize scientific realists as those who are generally disposed to affirm that the statements of mature scientific theories are (approximately) true, with successful reference to unobservable entities which is largely preserved across successive theories. One might reject any of these points (but for the first) and still be called "realist", but a realist will accept some combination of most of them. As the question of what scientific realism is (in a general form) has been answered, we turn now to to the arguments surrounding it.

2.2 The Pull Away from Scientific Realism: The Pessimistic Meta-Induction

Whilst there are several arguments that have been posited against the scientific realism, *extensional scienific realism*, the subject of this work, is motivated by answering the *pessimistic meta-induction* argument, and so it is on which we shall focus. In his summary of the argument Wray (2015) points to a variety of formulations, from a true inductive argument, to a *reductio ad absurdum.* In all cases, the first thing done is to point to the historical record and declare a wealth of well developed scientific theories that were, in their day, empirically successful,[2] but which we now reject. Laudan (1981) gives a substantial list of such theories as well as the claim that it could very easily be expanded:

- the crystalline spheres of ancient and medieval astronomy;
- the humoral theory of medicine;
- the effluvial theory of static electricity;
- "catastrophist" geology, with its commitment to a universal (Noachian) deluge;
- the phlogiston theory of chemistry;
- the caloric theory of heat;
- the vibratory theory of heat;
- the vital force theories of physiology;
- the electromagnetic aether;

between things one can see through simple means like the naked eye or a telescope being called "observable" on the one hand, and things whose observation requires complex machinery like electron microscopes being called "unobservable" on the other. Van Fraassen concedes that the distinction is vague, but maintains that this is not enough to dismiss the distinction, which still holds on intuitive grounds.

[2]A theory is taken to be empirically successful when it has many cases of experimentation and/or observation that confirm, the theory and none that refute it

- the theory of circular inertia;
- theories of spontaneous combustion (Ibid. p. 33).

Elements of this list might be subject to scrutiny. Psillos (1999) does precisely this, eliminating members on the grounds that they were not properly mature enough, or explanatoraly powerful, to be called empirically successful. This is hardly a principled approach, however, and it should be noted that Psillos eliminates only a portion of this list. Since, if Laudan is to be believed, his list of empirically successful and now-discarded theories is a fraction of the length it could have been, going through each list-member one at a time and crossing it off is unlikely to defeat the pessimistic meta-induction at the first hurdle. The anti-realist can simply add to the list in absence of a principle explaining why this cannot be done.

From here a simple induction in-line could be made:

Premise The scientific theories of the past, though empirically successful and believed true, were found to be false (see the above list).

Induction Therefore, though the scientific theories of today are empirically successful, it is unlikely that they are true.

As a naive formulation of the pessimistic meta-induction, this is a rather weak argument, but it serves to explain the kind of argument that could properly be called a pessimistic meta-induction.

Some authors, notably Laudan (1981) and Putnam (1978), frame the argument by targeting reference. Accordingly, the central terms of now-defunct scientific theories that were otherwise mature and empirically successful were found not to refer. Since the central terms of such theories were non-referring, their central claims could not be true, approximately or otherwise. For example, since there is no thing to which the term "phlogiston" refers– i.e. phlogiston does not exist–, the key claim of the phlogiston theory of chemistry, "burning is caused by the release of phlogiston" is false. By induction, the same can well be supposed of today's theories, successful as they may be. As Putnam puts it:

> One reason this is a serious worry is that eventually the following meta-induction becomes overwhelmingly compelling: just: as no term used in the science of more than fifty (or whatever) years ago referred, so it will turn out that no term used now (except maybe observation terms, if there are such) refers (1978, p. 25).

As noted, many realists– though not all (see Papineau, 2010)– will want to hold that the central terms of today's scientific theories are, at least for the most part, generally referring. To do otherwise would require a very particular notion of approximate-truth.

It might, however, be said that to focus entirely on reference is missing the bigger picture. Stanford (2015) grants that the central theoretical terms of scientific theories refer, precisely because this does not amount to much. As a result, if he is right, the scientific realist who proves the terms of their theory are referring has accomplished little in furtherance of their cause.

Stanford's point runs thus: In the first place, when rejecting a mature and empirically successful scientific theory for its successor, we rarely reject its key existential commitments (ibid. p. 401). Secondly, whether we take a term to refer is dependent on factors irrelevant to scientific realism, and so the continuity of reference of terms across theories is merely superficial. This is because the reference of a term comes from its use, which is subject to vicissitudes that are arbitrary to the question of scientific advancement (p. 403). For example, the term "atom" has gone through numerous revisions in meaning across its history simply because individual theorists wish to continue using the term to pay homage to previous theorists; not because the thought behind the term is substantively the same as in previous theories

> In sum, how a term will come to be extended or applied by speakers in novel, changed, or unexpected scientific contexts seems to owe as much the vicissitudes of later historical fortune or circumstance as anything else (ibid. p. 405).

Establishing continuity of reference, then, is more so a question of linguistic happenstance than scientific fact. His third point is simply that establishing existential claims, and therefore establishing a reference to a term, tells us almost nothing. "Atoms exist" tells us nothing about what atoms are, their internal properties or how they relate to other objects in the world. As regards to the pessimistic meta-induction, the driving force is not showing that the terms we use have continuously failed to refer, but instead that our beliefs about the objects to which we refer in mature, empirically successful scientific theories have turned out untrue:

> At the end of the day, then, what the historicist critic of realism is most fundamentally committed to is the idea that whether or not particular existential commitments of current theories are held to be true and whether or not particular terms are held to be referential, the central commitments of future theoretical orthodoxy will (or would) ultimately be separated from those of the present by differences as fundamental, profound, far-reaching, and unpredictable as those that separate our own theories from their historical predecessors (ibid. p. 415).

Summing up Stanford's position, we can say the following. Establishing that a term has reference says very little in the first place; what it does establish is determined largely by factors irrelevant to the debate concerning scientific realism; and the real bite of the pessimistic meta-induction is not inducing from a discontinuity of reference, but a discontinuity of beliefs about the things we refer to. In a defence of realism, we can say that an account which preserved the success of reference might be beneficial, at least to certain flavours of realism, but it is not sufficient to defeat the pessimistic meta-induction.

Turning from whether the pessimistic meta-induction should target reference, there is the question of the argument's form. The most dominant formulation of the pessimistic meta-induction, Laudan's, is not actually an induction, but is instead a *reductio ad absurdum*. Psillos (1999) reconstructs the argument as follows:

P1 "Assume that 'currently successful theories are approximately true'.

P2 'If currently successful theories are truth-like, then past theories cannot have been.

P3 'These ... false theories [of the past] were, nonetheless, empirically successful.

C Therefore, 'empirical success is not connected with truth-likeness and truth-likeness cannot explain success'.

Therefore, 'the realist's potential warrant for [the claim that currently successful theories are approximately true] is defeated'" (Psillos 1999, p. 102, in Wray, 2015, p. 64).

Far from an induction, this argument takes the form of *modus tollens*. Indeed, Timothy Lyons calls this formulation the *meta-modus tollens* (Lyons, 2002). Accordingly, the existence of mature theories that were, in their day, empirically successful, but which are known now to be untrue, falsifies the claim that "if a mature theory is empirically successful, then it is true". This presents a problem for scientific realism by undermining the grounds according to which one might call a scientific theory true. After all, since we adopt theories based on their empirical success, which does not speak to their truthfulness, then, lacking some other criterion for the adopting of scientific theory that entails such a theories truth, we have no reason to suppose that our currently adopted theories are true.

As a *reductio*, this argument is deductive, and so stronger, if its premises are shown true, than the inductive form of the pessimistic meta-induction. That said, it differs sufficiently from the likes of Putnam's formulation, in both form and its lack of focus on reference, that it is worth considering in its own right.

We have then, identified four different ways to formulate the pessimistic meta-induction. In the first place, there is the distinction between genuine inductions and *reductio ad absurdum* arguments. In the second, there is the distinction between reference-targeting arguments and non-reference targeting arguments.

2.2.1 How to Resist the Induction

Having summarized the pessimistic meta-induction, it does well to outline how a realist might go about defeating it. As established, each instance of the pessimistic meta-induction begins by painting a poor picture of the history of science, outlying a parade of now-discarded theories who, in their day, were mature and empirically successful. One might hope to head-off the induction by debunking the list it begins with. This, as has been said, without a principled reason for why such a list cannot exist, rather than merely going through its members one-by-one, is likely to lead to a barrage of further examples from the anti-realist, bringing the debate back to its beginning.

Putting this approach aside, the most viable avenue to defeating the induction is to establish some commonality between the theories of the past and their successors. If it could be shown that some element of past theories has survived scientific advancement, then we might at least be realists with respect to this element.

To illustrate, suppose T, a mature and empirically successful (for its time) scientific theory, has been succeeded by T_2, a theory with these same merits, which, in turn, has been succeeded by the theory accepted today, T_3. The set of claims made by T, T_2 and T_3 differ such that each set contradicts the other. It seems, then, that this is a case ideal for the anti-realist, who will point out the likelihood of T_4, which will contradict T_3. However, as it turns out, though the sum of claims made by each theory contradicts the claims of each other theory in the sequence, each nonetheless makes the claim, or set of claims, c. Both the theory of phlogiston and subsequent theories of combustion, for example, agree on some points regarding which substances are capable of burning.

To win a victory over the inductive formulation of the pessimistic meta-induction one can say that, though we might well find that there will be a T_4, the existence of c in all previous theories makes it likely, by an induction we might call optimistic, that T_4 will also contain c. The inductive formulation, then affords us no reason to doubt the truth of c. This tactic also affords success against the *reductio ad absurdum* formulation, since it cannot show that current scientific theories are untrue *with respect to c*. We start, as the argument bids us, to assume that our current scientific theory T_3 is truth-like *with respect to c*. **P2** of the above formulation says that "if currently successful theories are truth-like, then past theories cannot have been"(ibid.). To this, we can say that the current success of T_3, which, by assumption, is truth-like *with respect to c*, does not preclude the truth-likeness of T and T_2 *with respect to c*, since c is precisely what they all have in common. **P2**, then, may be rejected, at least in the case of c.

It can so be seen that realism may be saved, though perhaps only in some restricted way, by appeal to commonality across scientific theories throughout time.The commonality here invoked regards the same (set of) claim(s) being made across theories. This is a hypothetical for the purposes of an example. It may well be in practice that no such c can be found. This does not by itself eradicate hope for any kind of realism. c might be replaced with another commonality about which we can be realists (provided that it leads to empirical success, or else we have no reason to accept it in the first place regardless of philosophical argumentation). Perhaps each theory logically entails a claim rather than makes it outright, or, as *extensional realism* claims, there is commonality between the extensions of terms used in succeeding theories. c is used here to simply lay out the *form* of realist replies to the pessimistic meta-induction, rather than as an example of what the particulars of such replies must contain.

Moving from the hypothetical to the actual, a proper defence of realism must identify what this commonality *is*. A pleasing answer for the realist would be simply to say that whatever claims made by a theory that are responsible for its empirical success are retained by its successors, and so commonality is to be found there. This approach, though, runs the risk of being overly hopeful. Psillos (1999), for example, points to a "causal core" of descriptive claims. Accordingly, a term refers when there is some object that satisfies a set of descriptions that make up a causal core. Descriptions outside of the causal core may fail to be satisfied by the object to which a theorist wishes to refer, and so these descriptions are an untrue part of the theory to which they belong, but this is not terribly problematic since they do not determine reference, and so, as Psillos takes it, are not the key claims of the theory. By holding that a causal core of descriptions is maintained across theory change, Psillos hopes to preserve continuity of reference and establish a commonality to defeat the pessimistic meta-induction. For example, though the concept of "aether" is no longer in use, the set of descriptons that made up its causal core are still in use today, being satisfied by light waves and the electromagnetic spectrum. Thus, "aether" successfully referred, and commonality is in play.

Appealing as this sounds, Stanford calls it a pyrrhic victory:

> Of course, this account of the matter invites the realist to choose the core causal descriptions she associates with the central terms of past theories rather carefully, with one eye on current theories' claims about nature, so there is more than a whiff of *ad hocery* about the proposal. But even if we set this worry aside and assume that the realist can delicately titrate the core causal descriptions she associates with the crucial terms in successful past theories so as to render them referential by the lights of current theories, Psillos's victory will nonetheless remain a Pyrrhic one. The reason is that this case for the referential status of central terms in successful past theories simply invites from the historical record a renewed form of the pessimistic induction itself, this time concerning our ability to distinguish (at the time a theory is a going concern) which of our beliefs about an entity are actually part of its core causal description [...]
>
> The core of the problem is that nineteenth-century theorists themselves strenuously disagreed with the very assessment Psillos offers of what would have to be true of an entity in order for it to play the ether's causal role, that is, with his very claim about which descriptions of the ether enter into its core causal description (Stanford, 2003, p. 559).

Psillos' victory is pyrrhic because, to motivate an answer to the pessimistic meta-induction, Psillos must be overly selective of what descriptions belong to a causal core. In the first place, Psillos takes the constituents of the causal core of a theory to be different than what the theory's proponents would have taken it to be. As such, Psillos' position is arguably at odds with scientific practice, adopted merely for *ad hoc* philosophical reasons. Worse still, a new variety of pessimistic meta-induction is introduced. Since we can be mistaken as to what descriptions constitute a causal core, and so must be true, and which do not, and so can be false, it might be that we are entirely mistaken as to what makes up the causal core of today's theories. Indeed, for Psillos' position to work, successive theories show that we are continually mistaken about what makes up a causal core. As such, we can pessimistically induce that we are, at present, mistaken about the causal cores of our theories. Since Psilos' position allows us only to be realists with respect to causal cores, and we are likely mistaken about the causal cores, Psillos' realism is terribly insecure.

