Revista Română de Filosofie Analitică,
Romanian Journal of Analytic Philosophy
Vol. XVI, Iulie-Decembrie 2022, Nr. 2, p. 83-118
ISSN (ediția electronică): 1843-9969,
ISSN (ediția tipărită): 1844-2218
DOI: 10.62229/rrfaxvi-2/5

THINNING THE JUNGLE OF "UNCONCEIVED ALTERNATIVES." STANFORD'S ANTIREALISM MEETS EXPECTED UNIFICATIONS AND AVOIDABLE INCONSISTENCIES

IOAN MUNTEAN[1]

Abstract: K. Stanford (2006, 2009) has offered an antirealist argument (the "problem of unconceived alternatives", *PUA*) based on the argument that scientists are not able to grasp alternatives to a current scientific theory *T*. According to *PUA*, the mere existence of some epistemically inaccessible alternatives (*T'*, *T"*, …) weakens our trust in *T* and shakes the foundations of mainstream scientific realism. The realist may entertain the inkling that inter-theoretical relations (both existing and expected or 'hoped-for') play a role in accepting or rejecting *PUA*. The most celebrated intertheoretical relations, such as consilience, reduction, realization, emergence, equivalence, or approximation—whether prospective, expected, or realized—bear relevance to the conceivability of their alternatives. This paper presents an 'eliminative inference' based on an 'unification posit' that weakens the *PUA*. We employ first a minimal model of inter-theoretical unification couched in terms of the 'term identification' of the theoretical terms of two initially different theories, T_1 and T_2. Then we rethink unification as an 'ideological identification' where predicates in different theories are identified. Finally, we can envision a more sophisticated unification as entailment relations among T_1 and T_2 and their

[1] University of Texas, Rio Grande Valley and University of Illinois, Urbana-Champaign. Email: ioan.muntean@utrgv.edu. This material has benefited from the feedback of the audience and reviewers from the Philosophy of Science Association, the Society for Exact Philosophy, the Indiana Philosophical Association, and reviewers of the *RRFA*. Special thanks to Otavio Bueno, Ken Long, Andrei Mărăşoiu, and Connor Wilson for suggestions and support.

empirical grounds. In all these cases, we propose scenarios of inconceivability based on a minimal consistency requirement run against the "syntactic view" of scientific theories. The upshot of this mechanism is that *some* alternatives to T_1, which remain unconceived within the conceptual and ideological space of T_1, can be eliminated because they are inconsistent with empirical constraints on T_2. The overall space of 'serious' alternatives to both theories is 'thinned.' Consistency is a requirement that conditions inter-theoretical relations, mainly when the overlapping evidence supports theories. This argument illustrates in what sense *PUA* is lessened when scientists or scientific communities operate based on theoretical posits.

Keywords: *unconceived alternatives, K. Stanford, antirealism, conceptual space, B. Van Fraassen, P. Gärdenfors*

1. Introduction

Over the last five decades, several successful challenges to scientific realism have been formulated, including "the pessimistic meta-induction," the "underdetermination of theories by evidence," and the "problem of unconceived alternatives" (PUA). This paper is focused on the latter, which comes in several flavors. L. Sklar states that for any scientific theory T there are always incompatible alternatives (in what follow, alternatives to T are designated as a space of theories between which no rational choice can be made based on a priori "plausibilities, strength, parsimony, inductive confirmation, and so forth, relative to present empirical evidence" (L. Sklar 1981). Sklar's underdetermination of T by its alternatives is recurrent and transient. All alternatives are transient because new data may render some less attractive or make others more preferred, depending on the evidence we gather in the future (L. Sklar 1981; Lawrence Sklar 2000).

K. Stanford has advocated a different version of PUA and presented it as a comprehensive argument, supported by some historical evidence.[2]

[2] *PUA* is exposed in the 2006 book and promoted as "the new induction" (*NI)* in contrast with the mainstream pessimistic metainduction (PMI) (Stanford 2006). Other references to the *PUA* and NI are (Bhakthavatsalam and Kidd 2019; Chakravartty 2008; Devitt 2011;

PUA is premised on the idea that the most significant challenge to scientific realism arises from our inability to consider the full range of serious alternatives to a given hypothesis we seek to confirm and ultimately accept. Stanford's argument is an induction from the history of science (including all domains and all historical periods) to present and future science:

> We have every reason to believe that there are theoretical alternatives remaining unconceived by us whose grasp will be regarded by future scientific communities as absolutely fundamental and/ or a necessary precondition for conceiving of or even understanding the further accounts of nature that they themselves embrace. (Stanford 2006, 131).

In his book-length argument, Stanford puts forward a very general claim: scientists, in *every* scientific field and *at any time* in history, have found themselves in an "epistemic position in which [they] could conceive of only one or a few theories that were well confirmed by the available evidence, while subsequent inquiry would routinely (if not invariably) reveal further, radically distinct alternatives as well confirmed by the previously available evidence as those [they] were inclined to accept on the strength of that evidence" (Stanford 2006, 19). The book discusses at least three cases from the 19th-century history of biological inheritance: Ch. Darwin's "pangenesis theory", Fr. Galton's "stirp theory," and Weismann's "germ-plasm theory." The view ignored by all these scientists was gene regulation, so each of these scientists failed to envisage a relevant alternative to their theory about inheritance, the alternative that would have been accepted later by the scientific community (Stanford 2006, 132–33).

First, there is an essential caveat under which Stanford's argument is operational: because he focuses on unconceivability, *PUA* refers to scientists, not scientific theories. This sharply contrasts with the standard underdetermination problem, which emphasizes the falsehood of theories. Relevantly, Sklar's 'transient underdetermination' refers to both

---

Forber 2007; Godfrey-Smith 2007; Kukla 2010, 2010; Rowbottom 2019; Saatsi, Psillos, and Winther 2009; Stanford 2021, 2009, 2015b; Zamora Bonilla 2019).

the truth of theories and the scientists inability to conceive of alternatives to existing theories. The present paper offers an analysis of the inconceivability of radically different alternatives to a given theory T, which is arguably distinct from the mere existence of logical alternatives, which may not be different relevantly from T. Here, $\mathcal{T}_T = \{T', T'', T''' \ldots T^{(n)}\}$ is the countable set of alternatives of T, all of them being supported by the same evidence E as T but most of them being incompatible with T. This raises the immediate issue of the identity of theory T when compared to its unconceived alternatives. How do they differ from other members of $\mathcal{T}_T$ and from T? The fundamental difference here is epistemic, as given by all scientists at a moment t, none of the scientists can conceive any of the $T', T'', T''' \ldots$

From the history of science discussed in the book, Stanford infers that future scientists would see the space of our "theoretical grasp" as limited as we see old theories nowadays (*e.g.,* Weismann's) through the lenses of molecular cell biology:

> We have every reason to believe that there are theoretical alternatives remaining unconceived by us whose grasp will be regarded by future scientific communities as absolutely fundamental and/or a necessary precondition for conceiving of or even understanding the further accounts of nature that they themselves embrace (Stanford 2006, 131).

The current molecular genetics has been arguably supported empirically by the experimental data available to Weismann, whereas his theory has been rejected by the data made available subsequently (in the 1920s).

1.1. Conceivable problems with the Problem of Unconceived Alternatives (*PUA*)

There are probably several ways to reject Stanford's *PUA*. First, a realist can "trivialize" this "problem" by insisting it is not entirely different from the underdetermination of theories by evidence or the pessimistic metainduction. Hence, it is fundamentally vulnerable to the same arguments marshaled usually against antirealist arguments. If *PUA* is nothing more than *UTE* and *PMI* in disguise, then it can be addressed by the realist in the same way. *PUA* is important in many ways, but it generates new consequences for neither the scientific realism nor the antirealism positions.[3]

Second, one can reject the range to which *PUA* applies. Even if *PUA* is plausible in the case studies discussed in the book, what makes us think that the schema can be generalized beyond that historical context? As some reviewers have noted, Stanford's historical initial base is relatively small and restricted to a relatively short period in the history of biology (Ruhmkorff 2019; Votsis 2007). *PUA* is then a weak inductive argument, either because the sample used by Stanford is small, or because the sample is atypical.

However, one way to reconsider Stanford's *NI* is by applying it to other disciplines or other theories outside the scope of evolutionary biology. P.D. Magnus questioned the validity of Stanford's schema in a different context by addressing this question: were classical mechanics and the special theory of relativity equally confirmed in, let us say, the year 1780 (Magnus 2006). One can admit that carefully selected data from astronomical and terrestrial observation made before the 1780s would corroborate the special theory of relativity as an unconceived alternative to Newtonian mechanics.