To resist the pessimistic meta-induction, then, the realist must show commonality across theories. What these theories hold in common cannot be chosen for *ad hoc* reasons, lest the realist's resistance be pyrrhic. As shall be shown in the next chapter, *extensional realism* satisfies these conditions. For now, established is the general stratergy the realist must take.

2.3 The Pull Towards Scientific Realism: The Miracle Argument

All this said, a path to defeating the challenge to scientific realism does not establish *why* we should accept the position in the first place. For that, we turn to the miracle argument, which, in a general formulation, says that, if scientific theories are untrue, it would be highly improbable– i.e. miraculous– that they should enjoy empirical success. From this, it is concluded that, because there are an abundance of empirically successful scientific theories, the best explanation of this is their truth, hence we ought to be scientific realists. As Putnam puts it:

> But if [the gravitational field or the metric structure of space-time] don't really exist at all, then it is a miracle that a theory which speaks of gravitational action at a distance successfully predicts phenomena; it is a miracle that a theory which speaks of curved space-time successfully predicts phenomena; and the fact that the laws of the former theory are derivable 'in the limit' from the laws of the latter theory has no methodological significance (1978, p. 19).

For the scientific anti-realist to best this argument, then, an explanation of why we can have such theories as Putnam describes make accurate predictions which do not rely on their truth.

Van Fraassen offers an evolutionary explanation for why theories, accepted though untrue, are empirically successful. Van Fraassen holds that, rather than truth, science aims at *empirical adequacy* (1998). A theory is empirically adequate when:

> [...] what it says about observable things and events in this world, is true– exactly if it 'saves the phenomena'. A little more precisely: such a theory has at least one model that all the actual phenomena fit inside (ibid, p. 1069).

To exemplify, let T be a theory and o some observable phenomenon. For T to be empirically adequate with respect to o, there is some model $\mathcal{M}$ of T which entails e, which we can denote by writing $\mathcal{M}_T \vDash o$. Note that this requirement does not require the truth of o. Scientific advancement, so construed, is the creating of models that are increasingly empirically adequate. Suppose, to illustrate, that some second observable phenomenon o_2 was discovered. T and its replacement are such that, for some model $\mathcal{M}$ of T, $\mathcal{M}_T \nvDash o_2$, whereas, for every $\mathcal{M}$ of T_2, $\mathcal{M}_{T_2} \vDash o_2$. In plain English, this says that T fails to predict o_2, and T_2 does so. We therefore say that T is not empirically adequate, since it predicts something other than what we have observed, and so, all else being equal T_2 replaces T. These are the relevant aspects of the view van Fraassen calls *constructive empiricism.*

The evolutionary explanation for the success of scientific theories under constructive empiricism runs parallel to an explanation of natural selection. Just as, in nature, organisms with traits detrimental to survival are unlikely to survive and propagate, whereas those with traits conducive to survival survive and reproduce, a similar phenomenon occurs in the case of which scientific theories survive. Those which we find to be empirically unsuccessful are discarded, leaving as survivors only theories that are successful. At no point in this process is the truth or untruth of a theory invoked:

> I claim that the success of current scientific theories is no miracle. It is not even surprising to the scientific (Darwinist) mind. For any scientific theory is born into a life of fierce competition, a jungle red in tooth and claw. Only the successful theories survive—the ones which in fact latched on to actual regularities in nature (ibid. p. 1084).

As a reply to the miracle argument, this, by itself is insufficient. As Alan Musgrave (1998) points out, this approach misses the point of the miracle argument. Van Fraassen's account explains the claim "all presently accepted scientific theories are successful". The demand of the realist, however, is to explain the claim that, for a given accepted theory T, "T is successful":

> Of course, the Darwinian question is not "Why does the mouse run away from the cat?" but, rather, "How did this piece of mouse behavior evolve?" The Darwinian answers this question roughly in the terms suggested by van Fraassen: given an environment full of mouse-hunting cats, cat-fleeing mice are more likely to survive, reproduce, and pass their cat-fleeing behavior on to future generations. But the Darwinian explanation is not a substitute for the "intensional" one, for they are addressed to quite different questions.

> The Darwinian explains what the "intensionalist" postulates: that the mouse's perceiving the cat as an enemy (or, better, the mouse's genetically programmed behavioral response to cats) is adequate to the order of nature.
>
> Just as with cats and mice, so also with scientific success. It is one thing to explain why some theory is successful and quite another to explain why only successful theories survive. Van Fraassen's Darwinian explanation of the latter can be accepted by realist and antirealist alike. But to say that only successful theories are allowed to survive is not to explain why any particular theory is successful (ibid. p. 1100).

In any case, we can only select for successful theories based on observable data currently available to us. The evolutionary answer could do nothing to explain why scientific theories are successful in predicting phenomena currently unknown to us. In our above scenario, T_2 supplanted T because it allowed for o_2 whereas T did not. This, however, occurred only when o_2 became known to us. In realist terms, it would be miraculous for T_2 to ever predict o_3, something that will only become known in the future, unless T_2 turned out to be true. Therefore, since we routinely use scientific theorizing to make novel predictions, Van Fraassen's evolutionary explanation, as presented, will not do to answer the miracle argument.

None of this is to say that an anti-realist answer is impossible. Rather, two points should be taken from what has here been discussed. The first is that the anti-realist explanation of the success of science is not, *prima facie,* easy. Secondly, the appeal of scientific realism largely comes from the easiness of its answer to the problem. We have, therefore, in explaining the pull towards scientific realism, a criterion for realist theories: a realist account that does not explain the success of science is no more desirable for realists than an anti-realist theory.

2.4 Moving Forward

Going forward in our discussion of extensional realism, we have several things to keep in mind. We have a general characterization of realism, helping us to identify the roots of the view and how it stands apart from other realist theories; We have the chief hurdle it seeks to overcome: the pessimistic meta-induction; We have a criterion of evaluation regarding how well it, and its rivals, conform to the miracle argument by explaining the success of scientific theory. Now that the history of scientific realism has been covered, we are best situated to see what extensional realism contributes to its future.

Chapter 3

Extensional Realism

3.1 Extensional Realism: an Overview

As stated in the previous chapter, the chief realist strategy for the defeat of the pessimistic meta-induction is to identify some commonality across scientific theories throughout time, which we can point to and declare (approximately) truthful. In the case of De Haro's *extensional realism* (2020), this commonality is the extension(s)– i.e. reference(s)– of our terms, though *not necessarily* their intension(s)– i.e. sense(s).

The story runs as follows: a previous theory T_1 might make use of some term a, which refers, in certain domains and under the right conditions, to object x. The theory of today T_2, however, has abandoned the use of a, finding it to be conceptually problematic, failing to refer in key domains. Instead, term b is used. Despite this, we need not resort to the pessimistic meta-induction to conclude that T_1 was wantonly mistaken about x through its use of a. Instead, we can find that, in relevant domains and under the appropriate conditions, both a and b, despite being conceptually disparate (i.e. having different senses) referred to x. We say that, in such cases, a and b are *extensionally equivalent*. In this way, both T_1 and T_2 posses a degree of approximate truth. T_2 has replaced T_1 simply because new domains of inquiry have opened up, wherein b refers and a does not. With a proper notion of approximate truth that accounts for this (as shall be discussed presently), T_2 is shown to be more approximately true than T_1.

De Haro gives several examples where he believes this has occurred (manuscript b, p. 3). For instance, "gravitational force", as used in the Newtonian mechanics has been supplanted, in some domains, by "space-time curvature" in general relativity. It can, in fact, be shown that in these areas the former reduces to the latter, and so these are cases in which we can say the two are extensionally equivalent. General relativity has upseated Newtonian mechanics by using "space-time curvature" rather than "gravitational force" only in a select few areas. Indeed, for pragmatic purposes, it is still more convenient for some tasks to use the concept belonging to "gravitational force" than that belonging to "space-time curvature"; simple calculations predicting the movement of objects on Earth will find the first concept much more helpful than the second, even though the second can here be reduced to the first. From this, the extensional realist concludes that there is enough commonality of reference in the right domains, under the right conditions, to be realists about the extensions of both "gravitational force" and "space-time curvature". General relativity is more successful using its term in more domains than Newtonian mechanics using "gravitational force", and so the first theory is more approximately true. Both, however, find approximate truth in some areas, and so the pessimistic meta-induction loses its sting in this particular instance.

To advance extensional realism, two things are required. The first, which shall be detailed in the next section, is a general account of *how* the terms of disparate theories can be found to be extensionally equivalent in some areas but not other. The second, detailed in the section thereafter, is an account of approximate truth that uses this fact to defeat the pessimistic meta-induction.

3.2 Intensions and Extensions

Extensional realism is built upon a *causal-descriptive* theory of reference. This is a bridge theory between *causal* theories of reference and *descriptive* theories.

According to the second of these, referring terms are truncated descriptions, and a term successfully refers in a given context whenever there is an object or objects that satisfy that description. This view, often attributed to Russell (1911), holds that the meaning of the name "Joe Biden" in a given context is equivalent to a definite description, like, for example, "the current president of the United States". That there is an object in the world satisfying this description, "the current president of the united states" explains why the name "Joe Biden" refers. Were it that I, using the name "Joe Biden" as described, was mistaken in believing the United States has a president, it would be that I had failed to refer to anything by using the name.

As an analysis of scientific terms, a descriptive theory of reference would, by the pessimistic meta-induction, threaten that we cannot be justified in believing our language refers. We might, for instance, say that the term "electron," in a particular context, is equivalent to a description like "the negatively charged constituents of atoms". By the pessimistic meta-induction, we have no surety that our beliefs about what is and is not negatively charged are correct, or even that the concept of "negative charge" applies to anything at all. As such, we have no surety that "electron" refers.

On the other hand, there are *causal* theories of reference. Kripke (1980) is the leading modern proponent of such views. According to such views, the story of reference is a two-part process. In the first place, an object gets its name by an *initial baptism*, whereby some person confronted with the object in question gives it a name by which that person can refer to said object. If my subsequent use of that same name is part of a causal chain leading back to the initial baptism, then I have successfully referred to the same object. In line with our above illustration, causal theories of reference will have it that at some point, perhaps at the point of his birth, the parents of Joe Biden will have, after some interaction with him, given him the name "Joe Biden". This will have caused others to refer to his son by that name, which will in have done the same for others, and so on and so on. The success of my use of the name "Joe Biden" to refer to the current president of the United States is to be explained by my use of the term being the product of the chain of events that began with the initial baptism– i.e. the president's parents giving him his name. My own beliefs as to what descriptions Joe Biden satisfy have no bearing on whether my use of his name successfully refers.

In the context of scientific language, causal theories fare somewhat better against the pessimistic meta-induction. It might be found that our beliefs about electrons are radically untrue; that there is no thing to which the description "the negatively charged constituents of atoms" can properly be applied. The descriptive information our theories give about electrons, however, play no part in whether the term refers according to the causal picture. So long as we can trace our use of the term back to a proper initial baptism, which itself requires only some causal interaction with electrons, not any true theory of what they are, then we can safely hold that "electrons" refers.

Whilst this is a happier explanation, it is not by any large degree. Though reference for scientific terms has been saved, the causal theorists who takes this route has opened herself up to the objections of Stanford raised in the previous chapter. Since reference is devoid of descriptive content, establishing that it survives belief change says nothing useful to furthering scientific realism, and the question of whether or not a given term refers has been reduced to questions of mere happenstance.

A bridge theory would have the merits of descriptive and causal theories. From the former, it would take the advantage questions of reference are substantive questions, and so relevant to our scientific beliefs and, by extension, the question of scientific realism. From the latter, it would take the advantage that reference is not fixed by some general description, and so easily endangered by the pessimistic meta-induction. David Lewis calls such theories *causal-descriptive*: "descriptivism, global or local, in which descriptions are largely couched in causal terms" (1984, p. 226). This is a happy-medium in that it "[w]hen causal theories work, causal descriptivism does, too. When not, we need mixed theories, halfway houses between 'the new theory of reference' [i.e. a causal theory] and the old [i.e. a descriptive theory]" (ibid p. 227). This is the approach taken by De Haro.

The move begins with an appeal to classical semantics, according to which the meaning of a term is divided into an intension and extension. The intension of scientific terms are general descriptions. For instance, we might say that the intension of "electron" is something to the effect of "negatively charged particle with spin 1/2". The extension of a term– i.e. that term's reference– is determined by its sense alongside a context of utterance and circumstances of evaluation given by scientific theoires. The circumstances of evaluation are descriptive, and yet use causal terms– e.g. a place and a time. The key move made is that extensions are domain-specific. This means that a term t can refer to object x in one domain of scientific discourse, and y in another. This leads De Haro to encapsulate extensional realism in two points:

> (1) The fact that extensions are conceptually rich, and suffice for a realist commitment to both observable and unobservable entities. (2) The recommendation that, for ordinary scientific theories, one should expect intensions to fail, and that prudent scientists are very well aware of this (this awareness neither requires nor implies any pessimistic meta-induction). And thus, that there is no reason why one should have expected intensions in general to agree, or be committed to them (independently of circumstances of evaluation and context of utterance), in the first place. The differences in intension that are not tied to appropriate circumstances of evaluation and context of utterance are irrelevant, since the cautious scientific realist is not committed to them anyway (manuscript a, p. 15).