Third, one can weaken Stanford's *NI* by showing that his predicament is fundamentally sound for a large enough class of examples from the *past* history of science (included in the book), albeit current scientific theories and the way scientists think about science differ in some fundamental respects from historical cases. We have become better at approaching truth, using scientific standards, and acknowledging the schema of falsificationism in our current theories than scientists in the past have done. As some realists have noticed, our current and future theories are becoming more sophisticated than the historical cases at hand.

---

[3] This argument is expressed in (Chakravartty 2008; Ruhmkorff 2019) and by J. Saatsi in (Saatsi, Psillos, and Winther 2009; Saatsi 2019).

1.2. The maturity question: science or scientists?

In a more concrete sense, one can talk about the maturity of scientific communities and the maturity of their respective theories. According to this argument, most cases discussed by Stanford are not 'mature' theories, in the sense that some of our present theories are mature. Similarly, others would consider that scientific communities have become increasingly efficient in discovering alternatives to a given theory. P. Godfrey-Smith pointed out that scientific communities in our days are bigger, better connected, better organized, and can better explore the alternatives to mainstream theories (Godfrey-Smith 2007). The current communities are less vulnerable to the problem of "unconceived alternatives" than past communities of scientists because of some fundamental differences in their "epistemic status." Others would claim that *PUA* is less likely to be applied to the present and future of science, because something in the methodology, metaphysical commitments, and the general organization of science has changed radically since, let us say, the biology of the 19th century. As M. Devitt stated:

> […] we have very good reason to believe that we have been getting better and better at learning about the unobservable world; good reason to think that, aided by technological developments, there has been, over recent centuries, a steady improvement in the methodology of science. That's why our present theories are more successful" (Devitt 2011, 292).

Yet another key ingredient is the claim that the changes in the general advancement of science will not be so "dramatic," to put it that way, in the future. Stanford coined this attitude as "scientific catastrophism" (Stanford 2015a). The catastrophist postulates that the future of science will be quite different from its past or present and that the historically significant changes in science are mostly a thing of the past. Catastrophists weaken the power of the *PUA* from the past to the future by postulating that the history of science is not uniform. To the other camp, the "uniformitarians," significant changes will occur in the future at roughly the same rate as in the past: "In the course of further inquiry, those theories will ultimately be overturned, supplanted, or transformed in the manner of their historical predecessors" (Stanford 2015a, 877).

The realists usually retort by pointing out that the history of science is not uniform and that irreversible changes affect how we conduct the scientific endeavor from now on. When realists such as Godfrey-Smith and others emphasize that communities of scientists are more important (as opposed to individual scientists), they provide such a mechanism against *PUA* (Dellsén 2019; Godfrey-Smith 2007). To attack Stanford's *NI*, it is sufficient to show that when a community of scientists is larger, communicates more effectively, adopts a more effective methodology, separates relevant information from noise, etc., the scientific endeavor is less vulnerable to the problem of unconceived alternatives. Stanford shows that, on the contrary, the current situation is fundamentally different: theoretical orthodoxy, deep-rooted bias of the present institutionalized science, fosters scientific *conservatism* which, as his argument goes, was not present during the Scientific Revolution and the one or two centuries to follow:

> Today's scientific communities are almost certainly more effective vehicles for testing, evaluating, and applying theoretical conceptions of various parts of the natural world than were their historical predecessors, but I have argued that we have compelling reasons to believe that they are less effective than those same predecessors in conceiving, exploring, or developing fundamentally novel theoretical conceptions of nature in the first place (Stanford 2015b, 3931).

1.3. Assumptions of the current argument

This paper proposes a new framework for Stanford's debate with P. Godfrey-Smith, A. Chakravartty, J. Saatsi, and F. Dellsén. In the present framework, we set aside the sociological and historical differences between the past and the future of science *per se*. Separately from the value-laden historical context, one can see that science has been periodically controlled by certain 'posits': theoretical virtues or theories, ideals, background information, methodological maxims, metaphysical assumptions, and worldviews, among others. The posits are conceptualized as “external posits” by P. Vickers, “presuppositions” by Ph. Kitcher, or “idle elements” by St. Psillos (Philip Kitcher 1995; Psillos 1999; Vickers 2013). Sometimes these posits are adopted explicitly, sometimes presupposed tacitly. They do not show up explicitly in scientific inferences, so they can be called “indirect.” Nevertheless, at first glance, some posits can serve as eliminative entailments that decrease the number of unconceived alternatives. At various points in the history of science, these posits can be more or less effective in diminishing the strength of Stanford’s inductive argument.