We are to be realists about extensions, but not intensions. That is to say that we can reliably expect our terms to refer, and construct a notion of approximate truth from this fact, but can expect, for reasons of the pessimistic meta-induction, intensions to fail as general descriptions outwith a context of utterance and circumstances of evaluation. We can, therefore, (supposing our science has been properly conducted) expect the term "electron" to refer. What is referred to will vary according to domain, and there are likely domains in which the term has no reference. The general description, "negatively charged constituent of an atom", however, is not something we can, taken by itself, accept in a realist-friendly way.

More needs to be said to make this clear, specifically the role of circumstances of evaluation, contexts of utterance, and domains.

Of the first, it is said that circumstances of evaluation are to be equated with *circumstantial physical conditions*: "[t]hey are, for example, boundary conditions and other conditions about the empirical situation considered: including measurement or observational accuracies, the approximations made, in so far as these correspond to physical properties of the system considered (for example, low speed, or large orbital angular momentum, etc.)" (2020., p. 24). Where a domain is a possible world, as commonly understood in modal logics, each domain will have its own circumstantial physical conditions, as each world will have different physical occurrences therein.

The circumstantial physical conditions of a world (or domain) are determined by a *level of abstraction*.[1] This is a technical component of a scientific theory that determines physical states, by being composed of quantities of states and quantities relevant to the theory in question. De Haro provides an (inexhaustive list) of the constituents of a level of abstraction:

[1] Taken from Floridi, 2011

- The specification of a model, viz. the relevant kinematic and dynamical situation (including choices of initial and boundary conditions, velocities, energies, distances, assumptions about the population, environmental conditions, chemical composition, etc.).
- The values of the theory's and the model's free parameters (such as mass, coupling strength, chemical concentration, etc.).
- Extra-theoretical facts, in so far as these are modelled by, or incorporated into, the theory or model: including the allowed measurement or observational errors; other influences, such as noise, not described by the theory; aspects of the experimental— including material—conditions, weather conditions, etc.
- Approximations and idealisations (both theoretical and empirical) carried out in order to accommodate all of the above, i.e. to make the theory empirically adequate under the stated conditions and accuracies (ibid., pp. 25).

Where these are indices, the set of them together with assigned values composes a level of abstraction belonging to a theory. Changing the values of each index creates a new level of abstraction, and thus determines new circumstantial physical conditions, and so a new possible world.

As for the context of utterance, this is to be understood as a *context of application* (ibid., p. 24). Where a level of abstraction is descriptive, the context of abstraction contains causal aspects. They include "scientific methods and practices, the material preparation and composition of samples, experimental procedures, ostension and intentions of speakers, influences (or lack of them) that are not described by the theory, etc" (ibid.). It is in this area that descriptivism turns into causal descriptivism, since the likes of experimental procedures are not linguistic, but instead part of a series of causal processes of scientific investigation.

With this all established, the role of extensions comes in. Whilst we have mainly talked of extensions as belonging to terms, which amounts to the objects and relations in the world picked out by the components of a sentence, it should be kept in mind that what is said applies also to the extensions of sentences, which amounts to their truth values. The importance of extensions to the development De Haro's view is that terms can be *extensionally equivalent.* Two terms are extensionally equivalent with respect to a domain when they have the same extension in said domain. Note that for two terms to be extensionally equivalent, it is not required that they share the same intension. The intension of "forty-sixth president of the United States" is clearly different than that of "forty-seventh vice-president of the United States" though the extension of both is Joe Biden.

De Haro gives several examples of how extensional equivalence can come about in scientific inquiry. The first is *conceptual correspondence*, which comes about when two distinct concepts play the same conceptual role as one another under the right causal circumstances:

> Another example is the electron in the hydrogen atom. The correspondence between quantum and classical mechanics usually requires, as I mentioned above, setting some quantity to a large value, compared to Planck's constant h, i.e. doing an approximation or taking a limit. For the hydrogen atom, Bohr's correspondence principle entailed taking the limit of orbits with large quantum numbers (in units of h); in that limit, the classical orbital frequency of the valence electron coincides numerically with the frequency of the emitted radiation, and the electron can be treated as a charged oscillating classical source. Again, there is no intensional equivalence, but there is an extensional equivalence of the quantum and the classical system, in the limit considered. That is, the electron of the Bohr model of the atom is, pace Kuhn and Feyerabend, extensionally equivalent with the electron of the classical theory of radiation: the physical systems, and the possible experimental situations, to which these models can be applied are the same, even though intensionally they are of course very different (manuscript a., p. 28).

In this example we see that since, when numerical values are tweaked appropriately, concepts that are otherwise distinct, in this case two differing concepts of "electron", end up playing the same theoretical role, and can thus be said to be extensionally equivalent in the described incidence. Since numerical values belong to levels of abstraction, we say that adjusting the value has created a new level of abstraction, generating a domain (world) in which two terms are found to have the same extension. *Conceptual correspondence* thus highlights how levels of abstraction work for extensional realism by creating extensional equivalence.

On the causal side of things, we have *material correspondence.* De Haro points to the former at work in the case of mass in special relativity versus classical mechanics. There is material correspondence when the procedure used to measure both is the same, though they intensionally differ:

> Thus a single sequence or combination of experiments suffices to determine the mass in both theories, because of the material continuity between them, without which mere numerical agreement would be insufficient—which again highlights the importance of the context of application in establishing extensional equivalence. In other words, without identity of the bits of matter stuff one would not know which numbers one should measure (2020., p. 39).

Since one's method of measurement is not a linguistic or descriptive choice, but a physical activity belonging to the causal world, we can see this as the causal element of causal-descriptivism at work in determining reference, arising from the context of application.

There are other modes of correspondence that lead to extensional equivalency, such as *predictive correspondence* and *formal correspondence*, but this suffices for explanatory purposes.

With these concepts elucidated, we can begin to sketch the answer to the pessimistic meta-induction. We must hold that, in domains where previous theories were empirically successful, it was because of their extensional equivalence to the theories of today which supplanted them. New theories, then, are able to explain the success of the old, just as we can reasonably hope the theories of tomorrow can explain the success of those of today. A theory of approximate truth that captures this is presented as the hope for scientific realism.

3.3 E-Truth, I-Truth and Approximate Truth

In order to explicate approximate truth, De Haro first defines two applications of the notion of truth relative to any given scientific theory T, extensional truth (E-truth) and intensional truth (I-truth). According to the former:

> a sentence is **E-true**, relative to a possible world, iff its truth value, in that possible world allowed by the theory (with its specific circumstantial physical conditions), is 1 (2020, p. 46).

and according to the latter:

> a sentence is **I-true** iff it is true in all possible worlds to which the theory can be applied (under all possible circumstantial physical conditions), (ibid.).

For a truth value to be taken as 1, we can simply take to mean "true under the Tarskian disquotational view of truth".[2] We might then read the definition of E-truth as like to "for all sentences φ, "φ" is E-true relative to a possible world by the theory allowed by T iff, at w, φ. Since domains– i.e. possible worlds– are possible worlds, whose circumstantial physical conditions are determines by a given level of abstraction belonging to a theory, in order to determine the E-truth of a sentence s in T, it is necessary to state the level of abstraction.

To illustrate, take a theory T_1, which implies sentences s_1, s_2 and s_3, and levels of abstraction L_1, L_2 and L_3. For three worlds, w_1, w_2 and w_3, each sentence is true when evaluated from each level of abstraction at w_1 and w_2, but s_3 is false at w_3 when evaluated from level of abstraction L_3. Since, according to De Haro (2020, p. 48), a theory's intension is equal to the the union of domains in which all of its domains are true at all levels of abstraction. As such, we define the intension of T as $w_1 \cup w_2$. Given this, we say that T is E-true at w_1 and w_2, but E-false at w_3. We can say of T, then, that it is extensionally true in two domains, which we might say are relevant worlds of classical Newtonian mechanics, with low levels of , but extensionally false at w_3, a world of quantum physics, where energy levels are high.

We can thus say the following of T, which has turned out to be a Newtonian theory of physics. It is E-true in the domains $w_1 \cup w_2$ but E-false elsewhere. It

[2] See Tarki (1943). Essentially the disquotational veiw of truth is that a sentence "φ" is true iff φ

is not I-true, by virtue of their being a domain applicable to the theory where the presence of s_3 renders it E-false. It is (contingently) true by extension at some domains but its intension does not universally hold.

With this apparatus in place, it can be made clear what the extensional realist means when she says that the scientific theories of the day are approximately true. When she says this, she, as De Haro puts it, means we can be realists about the extensions of scientific theories, but not necessarily intensions, which means we can hold that our theories are extensionally true in relevant domains, relative to the levels, but that it may well be that our theories are intentionally false.

Compare with the various formulations of the pessimistic meta-induction outlined in the previous chapter. The first, the naive formulation, says that, since the mature, empirically successful theories of the past have turned out false, we can expect the same to happen in the future to our present theories. This we may grant, providing we assume the definition of "true" here is intentional, which is probable. It is very likely that, even if we consider a theory of today to be extensionally true in all relevant domains (which itself seems unlikely), that we shall eventually find some domain in which it is falsified extensionally. We need not grant, however, that our theories are bound to fail extensionally, since we may hold that mature theories of the past remain extensionally true in some domains, thus undermining the induction. For example, Maxwell's theory of the aether was replaced by the theory of electromagnetism. De Haro holds that, in some domains, the two theories are extensionally equivalent:

> [...] I wish to suggest that Maxwell's aether can be identified extensionally with the electromagnetic field itself, together with an appropriate set of Galilean frames of reference, where by 'an appropriate set of Galilean frames of reference' I mean a set of inertial frames of reference that move with low speed (much smaller than the speed of light) relative to earth and to each other. The electric and magnetic fields are indeed invariant under Galilean transformations within reasonable accuracy for the Galilean transformations between these frames, so long as their speeds are much lower than the speed of light, and so it makes sense to identify the aether with the electromagnetic field itself, together with the set of Galilean frames of reference. This is not to say that the two concepts are the same: for they are not. It means that the two terms are extensionally equivalent [...]. If true, this gives us an extensional equivalence between the new electromagnetic theory and Maxwell's old theory (2020, p. 41).

Where two theories are extensionally equivalent in a domain, the E-truth of one guarantees the E-truth of the other therein. So, by the E-truth of the theory of electromagnetism in a given domain, we can be assured of the E-truth of Maxwell's theory of the aether in that same domain, providing it is one wherein the two are extensionally equivalent.

This, in microcosm, is De Haro's notion of approximate truth. Generally, a theory T is more approximately true than T' if it is E-true in more domains.[3] Empirically successful theories are those which are E-true in domains in which they are employed. This notion of approximate truth is not, then, at odds with empirical success simply because theories are continuously replaced, as the pessimistic meta-induction would have it, because there is no need to concede that our old theories were not approximately true in the first place. We need only allow that our new theories are *more* approximately true than their predecessors. In the above, for instance, we need not say that the replacement of Maxwell's theory of the aether by theories of electromagnetism shows Maxwell to have created a falsehood; his theory was approximately true.

We can thus dispense with the naive formulation of the pessimistic meta-induction. By similar argumentation, we can dispense with the *reductio ad absurdum* formulations that threaten the connection between empirical success and approximate truth. Since we need not grant the existence of empirically successful theories that were not E-true, there is no problem to be found in explaining the success of theories by their approximate truth, as the miracle argument describes. So also can we deal with the formulations that target reference. Since we hold that E-true scientific theories are genuinely referring– in specific domains when paired with appropriate levels of abstraction and contexts of application– and empirical success is explained by E-truth, we have a link between empirical success and reference. Extensional realism, therefore, has ready answers to the various forms of the pessimistic meta-induction.

As to a Stanford-style charge that this victory is pyrrhic,[4] such worries do not apply. Stanford's concern was that a second pessimistic meta-induction could arise if we isolated some components of theories about which we could be realists, in that we might be continuously proven wrong about which components to isolate. Therefore, we could not merely claim to be realists about the causally important parts of a theory. This is not applicable in this instance, since extensional realism is merely a doctrine as to how to *interpret* scientific claims within a theory, and makes no distinction between one theoretical sentence and another. There is, likewise, no analogous worry that we are putting words into the mouths of scientists by giving an account of what is important in their own theory that might be at odds with their own. Extensional realism merely says that the terms of scientific theory, if such theories are properly developed, refer in the domains in which they are most commonly used. This can hardly be thought to be at odds with the intentions of scientists, since it makes no value-judgements by itself, nor does it direct us to reinterpret the intensions of scientific terms past what has been offered by the theory in which they are found.

We have, then, in extensional realism, a theory that defeats the various formulations of the pessimistic meta-induction with respect to the extensions of

[3]As shall be discussed later there is also the proviso that the domains in which T proves E-true are more relevant than those where T' proves the same. More shall be said on this later.

[4]See Stanford 2003

our language. It does so in a way that preserves reference. Note also that it accomplishes this in a substantial way, by tying extensions to intensions, which are theoretically significant. This avoids the complaints of Stanford (2015), which claim that purely causal theories of reference make preserving the extensions of our terms philosophically vacuous. Extensional realism appears to do everything asked of a realist theory of science. More needs to be said, however, to differentiate it from already existing realist theories, namely *structural realism.* This is the task of the following chapters.

Chapter 4

Structural Realism and Ramsey Sentences

4.1 Structural Realism: an Overview and Comparison

Among the kinds of Scientific Realism espoused today, one of the more prevalent is *structural realism.* Recall from chapter one that the strategy for defeating the pessimistic meta-induction is to identify some commonality across scientific development about which we can be realists. In the case of extensional realism, this commonality is the extensions of our scientific language, which can survive theory change in defiance of the meta-induction, as in the example of the extension of Maxwell's "aether" surviving the shift to talk of electromagnetic fields. In the case of structural realism, the commonality to be found is in the *structure* of the world described by a theory.