We take an ‘unification posit’ as a sought-for epistemic virtue of a community of scientists and, by extension, a virtue of a scientific discipline (hopefully constituted by more and more unificatory theories). The main aim of this paper is to investigate how unificatory posit imposed on the theories of one discipline can reduce the pool of alternative theories, even if these alternatives remain unconceived.

Stanford argues that today’s scientists are not more creative or efficient in their ability to exhaust the logical space of alternatives to a given theory. Regarding background information, Stanford believes that auxiliary hypotheses, although they may improve over time, are typically overlooked despite being equally well-supported by the available evidence. In other words, in such cases, the totality of evidence available at the time of an earlier theory’s acceptance characteristically offers equally compelling support for the combination of a later accepted alternative to that theory, together with the requisite alternative auxiliary hypotheses that would themselves later come to be accepted (Stanford 2006, 20). This paper argues that theories of an scientific discipline *D* at an earlier moment could be subsumed under an ‘expected unification’ posit, and this in itself would reduce space of reciprocally *consistent* unconceived alternatives in *D*.

The present argument does not attempt to show that contemporary scientific communities are better than past ones or worse than future ones: they profess different posits. The argument here is premised on the idea that communities of scientists may operate at different epochs, on a different set of theoretical posits, and seek different epistemic goals, which can be couched in terms of posits. Nevertheless, some of these posits endure over time or are more stable than methods, experimental setups, evidence, or theory choice standards. Each scientific discipline can undergo different stages, but one or more of its posits remain relatively unchanged. It is assumed here that posits such as parsimony, unification, explanatory power, predictive power, empirical adequacy, and coherence play a particular stable and enduring role. If the present argument is sound, then theoretical posits offer a general argument against the *PUA*, which weakens it, independently of the uniformity of the history of science, of different sociological factors, or the empirical success or failures of *D*. These theoretical posits are, it is assumed here, general enough and change relatively infrequently in the history of science. S. Schindler suggests that we should believe some of our theories because they (some only!) possess 'virtues' (non-ad hocness, consistency, unification, parsimony, or fertility) that extend beyond the evidence (Schindler 2018). This is the realist posit about scientific theories, and according to some statistical analysis, posing theoretical virtues is popular among scientists (Mizrahi 2022; Schindler 2022).

Middle positions between realism and antirealism have friendly attitudes towards such posits. In M. Massimi's perspectivism, standards of scientific conduct are relatively stable when compared to changes of perspectives (Massimi 2021). The present argument can be read as a conditional statement: "if this and that standard *S* is present and endures at some epoch in the history of *D*, that period is less vulnerable to the *PUA*, compared with other epochs in which *S* is abandoned or replaced frequently." The current paper identifies at least one standard *S* that may weaken the power of the *PUA*. This argument can be categorized as an argument for convergent realism, albeit weaker than arguments based on the need for diversity in the sampled history. It is not an argument based on the uniformity of the history of science, but can be used in conjunction with it.

2. Space of Unconceived Alternatives to Theories (UAT) and theoretical posits

What does Stanford mean by inconceivability and how do we identify elements in $\mathcal{T}_T$? Stanford and his critics express only informally the requirements on $\mathcal{T}_T$.[4] Stanford runs on a skeptical ticket against scientific knowledge by emphasizing that the *PUA* is an undisputable and unavoidable aspect of scientific life, at any moment in the history of science, contemporary scientists are no exception. As Stanford's *PUA* is primarily branded as an argument against scientific realism, its cognitive and epistemic dimensions are often overlooked. On the contrary, by situating his argument on the realist-antirealist map, we take Stanford's argument to tackle the cognitive inability of scientists to conceive alternatives directly. The present paper frames cognitive inability in terms of syntactic constraints on how scientists use the vocabulary of existing theories. Based on Stanford's own characterization, "unconceived alternatives" to our current theories are the result of our cognitive inability to conceive a different theory equally well confirmed by the existing and known empirical evidence. The failure to conceive alternatives to a theory is couched in the following sections in a first-order logic formalism. Imagine we have a finite set of objects and a finite set of conceivable predicates. One way to delineate the inability to conceive alternatives to a given theory that pairs some objects with some predicates is to keep the vocabulary of the language fixed and set a limit on our ability to pair the same objects with existing predicates. The other approach is to depopulate the signatures of relevant objects or predicates that are currently not used in our science but may become helpful in the future. Yet another, even more radical way is to introduce entirely new constants and predicates in $\mathcal{T}_T$ nonexistent in T. In the following, we aim to clarify these scenarios.