So understood, we take a scientific theory T to be a description of the world, or some fragment of it, detailing entities both observable and unobservable. In the case of unobservable entities, we cannot be realists about what our theory has to say about them intrinsically, but we can be realists about what T has to say with respect to their place in the structure of the world, namely their relations with other objects in the world.

There are many ways to flesh out this position. For illustrative purposes, we shall take John Worrall's influential formulation (1989). Worrall aims for a "best of both worlds" approach to navigate between the pull of the no miracle argument on the one hand, and the pessimistic meta-induction on the other. The no miracle argument appeals to Worrall because, at the empirical level, scientific theories have met with great success and have endured little alteration. However, at the non-empirical level, which deals with entities that cannot be observed, theories continuously go through great revolutions, which is suggestive of the pessimistic meta-induction. Worrall held that the structure of the unobserved world survived these revolutions, however, and could be used to explain the success of science whilst admitting to the continued change of theories on a non-structural level. The resultant realism about structure was, in Worrall's view, the great bridge between these two competing problems.

As for the question of what counts as a theory's structure, that is given by the mathematical aspect of a theory. Worrall takes this as being present in Poincare (1905). Accordingly, the equations used to denote the relations between entities tend to survive theory change, emerging in their successors. This Worrall exemplifies in the case of the shift from Newtonian mechanics to Einsteins relativity:

> The rule in the history of physics seems to be that, whenever a theory replaces a predecessor, which has however itself enjoyed genuine predictive success, the 'correspondence principle' applies. This requires the mathematical equations of the old theory to reemerge as limiting cases of the mathematical equations of the new. As is being increasingly realised, the principle operates, not just as an after-the-event requirement on a new theory if it is to count as better than the current theory, but often also as a heuristic tool in the actual development of the new theory [...] But the principle applies purely at the mathematical level and hence is quite compatible with the new theory's basic theoretical assumptions, which interpret the terms in the equations, being entirely at odds with those of the old. I can see no clear sense in which an action-at-a-distance force of gravity is a "limiting case" of, or "approximates" a space-time curvature. Or in which the 'theoretical mechanisms' of action-at-a-distance gravitational theory are "carried over" into general relativity theory. Yet Einstein's equations undeniably go over to Newton's in certain limiting special cases. In this sense, there is "approximate continuity" of *structure* in this case (1989, pp. 120-121).

The correspondence principle here invoked requires that structure is preserved throughout theory change by requiring the preservation of mathematical equations, which are merely structural in that, lacking an interpretation, they state nothing more than the relations between variables. Consider, for example, the equation $f(x) = f(y) + 1$. We might have a (non-scientific for the purposes of a simple explanation) theory T_1 about the make up of my house. Constitutive of T_1 is interpretation that holds the equation is true since f is, as a function, assigned the value of "the number of sockets in room x"; x is assigned the value of "living room"; and y the value of "bedroom". The equation, as interpreted by T_1 says, then, that I have one more socket in my living room than my bedroom. T_2 is a theory about my neighbour's house. It contains a similar interpretation that also makes $f(x) = f(y) + 1$ true. According to this interpretation, f is assigned the value "the number of lights in room x"; x is assigned the value "my neighbour's living room"; and y the value "my neighbour's bedroom". T_2 thus says that my neighbour has one more light in their living room than they do their bedroom. Obviously, the claim "I have one more socket in my living room than in my bedroom" is distinct in content from the claim "my neighbour has one more light in their living room than their bedroom". However, we can say that, in the context of their respective theories, the two claims are structurally identical at the mathematical level since both are properly stated as $f(x) = f(y) + 1$. Two claims, belonging to two different theories, are structurally identical when they are expressed mathematically by the same equation, and the difference in their linguistic content is to be explained by the difference in how the two theories interpret the terms of the equation.

That is what Worrall points to as the continuity of theory structure. Generally, when one theory replaces another, the mathematical equations of the former are found in the latter, but they nonetheless express different facts because the latter theory interprets the equations differently. This is why Worrall can point to continuity in the transition from Newtonian mechanics to the relativity of Einstein. There is no way in which gravity "approximates" a space-time curvature simply because a proper description of the one would be significantly different from a proper description of the other. The principle of correspondence, however, says that the equations used to describe the force of gravity by the Newtonian physicist can be found also amongst the equations of the relativity theorist, under interpretations that turn talk towards the new concept of space-time curvatures.

This, then, is the "best of both worlds" answer to the pessimistic meta-induction and the no miracle argument. As to the first, there is agreement that our beliefs about the objects to which we attempt to refer in our scientific theorizing change radically over time. This is established by the interpretations of our theories not tending to survive theory change– gravity does not approximate a space-time curvature according to structural realism. We cannot, therefore, be scientific realists in an unrestricted sense because of a pessimistic meta-induction targeted at theory interpretation. The success of science is nonetheless not a miracle, since, at the mathematical level, which is theoretical rather than empirical, there is commonality across theories via the survival of mathematical formulae. Our ability to use these formulae to achieve empirical success entitles us, so the structural realist says, to believe in their truth. Therefore, since mathematical formulae devoid of interpretation are structural, we may be *structural realists*.

In comparing structural realism and extensional realism, we find several points of departure in formulation. In the first case, whilst both ascribe commonality to scientific theories across time, and do this by identifying correspondences between theories, structural realism admits to only one kind of correspondence: mathematical correspondence. Two concepts mathematically correspond to one another when they play the same role in the same equations. This suffices for the survival of structure, but De Haro goes further. He identifies many more kinds of correspondence, as identified in the previous chapter.

This comes into play when Worall says he can see no way in which gravity approximates a space-time curvature. In a sense, there is agreement between the structural and extensional realist on this point, since extensional realism admits that the intension of the two terms differs substantially. However, extensional realism is a theory about scientific *concepts* rather than the objects of scientific study themselves. The extensional realist can make the claim that, in the right context– i.e. in the right levels of abstraction and under the right material conditions– the *concept* of "gravity" as found in Newtonian mechanics approximates the *concept* of "space-time curvature" insofar as the two are extensionally equivalent. This allows for the extensional realist to hold to the truth of non-mathematical sentences at an extensional level. In particular levels of abstraction, descriptive claims about gravity can be accepted by the extensional realist that are not structurally identical to claims the structural realist would accept independently of given levels of abstraction. This, De Haro points out, is because, where structural realism does not take levels of abstraction into account, from motivation to semantics extensional realism embraces them:

> Extensional scientific realism thus begins with the consideration of the level of abstraction (with the circumstantial physical conditions that it determines) and, where relevant,a context of application. And since these notions apply in all of semantics—the semantics of scientific theories being no exception—there is no reason why the realist should stop at structure: something that, by itself, is difficult to define without the mention of concepts and material procedures (De Haro, 2020, p. 17).

There are other differences as well, of course. For instance, Worrall motivates his brand of realism by appealing to the difference between the observable and unobservable. This distinction, as was pointed out in chapter one, has in the past been criticized, and it might well be thought an advantage that extensional realism makes no appeal to it and does not obviously require the distinction to hold true.

All this said, there are apparent distinctions enough to properly differentiate extensional realism from structural realism as formulated by Worrall. Since Worrall's work is so foundational to structural realism, it might be thought that this is all that needs be said on the matter. It shall be shown, however, that there is some formulation of structural realism that demands more from De Haro in order to avoid his theory collapsing into this variant.

4.2 Structural Realism and Ramsey Sentences

The kind of structural realism that presents a challenge to De Haro's extensional realism comes from Grover Maxwell. Maxwell (1962, 1970) claims that we can only be justified in believing scientific claims as true after the elimination of terms referring to unobservable entities and their intrinsic properties from our language. This comes from a reading of Bertrand Russell (1911), according to which we ought not use terms that directly refer to objects with which we are not acquainted. What Russell, and philosophers in general, mean precisely by being "acquainted" with an object is notoriously hard to explicate precisely. It suffices for explanatory purposes to say that we become acquainted with an object when it is part of our immediate sensual experience.[1] A term refers directly when it does not do so by description. I can, for example, refer to the tallest man in London by the description "the tallest man in London" because I aim to talk about whichever object satisfies this description. Therefore, the sentence "the tallest man in London is taller than me" indirectly refers, assuming there is some man in London taller than all others, because it is about a particular individual, despite the lack of a name for this person. To refer directly to the tallest man in London, I would use a non-descriptive item, such as the name "Steve" in the sentence "Steve is taller than me". The principle Maxwell takes from Russell, of which it is questionable whether Russell actually advocates is that I ought not attempt to directly refer to things with which I am not acquainted, hence, since I have never met the tallest man in London, I must use sentences of the former kind in place of the latter. So too does this rule apply to scientific talk. Since I am, by definition, not acquainted with unobservable entities, I can accept no sentence directly referring to them. By that same token, I cannot also accept as true any sentence refering to the intrinsic properties of unobservable entities, since I have not been acquainted with them either.

Right away, it seems dubious how any of this could begin to pose a problem for extensional realism. After all, as has been stated, extensional realism does not hold to the observable/unobservable entity dichotomy. Further, nowhere in De Haro's account of reference is it claimed we cannot directly refer to entities to which we are unacquainted. Whilst this is true, and so extensional realism cannot be said to be *in principle* the same as something approximating Maxwell's structural realism, if the two entail acceptance of the same set of sentences (and interpret the sentences in the same way) then, whilst the motivations of the two positions may differ, the two have the same practical consequences for what scientific claims we can claim to be true.. Since extensional realism is offered as a new kind of scientific realism, this is something to be avoided. To see how this might come about, what kind of sentences the Maxwellian structural realist would accept, and the problems for extensional realism originating thereby, we require an explanation of *Ramsey sentences*.

Ramsey sentences, introduced by Frank Ramsey (1929) and popularized by David Lewis (1970) serve as a way of eliminating theoretical terms. We shall draw mainly from Lewis in our explanation. Accordingly, non-logical terms are assigned to one of two categories: *t-terms* and *o-terms*. A t-term is a term introduced by our new theory from which we are attempting to derive a Ramsey sentence. "Space-time curvature" is a t-term in the theory of relativity, for instance, as is "electron" in atomic theory, since both terms were introduced by the theory in which they were found. T-terms, therefore, can be thought of as theoretical terms. Conversely, o-terms are terms for which we had an understanding prior to scientific theorizing. The category of o-terms is vast, and

[1]It was held by Russell that we could be acquainted with our own mental events. Whether this strictly falls under the umbrella of occurring within our sensual experience is a question for philosophy of mind and of no relevance here. We might well be working with a more restricted account of acquaintance than some philosophers have in mind, but excluded only are cases not relevant to the questions of this thesis.

contains most of our ordinary language terms. O-terms, then, can be thought of as "other terms" as opposed to theoretical terms. O-terms do not include logical terms, of the kind "and, or, not, etc..." that are usually dealt with in Boolean algebras.

Ramsey sentences eliminate t-terms via existential quantification. Take an arbitrary theory T. T is composed of many sentences t_1, t_2, t_3... The conjunction of all sentences of T, $t_1 \wedge t_2 \wedge t_3 \wedge ...$ is T's *postulate sentence*. To utter the postulate sentence of T is to enunciate the entire theory. T's Ramsey sentence, free of t-terms, is a logical derivation of the postulate sentence where all instances of t-terms are replaced by variables bound under an existential quantifier quantifier $\exists$.

To see how this works, consider what we can call a *t-sentence*, which is a sentence containing at least one t-term. The sentence "water is composed of hydrogen atoms and oxygen atoms" is a t-sentence since it contains the t-terms "hydrogen atoms" and "oxygen atoms". We can logically encode the sentence by saying that the predicate C denotes the relation "x being composed of y and z", w is the encoding of the term "water", h "hydrogen atoms" and o "oxygen atoms". Thus "water is composed of hydrogen atoms and oxygen atoms" becomes $Cwho$. From a t-sentence, one can derive an o-sentence, one containing only o-terms and logical terms by existential quantification over all t-terms. In this example, we have identified "hydrogen atoms" and "oxygen atoms" as our t-terms, and so, from $Cwho$, we derrive the logical consequence $\exists x, \exists y[Cwxy]$. In plain English, this would translate to "there exists at least one thing x, such that there exists at least one thing y, such that water is composed of x and y." In this way, we have replaced the t-terms in the t-sentence "water is composed of hydrogen atoms and oxygen atoms" with the purely logical terms "exists, x and y", resulting in an o-sentence containing only o-terms and logical terms.

To derive a Ramsey sentence from T's postulate sentence, we apply this process to all the t-sentences that make up the conjuncts of the postulate sentence. From $t_1 \wedge t_2 \wedge t_3 \wedge ...$, we derive the sentence $\exists x_1 \exists x_2 \exists x_3 ...[o_1 \wedge o_2 \wedge 0_3 \wedge ...]$, where, for every number n, o_n is t_n but with all t-terms replaced by variables. The Ramsey sentence of a theory, therefore, is a conjunction of o-sentences, containing variables bound by existential quantifiers. We have eliminated terms like "electrons" and "space-time curvature" and instead use terms like "thing x_1" and "thing x_2". In so doing, however, much of the descriptive content of the original sentence has been lost, since the intensions of "electrons" and "space-time curvature" are no longer invoked.

Having explained the process of deriving Ramsey sentences, we can explicate Maxwell's structural realism. Maxwell takes the theoretical terms of science to be those denoting unobservable entities and their intrinsic properties– i.e. given names for unobservables like "electrons" and the one-place predicates to which we apply them, such as "negatively charged". Whilst we cannot be realists about postulate sentences containing such terms, we can regarding corresponding Ramsey sentences. Hence, we cannot accept the truth of the t-sentence "electrons are negatively charged" (Ne), but we can accept the truth of the o-sentence "there exists at least property X, such that there exists some object x such that x is X" $(\exists X, \exists x[Xx])$.

This is a form of structural realism since, though we cannot be realists about the intrinsic nature of unobservable entities, because that would require giving them names and using one-place predicates, in violation of Maxwell's rule regarding direct reference to things with which we are not acquainted. Nevertheless, we can be realists about the *existence* of some unobservable entities and how they relate to one another, as this does not require the names of any unobservables or the application of one-place predicates. This is structural, since it is, similarly to Worall, talk of variables without interpretation and how they relate to one another. Rather than talk about the nature of electrons and atomic nuclei, we talk about how thing x_1 relates to thing x_2, similarly to what might be done in a mathematical equation lacking interpretation. Where Worrall is a realist about the mathematical structure of the universe, Maxwell is a realist about its logical structure.

4.3 Extensional Realism and Ramsey Sentences

With all this established, we turn to the implications for extensional realism. Extensional realism holds that the scientific claims we should adopt are the ones that are E-true at more worlds and in domains of greater significance. The relative question of a domain's significance is left to the judgement of individual scientists:

> On this approach, one recognises that theory choice involves judgment, and that many qualitative arguments and considerations can be used that give reasonable verdicts, thus illuminating the history of science. Such qualitative arguments involve something like a quantitative reasoning: but more in the way of making a list of significant reasons for and against a general argument, i.e. without the pretence of the procedure being exact or comprehensive.

If theory T' is E-true at more worlds than T, then, assuming the worlds in which T is E-true are a subset of those worlds in which T' is E-true, then T' is to be adopted over T. Where this is not the case, we apply the qualitative judgements spoken of above to determine whether the T' are of more significance than the T worlds to determine which theory to choose. It shall be argued that this threatens to collapse extensional realism, in practice (though not principle), into an extreme structural realism.

In the first place, it is clear that the extensional realist must be committed to the E-truth of the o-sentences corresponding to all t-sentences sentences they believe to be E-true at any given world (domain). This follows as such o-sentences are logical consequences of corresponding t-sentences, and so the latter cannot fail to be E-true where the former succeeds, no matter the level of abstraction or context of application. Therefore, since any given theory's postulate sentence is just a conjunction of t-sentences (and perhaps some o-sentences), the corresponding conjunction of o-sentences that is its Ramsey sentence must be E-true in at least as many worlds as the postulate sentence.

From here, we can go a step further and say that, for any given postulate sentence that is not I-true, there is a corresponding Ramsey sentence E-true in more worlds than the postulate sentence is. It is clear that this cannot be the case for I-true postulate sentences, since they are E-true at all worlds, and so a world cannot be found in which the Ramsey sentence is E-true and the postulate sentence not. Since De Haro claims we are not entitled to believe our theories I-true, however, such theories are not relevant. Where a theory is not I-true, depending on what terms are taken to be t-terms, a Ramsey sentence can more easily be made E-true.

Take now what we shall call *extreme logical structural realism.* This takes it that all terms that are not logical terms are t-terms. From this, the extreme logical structural realist takes it that, not only can we not be realists about postulate sentences, we are only permitted to be realists about Ramsey sentences which consist entirely of existentially quantified variables and logical connectives. Intuitively, such sentences are easy to make E-true.

Consider, to exemplify this, the arbitrary Ramsey sentence $\exists x_1 \exists x_2 \exists X[Xx_1x_2]$– there exists at least one thing x_1, such that there exists at least one thing x_2 such that there exists at least one relation X, such that x_1 is related by X to x_2. This sentence can be shown to be true at any world consisting of at least two objects and one relation.[2]

Indeed for any postulate sentence T which refers to some number of objects and relations, one can imagine infinite worlds with the same number of objects structured in such a way as to make E-true a Ramsey sentence the extreme logical structural realist would accept, which we can call an *extreme Ramsey sentence.* A strict subset of these worlds, in the case of non-I-true theories, are the worlds in which the postulate is E-true, since the extreme Ramsey sentence is a logical consequence of the postulate sentence.

Given this, and De Haro's account of scientific progress, we are threatened by a collapse from extensional realism to extreme logical structural realism. Extreme Ramsey sentences are E-true at more worlds than postulate sentences, as well as E-true in all the worlds in which the postulate sentence is E-true. The domain of the postulate sentence cannot be more useful than the domain of the extreme Ramsey sentence, since the domain of the latter is contains the former.[3] Science should be progressing, according to this picture, to the development of extreme Ramsey sentences, an absurd notion given the emptiness of their content, and certainly not what we want from extensional realism.

Even if one does not go so far as extreme logical structural realist, there

[2]In fact, it can also be shown to be true at any world with one object, where the relationship substituted by X is self-identity.

[3]It might here be objected that this is not the case, because the additional worlds of the extreme Ramsey sentence are somehow unnecessary baggage and thus make the Ramsey sentence less relevant. Were this to hold, however, than the domain of an I-true sentence must always be thought to be irrelevant, being E-true in all worlds. This is an unhappy line to take as I-true claims are supposed to be the golden standard for which we strive but can never reach.

is still the general threat of some form of structural realism based on Ramsey sentences not so extreme. A sentence involving the term "electron" requires for its E-truth electrons to behave in a certain way. That same sentence, where "electron" is replaced by "thing x" requires only that any old thing behave in that way. It is much easier to make this sentence true, and thus E-true, and so it is probable that there is some world where the second sentence is E-true and the first not.

The challenge for extensional realism is to find a way to compare theories that blocks this. Ideally such an account would say science progresses to sentences that are more approximately true weighted against their informativity. The problem with the Ramsey sentences described above is that they are uninformative, especially in the case of extreme Ramsey sentences, which are so nebulous as to say almost nothing about the world. The solving of this challenge along these lines is the subject of the next chapter.

Chapter 5

Choosing Between Theories: Verisimilitude and Truthlikeness

5.1 Verisimilitude

The problem outlined in the previous chapter is the problem of how to choose between competing theories when relative truth alone will not do. The danger for extensional realism is that the best theories are those that are the most approximately true, and the most approximately true theories are the most general or nebulous– *i.e.* the least useful. To solve this problem, we need a theory of what Karl Popper calls *verisimilitude*, which weighs theories both by their truth and their content. With this understanding, the value of a theory is a measure of its closeness to truth and the useful consequences we can derive from it.

Popper was not himself a realist. Rather, he was, as he termed it, a *falibalist* (1963). To a falibalist, science is not in the business of putting out true theories (for all the reasons discussed in chapter one). Rather, science puts out theories that have not yet been falsified. We select our best theory that lacks refutation, and hold to it until such refutation is found, in which case it is replaced by a subsequent theory until that, in turn, can be falsified. The question then becomes *how* to determine which theories are the best. Truth is one factor, an arguably puzzling position since Popper denies the truth of scientific theories. According to him, science is (at least in part) a search for truth, but one that is doomed to failure:

> For the fact is that we too see science as the search for truth, and that, at least since Tarski, we are no longer afraid to say so. Indeed, it is only with respect to this aim, the discovery of truth, that we can say that though we are fallible, we hope to learn from our mis-

> takes. It is only the idea of truth which allows us to speak sensibly of mistakes and of rational criticism, and which makes rational discussion possible–that is to say, critical discussion in search of mistakes with the serious purpose of eliminating as many of these mistakes as we can, in order to get nearer to the truth. Thus the very idea of error–and of fallibility–involves the idea of an objective truth as the standard of which we may fall short (ibid., p. 229).

Truth plays a role in the falibalist scheme by being constitutive of falsification. There is no sense in falsifying a theory that never tried to be true in the first place, and we could never call a theory mistaken unless we could recognize where it diverged from the truth. Truth, then, has a role in the falibalist picture, but a merely regulative one. Were science merely the search for truth, under this picture, it would be an idle mental exercise whose theories would have extremely limited value. Rather, a successful scientific theory must also contain interesting truth:

> Thus we accept the idea that the task of science is the search for truth, that is, for true theories (even though as Xenophanes pointed out we may never get them, or know them as true if we get them). Yet we also stress that truth is not the only aim of science. We want more than mere truth: what we look for is interesting truth– truth which is hard to come by. And in the natural sciences (as distinct from mathematics) what we look for is truth which has a high degree of explanatory power, which implies that it is logically improbable (ibid.).

Though distinct in motivation, Popper's goals are symmetric to ours regarding extensional realism. Approximate truth alone leads us to the adopting of extreme Ramsey sentences above all else. A selection process for theories that excludes those that lack interesting truths would be ideal for our situation. Popper calls this the *verisimilitude* of a theory. A theory of maximal verisimilitude would not only be utterly true, it would yield useful consequences. Extreme Ramsey sentences have a low verisimilitude since, though they do well under the first count, they fail on the second.

The explanation of Popper's verisimilitude laid out here draws from Popper (1963, 1976) as well as the criticisms offered by Pavel Tichy (1974). Popper gives two accounts, the first formulated in logical terms and the second in probabilistic terms. Since extensional realism is a semantic theory, and we wish to see what it can incorporate from Popper, we shall focus solely on the first.

We define a language $\mathcal{L}$ as containing a number of expressions $a, b, c....$ Popper's verisimilitude focuses on the *content* of an expression. To see this, he considers the intuitive judgements that bid us pick one theory, t_2, over another, t_1. Popper gives a detailed list, including t_2 explaining more facts than t_1, suggesting more experiments, making more precise assertions, *etc.* All of this is to be found in the content of a theory. Popper thus defines a content function C such that $C(a)$ denotes the content of a, which is the set of a's logical consequences.

With this established, we can define an expression's relative verisimilitude as a measure of its truth-content set against its false-content as compared to that same measurement carried out on another expression. To do this, we define for $\mathcal{L}$, T, which is the set of true expressions in $\mathcal{L}$.. From here, we define a truth-content function $C_T(a)$ such that, for any value of a, $C_T(a) = C(a) \cap T$. $C_T(a)$ thus denotes the set of a's logical consequences that are true. Note that when a is itself true–i.e. when $a \in T-$ it follows that $C_T(a) = C(a)$, meaning that the entirety of a's content is truth-content.

The false-content of a, $C_F(a)$ is defined alongside *relative content*. The relative content of a given b, which we can define as $a \oplus b$, is the content of a that is not found in b, meaning that $C(a) \oplus C(b) = C(a) \cup C(b) \setminus C(b)$. The false-content of an expression a is the relative content of a given the truth-content of a– i.e. $C(a) \cup C_T(a) \setminus C_T(a)$. Tichy (1973, p. 156) proves, given the set of false expressions in $\mathcal{L}$, F, that $C_F(a) = C(a) \cap F$.

To determine the relative verisimilitude of two expressions using these functions, the two must be comparable. Any two expressions a and b are comparable with respect to truth-content iff $C_T(a) \subseteq C_T(b)$ or $C_T(b) \subseteq C_T(a)$, and with respect to false-content iff $C_F(a) \subseteq C_F(b)$ or $C_F(b) \subseteq C_F(a)$. If otherwise, then the two expressions have nothing to do with one another, having such divergent content.

Finally, we can state the definition of (relative) verisimilitude. Popper takes it that, for theories t_1 and t_2, t_2 has more verisimilitude than t_1 iff one of two conditions are met:

> **a.** the truth-content but not the falsity-content of t_2 exceeds that of t_1 OR
>
> **b.** the falsity-content of t_1, but not its truth-content, exceeds that of t_2 (1963, p. 156).

In our notation, we may state this formally as follows. For any two expressions a and b, a has more verisimilitude than b iff a and b are comparable with respect to truth-content and false-content, and either $C_T(b) \subset C_T(a)$ and $C_F(b) \not\subset C_F(a)$, or else $C_F(a) \subset C_F(b)$ and $C_T(a) \not\subset C_T(b)$.

Popper provides, then, a way of deciding between theories according to how truthful (and false) the content of these theories are found to be. We can say, for example, that a theory that a true theory is of more use than a tautology because its truth content is higher and falsity content is not. Consider, for instance, the theory t as opposed to the tautology $\neg(t \wedge \neg t)$. Since $\neg(t \wedge \neg t)$ is a tautology, and is thus a logical consequence of t, it trivially follows that the two are comparable with respect to truth-content and false-content. It follows further that $C_T(\neg(t \wedge \neg t)) \subset C_T(t)$, since the entirety of the former's content is entailed by the latter, and that, since $C_T(\neg(t \wedge \neg t)) = C_T(t) = \emptyset$, $C_F(\neg(t \wedge \neg t)) \not\subset C_F(t)$. Therefore, thee theory t is of a higher verisimilitude than the tautological instantiation of the law of non-contradiction. It might, therefore, be hoped that Popper's verisimilitude neatly maps onto an extensional realist framework to solve our problem, the idea being that extreme Ramsey sentences must be of low verisimilitude, hence why we do not select them over postulate sentences.

De Haro considers this briefly but rejects it, for two reasons, practical and logical:

> The logic of Popper's proposal turns out to be problematic because the definition does not allow the comparison of false theories. And, although the literature has proposed interesting solutions to this problem: I do not think that an explication of Popper's proposal is needed to reply to the pessimistic meta-induction argument that working scientists are confronted with. For working scientists do not normally calculate truth measures when they judge rival theories: and, since I do not here aim to reform scientific practice, but rather to reply to a conceptual argument against this practice, I will leave such quantitative and formal projects aside.
>
> This leaves a more qualitative approach as the interesting approach for our project. On this approach, one recognises that theory choice involves judgment, and that many qualitative arguments and considerations can be used that give reasonable verdicts, thus illuminating the history of science. Such qualitative arguments involve something like a quantitative reasoning: but more in the way of making a list of significant reasons for and against a general argument, i.e. without the pretence of the procedure being exact or comprehensive (De Haro, 2020, p. 51-52).

The practical objection to the use of verisimilitude is that everyday scientific practice does not involve the quantitative act of counting true and false consequences. In practice, theories are selected on the basis of quantitative reasoning, meaning according to the strengths of their respective arguments. This is true, though need not be a problem for the Popperian. Defenders of verisimilitude can acknowledge this, along with the explanatory importance of their theory, by saying that the qualitative approach has its foundations in the quantitative. That is to say that the *reason* given arguments are used in the selection of theories is that they point to theories having high truth-content. For example, the defence of a theory on the grounds of its explanatory power is qualatative, but explanatory power is important in a theory because it is used to derive more, and more useful, conclusions. The theoretical role of verisimilitude, the Popperian can say, is not to describe established scientific practice, but to explain the importance of practice in theory-selection. A good scientific theory has strong arguments in its favour, and the arguments are strong because they point to the high verisimilitude of a theory. In any case, if we grant, as De Haro does, that discussion of verisimilitude is of little scientific importance, we can hold still to its philosophical role in a defence of a philosophical thesis like scientific realism.

The more concerning objection is the point that Popper's explication of verisimilitude renders it impossible to compare false theories. Tichy (1974, p. 156-157) proves that the definition of verisimilitude (as, basically, laid out above) yields the result that no false theory can have any more verisimilitude than any other false theory. To see whether this applies in the case of extensional realist semantics, it is necessary first to show how such semantics would incorporate verisimilitude. De Haro imagines that the verisimilitude of a theory under extensional realism could be given by weighing the domains in which a theory is E-true against those in which it is not: "Indeed Popper's proposal, applied to E-truth as in this thesis, is to count the number of extensions of a theory with value 1 and subtract from it the number of extensions with value 0."

Taking this in board, we define Ce as denoting *extensional content*, where, for any a, $Ce(a)$ denotes the set of worlds where a has a truth value. $Ce_T(a)$ denotes the set of worlds where a has truth value one (given some level of abstraction), and $Ce_F(a)$ denotes the set of worlds where a has truth value 0 (given some level of abstraction). We have now defined extensional content, *extensional truth-content* and *extensional false-content*. We proceed as before with the rules regarding comparable truth content and false content:

> For any two expressions a and b, a is comparable to b with respect to truth-content iff $Ce_T(a) \subseteq Ce_T(b)$ or $Ce_T(b) \subseteq Ce_T(a)$, and comparable to b with respect to false content iff $Ce_F(a) \subseteq Ce_F(b)$ or $Ce_F(b) \subseteq Ce_F(a)$.

This gives us the following definition of extensional verisimilitude:

> For any expressions a and b, a has more verisimilitude than b iff a and b are comparable with respect to truth-content and false-content and either: $Ce_T(b) \subset Ce_T(a)$ and $Ce_F(b) \not\subset Ce_F(a)$, or else $Ce_F(a) \subset Ce_F(b)$ and $Ce_T(a) \subset Ce_T(b)$.

Now comes the question of what it means to compare two false theories. There are two things this might mean: either false in the conventional sense, which we can take to mean I-truth, or false as relates to De Haro's notion of approximate truth.

If approximately true means "E-true at no less than one world" (with more approximately true sentences being E-true at more worlds), then approximately false we can take to mean "E-true at no worlds". On this understanding, any approximately false theory t would be such that $Ce_T(t) = \emptyset$, and, where W denotes the set of all worlds, $Ce_F(t) = W$. Hence, for any two approximately false theories, t and t', $Ce_T(t) = Ce_T(t')$ and $Ce_T(t) = Ce_T(t')$. It therefore can never be the case that $Ce_T(t') \subset Ce_T(t)$ and $Ce_F(t') \not\subset Ce_F(t)$, or else $Ce_F(t) \subset Ce_F(t')$ and $Ce_T(t) \subset Ce_T(t')$. Therefore, on this understanding of approximate truth, no approximately false theory can have any more verisimilitude than any other.

As for understanding a false theory as "not I-true", here we fair better. W consist of w_1, w_2 and w_3. Given appropriate levels of abstraction, t is E-true at w_1 and E-false at w_2 and w_3. Likewise, t' is E-true at w_1 and w_2, but E-false at w_3. Since both are E-false at w_3, neither are I-true. t is comparable to t' with respect to extensional truth-content, since $Ce_T(t) = \{w_1\}$ and

$Ce_T t' = \{w_1, w_2\}$, meaning $Ce_T(t) \subseteq Ce_T(t')$. Likewise, it is comparable with respect to false-content since $Ce_F(t') = \{w_3\}$ and $Ce_F(t) = \{w_2, w_3\}$, meaning that $Ce_F(t') \subseteq Ce_F(t)$. By similar reasoning reasoning, $Ce_T(t) \subset Ce_T(t')$ and $Ce_F(t) \not\subset Ce_F(t')$. It follows from this that t' has more extensional verisimilitude than t, despite neither being I-true. Two false theories (on this understanding of false) can differ in verisimilitude. It remains true, however, that if both fail to be I-true by virtue of failing ever to be E-true, than one can have no more verisimilitude than the other. De Haro's logical complaint holds water.

Worse still, this account of extensional verisimilitude fails even to solve the problem we intended it for. A postulate sentence can never have more extensional verisimilitude than a corresponding Ramsey sentnece, and will typically have far less. Let p be a postulate sentence and rp a corresponding Ramsey sentence. Since $p \vDash rp$, for every world (and level of abstracion) such that p is E-true, so too is rp. It immediately follows that both p and rp are comparable with respect to extensional truth-content, since $Ce_T(p) \subseteq Ce_T(rp)$ and, since rp cannot be E-false without the same holding for p, $Ce_F(rp) \subseteq Ce_F(p)$. Suppose p has more extensional verisimilitude than rp, it follows that either $Ce_T(rp) \subset Ce_T(p)$, or else $Ce_F(p) \subset Ce_F(rp)$. Neither of these can be true, since both imply the existence of a world wherein p is true but its logical consequence rp is not. It can never be, therefore, that a postulate sentence has greater extensional verisimilitude than any of its corresponding Ramsey sentences. By similar reasoning, since, for any given postulate sentence and corresponding extreme Ramsey sentence, it is very likely the case that the extreme Ramsey sentence is E-true at more worlds, and it can never be E-false at more worlds, it will often be that an extreme Ramsey sentence has more extensional verisimilitude than its corresponding postulate sentence.

Even setting aside the question of comparing false theories, extensional verisimilitude, as laid out, is wholly inadequate to stop the collapse from extensional realism to extreme logical structural realism. It shall soon be argued that this is salvageable so long as we deny the initial premise that an adaptation of Popper's verisimilitude onto extensional realism involves the comparing of sets of worlds against each other. For now, though, we turn to non-Popperian attempts to measure verisimilitude, using possible worlds, to see if they are any more hopeful.

5.2 Truthlikeness and Distances

An alternate approach to the problem of choosing between theories is truthlikeness following possible worlds. Tichy (1974), in his criticisms of Popper's verisimilitude, concluded that a proper theory of truthlikeness would be one that took into account the *distance* between a proposition and the truth. Risto Hilpinen takes this idea to heart. He builds on Popper's verisimilitude in two ways, fleshing out an idea of *closeness to truth* and *informative content*:

> the truthlikeness of a proposition depends partly on its closeness to truth (or its degree of approximate truth [in the sense of degrees]),

> partly on the amount of information conveyed by the statement regarding its subject-matter (Hilpinen, 1976, p. 21).

This, if it could be made to work for extensional realism, could suit our needs perfectly. The extensional realist has their own account of approximate truth– which we can take in Hilpinen's terminology to mean "closeness to truth"– but has no baked-in theory of informative content. It might well be shown that, though Ramsey sentences are more (or at least no less) approximately true than postulate sentences, because the former kind lacks so much in informative content, their truthlikeness is lower than that of postulate sentences. If this is so, we can explain the need to pick postulate sentences over Ramsey sentences in our theory crafting via a Hilpinen-style account, thus preventing the collapse into extreme logical structural realism.

To do this, we must first lay out Hilpinen's theory to see how it might be adapted. A model of truthlikeness is a trupple $\langle U, ||, \mathcal{N} \rangle$. U denotes the set of all possible worlds, and $||$ is a function from propositions to sets of worlds such that, for any propositions A and B, $| A |$ is the set of worlds wherein A is true, and:

$$| \perp | = \emptyset$$

$$| \top | = U$$

$$| \neg A | = U \setminus | A |$$

$$| A \wedge B | = | A | \cap | B |$$

$$| A \vee B | = | A | \cup | B |$$

$\mathcal{N}$ invokes the concept of similarity. The idea being that A is approximately true when it is true at worlds similar to our own, and that, the more informative a proposition is, the more worlds it precludes that are similar to our own. Right away we run into some problems. The idea of ranking worlds according to similarity has been seen before. David Lewis (1979) famously did so in his explication of counterfactuals, holding that the statement "If A were true then B would be true" is true when, at all the most similar worlds where A is true, B also is true. The question then becomes *how* we judge one world to be more similar than another. This is not something easily done, there is no theoretically neutral answer, and it may be objectionable that extensional realism should have to burden itself with doing so. For now, we shall proceed with the idea that we can have some intuitive idea of when one world is similar to another and work around that. For some, this may be Hilpinen-style apprach falling at the first hurdle, but for others, who are more forgiving of similarity being consigned to an intuitive notion, or who have their own idea of similarity they wish to apply this to, it is worth continuing.

With that preamble out of the way, we define, for any u, $\mathcal{N}_u$ as a set structuring the worlds similar to u.[1]For any u, N_u is comprised of different "nested

[1]Hilpinen goes further and defines N_u^i as the set of worlds similar to u with respect to i.

spheres". These spheres order $\mathcal{N}$ according to similarity, meaning that one given sphere might be more similar to u than another. These spheres are subsets of U and satisfy the two following properties:

1 For every $K, L \in \mathcal{N}_u$, $K \subseteq L or L \subseteq K$.

2 u always belongs to any $K \in \mathcal{N}_u$, $u \in K$.

These are spheres in the sense that they are groups of worlds and all have their origin at the same point in modal space: u. The idea is that, the smaller the sphere, the more similar its worlds with respect to the origin point. For instance if $K \in \mathcal{N}_u = \{u\}$, then K is the sphere most similar to u, since it is identical to u. Where $K = N_u$, this sphere contains is the most dissimilar to u, containing far away worlds that are removed in similarity from u. $K \subseteq L$ may be interpreted, in this context, as saying K is at least as similar to u as L.

The function that determines distance from truth E makes use of the nested spheres of $\mathcal{N}$, in the following way:

> For any world u and proposition A, $E_u(A) = \{K | K \in \mathcal{N}_u$ and $K \cap \mid A \mid = \emptyset\}$.

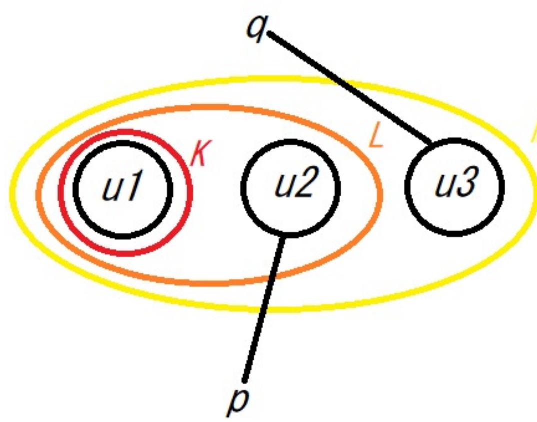
$E_u(A)$, then, is a measure of the number of nested spheres that can be made out of the set of worlds similar to u wherein A is false. The idea is that, for any A and u, the closer to the truth (at u) that A is, the smaller $E_u(A)$ becomes. If, for example, A is true at u, then $E_u(A) = \emptyset$– i.e. A is maximally close to the truth at u. If A is false, but true at some similar worlds to u, then $E_u(A)$ will be non-empty. The size of $E_u(A)$ in this case is dependent on how far away in logical space from u the worlds wherein A is true are to be found, as according to the ordering on nested spheres imposed by 2. If A is true at a world nearby to u, then the number of nested spheres in $\mathcal{N}_u$ wherein A is false will be lower than if A were true only at worlds further away. A contradiction cannot be further from the truth, and this is supported by the fact that $E_u(\mid \perp \mid) = \emptyset$.

Intuitively, then, Halipien's approximate truth runs as follows. Worlds are ordered according to how similar they are to each other. A proposition is close to the truth, as described by a given world u, when the worlds close u in the ordering of nested spheres of similarity are ones where that proposition is true. Likewise a proposition is far from the truth when the nearest worlds to u are ones where that proposition is false. Here we have closeness to truth using the notion of distance in logical space. It is possible to measure the relative closeness to truth of two propositions using this. A will be at least as close to the truth, as described by u, as B providing that the nearest worlds to u where B is true are ones where A is true. This means that it is not possible to create more nested spheres in worlds similar to u wherein A is false than B. This yields the following definition of relative closeness to truth:

For the sake of simplicity, and because we have already identified problematic elements of the notion of similarity and do not wish to burden ourselves by dealing with the concept more than need be, discussion of this is omitted here. Nothing said from here on cannot easily be adapted to meet this additional notation, however, and readers are free to do so should they wish.

A is at least as close to the truth, as described by u, as B is iff $E_u(A) \subseteq E_u(B)$.

Take the following example concerning propositions p and q, where $U = \{u_1, u_2, u_3\}$, and $\mathcal{N}_{u1} = \{K, L, M\}$:



At u_1, neither p nor q is true. p, however, is closer to the truth than q since the nearest world where p is true is closer to u_1 than the nearest world where q is true. More formally, $E_{u_1}(p) = \{K\}$ and $E_{u_1}(q) = \{K, L\}$, hence $E_{u_1}(p) \subset E_{u_1}(q)$, meaning p is at least as close to the truth as described at u_1 as q, but not the other way around, and so p is closer to the truth as described at u_1 than q.

The question of how to measure closeness to truth thus settled, Hilpinen turns to the question of informativity (truthlikeness being a measure of both). A popular idea is to equate the information expressed by a proposition to the set of worlds wherein that proposition is true. We are informed by a proposition when we are able to eliminate from our epistemic framework worlds incompatible with this set.[2] A simple idea, then, is that, the more informative a proposition, the more worlds it eliminates for us as epistemically possible. For example, the hypothetical proposition p, which is true only in the actual world, is hugely informative (in the actual world). Upon learning p, an agent in the actual world would consider the only epistemically possible world to be the actual world, and could thus be certain in their knowledge. Conversely, the proposition q is true at a great many worlds, and so eliminates few possibilities. All else, being equal, an agent learning q would still consider a multititude of worlds as epistemic possibility and could not have certainty as to which of these is the actual world. Hilpinen runs with this basic framework, though he is interested not only in how many worlds a proposition eliminates, but in how relevant they are (i.e. how similar they are to the actual world).

[2]See Stalnaker, 1999 as one of the chief defenders of this position.

To this end, Hilpinen adopts the following measure of a proposition's informativity:

> Where $I_u(A)$ denotes the informativity of proposition A and world u, $I_u(A) = \{K \mid K \in \mathcal{N}_u$ and $\mid A \mid$ is a non-empty subset of $K\}$.

$I_u(A)$ is the set containing all nested spheres wherein $\mid A \mid$ is contained in its entirety. The more of such nested spheres there are, the more informative a proposition is since A eliminates more worlds as epistemic possibilities. The use of nested spheres ensures also that this only measures worlds similar to u– i.e. relevant worlds. Supposing A is true at worlds entirely dissimilar to u, it follows that there is no $K \in \mathcal{N}_u$ such that $\mid A \mid$ is a non-empty subset of K. From the perspective of u, then, A is a completely uninformative proposition– i.e. $I_u(A) = \emptyset$– because it deals with epistemic worlds far removed in similarity from the actual world. Tautologies and contradictions are also uninformative, the first because they eliminates no worlds as possibilities, since tautologies are true at all worlds. If A is a contradiction, then $\mid A \mid$ is empty, and $I_u(A)$ accepts only nested spheres wherein $\mid A \mid$ is a non-empty subset of such spheres, and so must itself be empty. Returning to the above figure describing propositions p and q, we can say that p is more informative than q at u_1 since $I_{u_1}(p) = \{L, M\}$ and $I_{u_1}(q) = M$, meaning that $I_{u_1}(q) \subset I_{u_1}(p)$– i.e. q is less informative than p at u_1.

Having a notion now of both closeness to truth and informativity, Hilpinen can define truth likeness:

> A is at least as truthlike as B with respect to world u iff $E_u(A) \subseteq E_u(B)$ and $I_u(B) \subseteq I_u(A)$.

This means that A is at least as truthlike as B when A is at least as close to the truth as B and at least as informative as B. Note that there can be cases where A and B are incomparable with respect to truthlikeness, either when $E_u(A) \nsubseteq E_u(B)$ and $E_u(B) \nsubseteq E_u(A)$, or else when $I_u(A) \nsubseteq I_u(B)$ and $I_u(B) \nsubseteq I_u(A)$. In such cases, the worlds dealt with by A and B are distinct in similarity enough for the two propositions to be so dissimilar in content as to be incomparable.

All this laid out, we must then ask whether some version of truthlikeness along these lines could be laid out that is complimentary to the extensional realism project, what we might call *extensional truthlikeness*. The idea is that extreme Ramsey sentences, failing to be informative, will fail also to be truthlike. Right away, there is the problem of which world to designate as the actual world, since truthlikeness is reletavized to a given possible world. The move here must surely be to say that, since worlds on the extensional realism picture are domains, the actual world at any given point is whatever domain we happen to be working in. Say that we are using Newtonian theory in a given level of abstraction that means we are working in domain w_1. An evaluation of truthlikeness is ultimately a practical evaluation of usefulness, and so practical considerations like domain and level of abstraction are to be employed. In evaluating propositions for truthlikeness in such a case, we use w_1. We do not, then, have an absolute notion of truthlikeness, but rather an idea of truthlikeness within a given domain. We shall return to this idea below, but there is a much more pressing issue: the tension between Hilpinen's informativity and De Haro's approxmate truth.

First we define extensional informativity. $I_{d,L}(A)$ denotes the informativity of A at domain d– i.e. world d– and level of abstraction L. An extensional realist version of Hilpinen's approach must come with some orderings of similarity.[3] N_u, again denotes all domains similar to N_u– insofar as scientific domains can be similar to one another. It is certainly appealing to say that domains of quantum physics are more similar to domains of Newtonian mechanics than, say, cellular biology. We can assume a system of nested spheres amongst these worlds no different to Hilpinen. Where Hilpinen's evaluation function $||$ assigned to propositions the worlds in which they are true, we require something different, dealing with extensional truth. Instead, we say that, for any L, $\mid A \mid_L$ denotes the set of worlds where A is E-true given the level of abstraction L. Likewise:

$\mid \perp \mid_L = \emptyset$

$\mid \top \mid_L = D$ (where D is the set of all domains).

$\mid \neg A \mid_L = D \setminus \mid A \mid_L$

$\mid A \wedge B \mid_L = \mid A \mid_L \cap \mid B \mid_L$

$\mid A \vee B \mid_L = \mid A \mid_L \cup \mid B \mid_L$

Our definition of informativity, with all of this taken into account, comes out as:

$$I_{d,L}(A) = \{K \mid K \in \mathcal{N}_u \text{ and } \mid A \mid_L \text{ is a non-empty subset of } K\}.$$

And, for measuring relative infromativity with respect to a domain and level of abstraction:

> A is at least as informative as B with respect to domain d and level of abstraction L iff $I_{d,L}(A) \subseteq I_{d,L}(B)$.

This is an espescially helpful definition. It is observable that an extreme Ramsey sentence is E-true at any world with basic logical structures– which means they are likely E-true at worlds outside N_u for any given level of abstraction, meaning that an extreme Ramsey sentence ER would likely be such that $I_{d,L}(ER) = \emptyset$. In any case, where A is neither a tautology nor a contradiction, and $A \vDash B$ but not $B \vDash A$, as is the case for the postulate sentences of science and their corresponding Ramsey sentences, it will always be the case that $\mid A \mid_L \subset \mid B \mid_L$. Hence, by definition of $I_{d,u}$, $I_{d,u}(B) \subseteq I_{d,u}(A)$, meaning a postulate sentence will always be at least as informative as a corresponding Ramsey sentence. Further, assuming A is E-true at the domain in question, then, since

[3] The questionable nature of this has been adressed above.

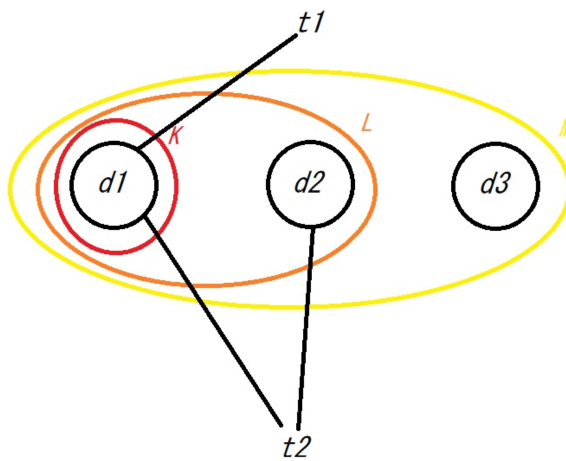
$| A |_L \subset | B |_L$, it will usually be the case that a postulate sentence comes out as extensionally more informative than its Ramsey sentences. Afterall, the more worlds a proposition is E-true at, the less it excludes, and a Ramsey sentence will usually be E-true at more worlds.

So far, so good. The problem, however, is that this theory clashes with De Haro's picture of scientific progress. The general idea was that subsequent theories were typically extensionally equivalent to their predecessors at all the worlds where the latter were E-true, and then themselves E-true at some additional worlds. Take, to illustrate this theory t_1 and its successor t_2. Given three domains d_1, d_2 and d_3 and the level of abstraction L_1, t is E-true at d_1 and E-false elsewhere. Its successor t_2 is extensionally equivalent to t at d_1 given L_1, as well as $E - true$ at d_2 and E-false at d_3. On De Haro's view, t_2 is more approximately true than t_1 because it is E-true in more domains as well as the same domains as t_1. On the other hand, our notion of informativity had it such that propositions were *less* informative when they are E-true at more worlds. This implies that, the more approximately true a theory is than its predecessor, the less informative it is at worlds where the two are extentionally equivalent.

Assume, by way of example, that d_1, d_2 and d_3 are similar worlds.[4] Each is a part of $\mathcal{N}_{d_1}$, which consists of the three nested spheres $K = \{d_1\}, L = \{d_1, d_2\}, M = \{d_1, d_2, d_3\}$ (see below). Therefore $I_{d_1,L_1}(t_1) = \{K, L, M\}$ and $I_{d_1,L_1}(t_2) = \{L, M\}$. Therefore, $I_{d_1,L_1}(t_2) \subset I_{d_1,L_1}(t_1)$. This means that t_1 is more informative at d_1 than t_2. This is a disastrous result. No theory should be more informative than another in a domain where the two are extensionally equivalent. There is really no intuitive reason why t_1 should be considered more intuitive than t_2, considering, where t_1 says anything, it says the same as t_2, and t_2 says more truthful things in more domains. In order for the Hilpinen approach to work, one must abandon De Haro's approximate truth, but that was a crucial element in the extensional realist picture. To do so would throw the baby out with the bathwater.

[4] Arguably this kind of assumption will usually be a fair one, since logical space is infinitely large and one needs a particular metric for isolating which worlds to study

t_2 is less informative at d_1 than t_1, despite t_2 being more approximately true.



Truthlikeness, as understood using distances between worlds, appears to be a dead-end. Extensional realism says that (generally), the more domains in which a theory is E-true, the more approximately true it is. Truthlikeness says that, the more domains in which a theory is E-true, the less informative it is at any of those domains. The two are irreconcilable. There remains something to take away from all this, which is the idea of reletavizing truthlikeness to a given world. A similar idea, relativing Popper's verisimilitude to particular domains, can, it shall now be argued, strengthen the case for verisimilitude preventing the collapse from extensional realism to extreme logical structural realism.

5.3 A Return to Verisimilitude

In the previous section, we saw that, in order to incorporate truthlikeness into extensional realism, it would have to be reletavized to a domain. This stands in contrast to verisimilitude as imagined by De Haro, which is aggregated across the possible worlds in a multiverse. Instead of measuring truth content by the logical consequences of a theory in the actual world, verisimilitude is measured by comparing sets of worlds.

The content of a sentence is dependent on the references of its constituent parts. The references of the constituent parts of a sentence are dependent on the possible world they are evaluated in. The sentence "the tallest person in London owns a hat" has differing content in the world where Harry is the tallest person in London than it does in the world where George is the tallest. The truth-content of a sentence, being simply the part of a sentence's content that is true, must also be dependent on the world of evaluation, as must also be the case for false content, and so, verisimilitude.

With this in mind, entertain the following notion of extensional verisimilitude. Let $C'_{d,L}(A)$ denote the content of A at a given world and level of abstraction.[5] $C'_{d,L}(A)$ denotes the set of theorems derivable from A in d given L. $T'_{d,L}$ is the set of E-true sentences at d given L. We can thus derive extensional truth-content, $TC'_{d,L}$, as follows:

> For all A, d and L, $TC'_{d,L}(A) = C'_{d,L}(A) \cap T_{d,L}$.

Returning to our example, the extensional truth-content of the sentences "the tallest man in London owns a hat" includes the sentence "there are hats" in the world where hats exist, but not in the world where there are none.

Much the same can be said for the extensional false-content of a sentence, $FC'_{d,L}$. Where $F'_{d,L}$ denotes the set of sentences that are E-false at d given L:

$$FC'_{d,L}(A) = C'_{d,L}(A) \cap F'_{d,L}.$$

The truth-content and false-content of a proposition at a given world and level of abstraction are thus equivalent to the intersection of that sentence's content at that world and the extensionally true sentences and extensionally false sentences of the world respectively. Whereas the verisimilitude in §5.1 defined content in terms of sets of *worlds*, we have moved instead to talk of the sets of true (and false) *sentences* and particular worlds.

Given this, we can define relative verisimilitude. First, we outline what it takes for two theories to be comparable with respect to extensional verisimilitude:

> A and B are comparable with respect to verisimilitude at d given L iff $TC'_{d,L}(A) \subseteq TC'_{d,L}(B)$ or $TC'_{d,L}(B) \subseteq TC'_{d,L}(A)$, and $FC'_{d,L}(A) \subseteq FC'_{d,L}(B)$ or $FC'_{d,L}(B) \subset FC'_{d,L}(A)$.

The first condition, $TC'_{d,L}(A) \subseteq TC'_{d,L}(B)$ or $TC'_{d,L}(B) \subseteq TC'_{d,L}(A)$ means that the two are comparable with respect to extensional truth-content (at a given domain and level of abstraction) and the second, $FC'_{d,L}(A) \subseteq FC'_{d,L}(B)$ or $FC'_{d,L}(B) \subset FC'_{d,L}(A)$, means they are comparable with respect to extensional false-content.

Finally relative extensional verisimilitude is defined:

> For any sentences A and B, A has more extensional verisimilitude than B at d given L iff A and B are comparable with respect to extensional truth-content and extensional false-content at d given L and either: $TC'_{d,L}(B) \subset TC'_{d,L}(A)$ and $FC'_{d,L}(B) \not\subset FC'_{d,L}(A)$, or else $FC'_{d,L}(A) \subset FC'_{d,L}(B)$ and $TC'_{d,L}(A) \not\subset TC'_{d,L}(B)$.

It immediately follows from this definition that, at any world where a postulate sentence is E-true, it will have a higher extensional verisimilitude than

[5] We have established already that the content of a sentence will vary from world to world (or domain to domain). Since De Haro has it that reference is fixed according to a level of abstraction, and so content must also be fixed by a level of abstraction, L is used to make this clear. It may be that use of L is unnecessary, but nothing of importance hinges on this.

its Ramsey sentences, since both will have empty extensional false-content sets, and the extensional truth-content set of a Ramsey sentence is a proper subset of the extensional truth-content set of a postulate sentence.

This may be spelled-out more formally. Assume that A is E-true, at some d given some L, and that $A \vdash B$ but not $B \vdash A$. Since $A \vdash B$, $B \in TC'_{d,L}(A)$. Therefore, given our revised definition of truth-content, $TC'_{d,L}(B) \subseteq TC'_{d,L}(A)$. Because A is E-true, and so B is E-true, neither have false-content at u, and so $FC'_{d,L}(A) = FC'_{d,L}(B) = \emptyset$, from which it follows that $FC'_{d,L}(A) \subseteq FC'_{d,L}(B)$. From our definition of relative extensional verisimilitude, we conclude that, because $TC'_{d,L}(B) \subseteq TC'_{d,L}(A)$ and $FC'_{d,L}(A) \subseteq FC'_{d,L}(B)$, A and B are comparable at u with respect to verisimilitude.

Having established that the two are comparable, it is then easy to show that A has the higher extensional verisimilitude at u. Because it is not the case that $B \vdash A$, $B \notin TC'_{d,L}(A)$. Therefore, $TC'_{d,L}(B) \neq TC'_{d,L}(A)$. This means that, because $TC'_{d,L}(B) \subseteq TC'_{d,L}(A)$, it must be that $TC'_{d,L}(B) \subset TC'_{d,L}(A)$. Likewise, because $FC'_{d,L}(A) = FC'_{d,L}(B)$, it follows that $FC'_{d,L}(B) \not\subset FC'_{d,L}(A)$. By our definition of extensional verisimilitude, then, A has the higher extensional verisimilitude.

Generalizing from this, we can conclude that, for any d where it is such that $A \vdash B$ and not $B \vdash A$, where A is extensionally true, A has a higher extensional verisimilitude than B. Since it is always the case, for any postulate sentence p and corresponding Ramsey sentence r, that $p \vdash r$, whenever a postulate sentence is E-true at a world, it has a higher extensional verisimilitude than any Ramsey sentence that can be derived from it.

Compare this result to the verisimilitude of §5.1. That verisimilitude ensured a Ramsey sentence would always have at least as much verisimilitude as a corresponding postulate sentence, and frequently more so. Extensional verisimilitude avoids this by considering sentences rather than worlds, and reletavizing extensional verisimilitude to domains.

We now turn to the question of how to use this definition in defence of extensional realism. Recall that the whole topic of verisimilitude came about out of a desire to explain how we pick theories on the extensional realist framework. Here, revitalization comes in. When we pick theories, we do so in a domain; the best theory of electrons is in the domain of physics, whereas the best theory of cells is in the domain of biology, for instance. We say, then, that a theory is best, in a given domain, when it has a higher extensional verisimilitude than its rivals.

We can maintain that a theory is more approximately true when it is extensionally true at more worlds, as laid out by De Haro. In this sense, extreme logical Ramsey sentences are very approximately true. But, in the domains where postulate sentences are E-true, they are the better theory. Extreme logical Ramsey sentences are only the better theory in domains wherein postulate sentences are E-false. This is not a heavy admission to make, because, despite this, we still say they are not very good scientific theories in those domains, owing to their low verisimilitude. The threat of collapse into extreme logical structural realism comes about from the premise that the best theories *sim-*

pliciter are the ones that are more approximately true. We avoid this problem by denying that there is such a thing as best *simpliciter*. There is only best in a given domain.

With regards to the pessimistic meta-induction, we can still say that the theories we accept today are approximately true, and more so than the theories of the past, but we need not commit ourselves to saying that they are the theories we accept today precisely *because* they are more approximately true. That way extreme logical realism lies. We say only that the theories of today are the best in the domains in which they are used because of their high extensional verisimilitude (which requires them to be extensionally true).

There is still the question of comparing the extensional verisimilitude of false theories. Tichy's proof that no two false theories can have any more or less verisimilitude in the actual world can be applied across all worlds, and so relatavization to different domains has not avoided this problem. There are two things to say here. Firstly, work has been done to deliver systems of verisimilitude that avoid this problem using relevance logics, which could be adapted to an extensional realist framework. Ken Gemes (2007) offers a particularly interesting account on this score. We need not go into that here because of the second point, that where the inability to compare the verisimilitude of false theories was a problem for Popper, it is not so for the extensional realist.

Popper was working with a binary system of truth and falsity in the actual world, and declared that all theories are (at least very likely) to be false. For science to progress at all under this, it is necessary to be able to compare false theories, otherwise no scientific theory has any worth. This is not the case for extensional realism, because extensional realism admits to extensionally true theories. For theory evaluation, we are able to compare theories that are true at a domain, given a particular level of abstraction. When we have theories already present that are, in a sense, true, it is unimportant for scientists to have to compare two theories known to be false. It is only of much importance to be able to compare false theories if one holds that the theories of science are routinely false, which the anti-realist accepts but the realist does not. That is not to say that a system of verisimilitude that allows for the comparison of false theories is of no importance at all, only that it is of limited importance to the extensional realist.

Chapter 6

Conclusion

With this adapted notion of verisimilitude, extensional realism stands as a truly novel reply to the pessimistic meta-induction. The threat of structural realism, and worse still extreme logical structural realism, is negated and we have a picture that explains not only why the claims of science are (approximately) true, but why the truths of science are interesting.

It is noteworthy that, in order to accomplish this, extensional realism is not here proposed as a purely semantic view, but contains a syntactic element as well. Very broadly, we can call a view semantic if it is model-theoretic and syntactic if it is proof theoretic. The extensional realist account of approximate truth is semantic in that it models truth values at possible worlds, so too is the truthlikeness approach. On the other hand, the verisimilitude offered is proof-theoretic, concerning what sentences can be derived from others. The debate about whether scientific theories should be formulated semantically is long standing.[1] Recent work, however, has suggested that semantic approaches and syntactic approaches need not be mutually exclusive. Lutz (2017) calls the debate illusory, holding that there is little practical distinction between semantic and syntactic approaches, and Frigg (2022) goes so far as to say that a proper formalization of scientific theory requires both in tandem. This idea has played out here. Our theory of approximate truth is semantic, but semantic attempts to rank theories– i.e. truthlikeness and the initial spelling out of verisimilitude– were insufficient to solve the problems of chapter four. It was only the addition of a syntactic verisimilitude that could do so; we judge what is (approximately) true semantically, and what truths are useful syntactically.

Recall that the chief method of beating the pessimistic meta-induction has been to identify some commonality across the history of science about which we can be realists. This is insufficient. One must go further and say why this commonality is important. The extreme logical structural realist is a realist about propositions that are of no concern to anybody. Such a position is both true and empty. The extensional realist has identified a host of scientific truths,

[1] See Lutz (2017) for an overview. Full discussion is outwidth the scope of this concluding chapter.

and, if these arguments are acceptable, is now equipped to say which are the useful truths, and to do so in a way that is not mere structural realism.

Chapter 7

References

De Haro Ollé, S. (2020), "On Inter-Theoretic Relations and Scientific Realism", PhD thesis, University of Cambridge.

Cited also: "An Extensional Scientific Realism (I) The Continuity of Extensions", *Manuscript a* (forthcoming); and "An Extensional Scientific Realism Part II", *Manuscript b* (forthcoming).

Floridi, L., *The Philosophy of Information*, (Oxford University Press, 2011).

Frigg, R, *Models and Theories: a Philosophical Inquiry*, (Routledge, 2022).

Gemes, K., (2007), "Verisimilitude and Content", *Synthese*, pp. 293-306.

Hilpinen, R., "Approximate Truth and Truthlikeness", in *Formal Methods in the Methodology of the Empirical Sciences*, eds. Przełecki, M., Szaniawski, K. and W'Ojcicki, R., (D. Reidel Publishing Company, 1976).

Kripke, S.A., *Naming and Necessity: Lectures Given to the Princeton University Philosophy Colloquium*, (Harvard University Press, 1980).

Laudan, L., (1981), "A Confutation of Convergant Realism", *Philosophy of Science*, 48.1, pp. 19-49.

Lewis, D., (1970), "How to Define Theoretical Terms", *Journal of Philosophy*, 67.13, pp. 427-446.

Lewis, D., (1979) "Counterfactual Dependence and Time's Arrow", *Noûs*, 13.4, pp. 455-476.

Lewis, D., (1984), "Putnam's Paradox", *Australasian Journal of Philosophy*, 62.3, pp. 221-236.

Lutz, Sebastian, (2017) "What was the Syntax-Semantics Debate in the Philosophy of Science About?", *Synthese*, 95.2, pp. 319-352.

Lyons, T.D., "Scientific Realism and the Pessimistic Meta-Induction", in *Recent Themes in the Philosophy of Science: Scientific Realism and Commonsense*, eds. Clarke, S. and Lyons, T.D., (Springer, 2002), pp. 63-90.

Maxwell, G., "The Ontological Status of Theoretical Entities", in *Scientific Explanation, Space, and Time: Minnesota Studies in the Philosophy of Science*, eds. Feigl, H. and Maxwell, G. (Minnesota University Press, 1962), pp. 181-192.

Maxwell, G., (1970), "Structural Realism and the Meaning of Theoretical Terms", *Minnesota Studies in the Philosophy of Science*, 4, pp. 181-192.

Musgrave, A., "Realism Versus Constructive Empiricism", in *Philosophy of Science: The Central Issues*, eds. Curd, M. and Cover, J.A. (Norton and Company, 1998), pp. 1088-1114.

Papineau, D., (2010), "Realism, Ramsey Sentences and the Pessimistic Meta-Induction", *Studies in History and Philosophy of Science Part A*, 44.4, pp. 375-385.

Poincaré, H., [1905], *Science and Hypothesis, the Complete Text*, eds. Frappier, M. and Stump, D.J, trans. Frappier, M, Smith, A. and Stump D.J. (Bloomsbury, 2017).

Popper, K.R., *Conjectures and Refutations*, (Routledge, 1962).

Popper, K.R., (1976), "A Note on Verisimilitude", *British Journal for the Philosophy of Science*, 27.2, pp. 147-159.

Psillos, S., *Scientific Realism: How Science Tracks Truth*, (Routledge, 1999).

Psillos, S, (2000), "Carnap, the Ramsey Sentence and Realistic Empiricism", *Erkenntnis*, 52.2, pp. 253-279.

Putnam, H., *Meaning and the Moral Sciences*, (Routledge, 1978).

Ramsey, F.P., [1929], "Theories", in *The Foundations of Mathematics and Other Logical Essays*, eds. Braithwaite, R.B. and Moore, G.E. (Kegan Paul, Trench, Trubner, & Co., 1931), pp. 212-236.

Rowbottom, D.P., (2014), "Aimless Science", *Synthese*, 191.6, pp. 1211-1221.

Russell, B., (1911)"Knowledge by Acquaintance and Knowledge by Description", *Proceedings of the Aristotelian Society*, 11.108, pp. 108-128.

Stanford, P.K., (2003), "Pyrrhic Victories for Scientific Realism", *Journal of Philosophy*, 100.11, pp. 553-572.

Stanford, P.K., (2015), "'Atoms Exist' is Probably True, and Other Facts that Should Not Comfort Scientific Realists", *Journal of Philosophy*, 112.8, pp. 397-416.

Tarski, A., (1943) “The Semantic Conception of Truth and the Foundations of Semantics”, *Philosophy and Phenomenological Research*, 4.3, pp. 341-376.

Tichy, P., (1974), “On Popper’s Definitions of Verisimilitude”, *British Journal for the Philosophy of Science*, 25.2, pp. 155-160.

van Fraassen, B., “Arguments Concerning Scientific Realism”, in *Philosophy of Science: The Central Issues*, eds. Curd, M. and Cover, J.A., (Norton and Company, 1998), pp. 1064-1088.

Worall, J., (1989), “Structural Realism: The Best of Both Worlds”, *Dialetica*, 43.1, pp. 99-124.

Wray, K.B., (2015), “Pessimistic Inductions: Four Varieties”, *International Studies in the Philosophy of Science*, 29.1, pp. 61-73.