One suggestion, repeated in different contexts, is that inconceivability is linked to scientists' inability to imaginatively exhaust a *space* of plausible, scientifically serious, and reasonable candidate

[4] See the *Synthese* volume on *PUA* and its introduction in (Bhakthavatsalam and Kidd 2019).

theories for a given set of phenomena before "proceeding to eliminate all but a single contender" (Stanford 2006, 29, 31, 32). Let us assume, informally, that alternatives to *T*, referred here as the space of 'Unconceived Alternative Theories' (*UAT*), form a space that scientists can, in principle, explore. One task of the present approach is to develop a more precise formulation of this *UAT* space and ways to reduce it. It is often repeated by Stanford and Sklar that the space of important, and serious, alternatives to a theory is "indefinitely large:" at a given moment, a scientist or even a large community of scientists can embrace only a subspace of such a space. Genuine conceptual improvements in the present, compared to past science, mean that:

> we enjoy the luxury of conceiving of and considering an ever-larger space of serious theoretical alternatives. Of course, even if the space of unconceived alternatives contained only a finite number of well-defined possibilities, we would seem to have little reason to believe that we are presently at the end of an exhaustive search of it and have finally reached the point at which serious unconceived possibilities no longer pose any real danger to our theoretical science in a given domain (Stanford 2006, 133).

Stanford warns us that the space of alternatives has a "vague and indefinite character, with members that are difficult if not impossible to individuate sharply or unequivocally: an indefinite number of alternative possibilities are neglected" (Stanford 2006, 133). But the argument for such a claim, according to Stanford, is mainly historical:

> While there are certainly cases of eliminative inferences in which we can justify restricting our attention to some small part of the space of possibilities […] our historical investigation will suggest that in the case of fundamental theoretical science it is often a consequence of our failure to conceive of the serious alternative possibilities that do in fact exist that we embrace the substantive assumptions needed to restrict the space of theoretical alternatives under consideration to a comparatively small and/or well-behaved set (Stanford 2006, 41).

Stanford presages us to accept that the space of alternatives has a "vague and indefinite character, with members that are difficult if not impossible to individuate sharply or unequivocally: an indefinite number of alternative possibilities are neglected" (Stanford 2006, 133).

2.1. Downsizing the *UAT* space

It is part and parcel of the present argument to provide a formal framework for a version of the eliminative argument that can reduce the impact of the *UAT* spaces on the practice of science by scientists and on scientific realism. We restrict the present analysis parsimoniously to a first-order logic, a rather orthodox form of the "received view" on scientific theories and the theoretical language (TL_T) of the theory, without insisting too much on the observational language (OL_T) or their correspondence functions (CF_T).[5]

When and how do theories have unconceivable alternatives? There is a trivial answer to this question: always! At any time *t*, in any scientific discipline *D* and any theory *T* within it, there are alternatives (conceived or not by the scientists practicing *D*), supported at *t* by the same known evidence *E* that supports *T*, just because any *T* is underdetermined by *E*. This trite answer is well heeded by Stanford: the underdetermination of any *T* by evidence *E* is not the same as *PUA*. Although there is an infinity of trivial, conceivable alternatives to *T* equally well supported by *E*, not all of them are 'serious,' distinct enough from each other (and from *T*), and truly unconceivable. Nevertheless, what are the *serious* and *distinct* alternatives to *T* still unconceivable within *D* (i.e., not by the individual scientist, but all the scientists practicing *D*)?

Relating the informal approach of the *NI* to the syntactic view and scientific unification requires some conceptual clarifications and definitions that both proponents and critics of *PUA* could accept. For the present purposes, we choose a minimal formalism suited for bringing together Stanford's *PUA* and the unification posit in the received view framework.

[5] We follow here the notations and standard syntactic conventions from (Carnap 1995 (1966); Hempel 1966).

To challenge *PUA*, any realist argument must articulate clearly the "truly inconceivability" of *T*'s alternatives. Domain *D* and a definite empirical evidence *E* to *D*, one can imagine a set $\Theta_E$comprised of all theories in *D* empirically supported by *E*. In simple probabilistic terms:

$$\forall T(T \in \Theta_E \rightarrow \left(pr(T|E) > pr(T)\right)$$

Within Θ_E there is a 'field' Υ_E defined as a set of theories accepted, known, and grasped by scientists practicing that discipline, together with the accepted body of evidence *E*. The complementary set $\overline{\Upsilon}_E$ is simply the set of unknown and inconceivable theories supported by E.[6] We assume now these rather inconspicuous claims:

(1) Any theory T in Υ_E is known and accepted by the scientists practicing in *D* at *t*. In a formal notation, $\forall T(T \in \Upsilon_E \rightarrow (K_t(T,D) \wedge A_t(T,D))$, where $A_t(p,C)$ is a two-place predicate that formalizes the acceptance of a proposition p by the community *C* at *t*, and $K_t(p,C)$ formalizes that p is known by the community *C* at *t*.[7]

Acceptance of a scientific theory must meet some conditions, even if, as Quine points out, "acceptability depends on a weighing of the total evidence" (Quine and Ullian, 2007). Minimally, they are at least:

(2) Acceptance in *D* that *E* supports *T* at *t* in a probabilistic term: it is true that the probability of 'T given E', *pr(TE)*, is higher than the probability of *T*: $\forall T(T \in \Upsilon_E \rightarrow A_t('pr(T|E) > pr(T)'$,D);

(3) Acceptance in *D* that *T* is 'true' at t: $A_t('T'$,D). This may mean in a strong form that if s is a scientist, then she believes '*T* is true' or, in an inferential way, that the negation of T entails a contradiction given rules of inference in D: $\forall s(s \in D \rightarrow A_t('T', s)$. In terms of entailment: $\forall s(s \in D \rightarrow A_t(\sim T \vdash_D \perp, s)$.

---

[6] We assume that $\overline{\Upsilon}_E \cup \Upsilon_E = \Theta_E$ is a partition of Θ_E.

[7] We do not assume here any particular relation between predicates *A* and *K* except that accepted theories are "known" and that "to be known" does not entail "to be accepted."

2.2. A functional definition of unconceivability

What is exactly the field of theories Υ_E?At a more refined level of analysis, other requirements for accepting T may include predictive and/or explanatory powers, conformity with current scientific standards (methodologically, epistemically, etc.), fecundity, usefulness, etc. However, for the present purpose is adamant to see that the D community can, in principle, expect T to enter into some *future* inter-theoretical relationship with another accepted theories from Υ_E. Then T is accepted as part of a larger theoretical field composed of different theories, models, etc. in Υ_E and within this theoretical field some theories may become reduced, eliminated, or… unified. This is a holistic acceptance that constitutes a theoretical posit relevant to NI and PUA. In this paper, we focus on the unifying inter-theoretical posit that can potentially reduce the power range of PUA.

DEFINITION 1. At any moment t, for any theory T in Υ_E there is a non-null unconceivability function $U_t : \Upsilon_E \to \overline{\Upsilon}_E$ relating one theory in $\Upsilon_E$to a countable set of unconceived alternative theories $\mathcal{T}_T \in \overline{\Upsilon}_E$,such that $\mathcal{T}_T = U_t(T)$:

- I) At t, E supports all alternative theories in the inconceivability co-domain of T:

$$\forall T'\big(T' \in \mathcal{T}_T \to (pr(T'|E) > pr(T')\big)$$

- II) At t, none of the theories in $\mathcal{T}_T$ is known (grasped) or accepted in D:

$$\forall s \forall T'\big((s \in D \wedge T' \in \mathcal{T}_T) \to (\sim K_t(T', s)\big)$$

- III) Unbeknownst to all scientists in D, there is at least one theory in $\mathcal{T}_T$ that E supports better than it supports T:

$$\exists T' \Big(T' \in \mathcal{T}_T \wedge \big(pr(E|T') > pr(E|T)\big) \wedge\sim K_t(T', D)\Big)$$

- IV) The degree of the inconceivability of T is not related to one alternative T', but to the size of the minimal inconsistent set of T and $\mathcal{T}_T$: $\mathrm{MI}(T \wedge \mathcal{T}_T)$.

We do not envisage inconceivability as a characteristic of T, but as a function $U_t: \varUpsilon_E \to \overline{\varUpsilon}_E$ which relates at t some accepted T in $\varUpsilon_E$ (by all its practitioners $\forall s \in D$) and empirical evidence E that supports T at t (a moment or an interval of time) to a subset $\mathcal{T}_T$ from the set of the unconceived theories supported by E: $\mathcal{T}_T \in \overline{\varUpsilon}_E$.

The minimal inconsistent set of A is defined as a subset of A that is inconsistent, but any sub-subset of the latter is consistent:

$$\mathrm{MI}(A) = \{A' \in A; A' \vdash \perp; \forall A'' \in A'; A'' \nvdash \perp\}$$

(Benferhat, Dubois, and Prade 1997).

Removing one element from MI(A) makes it consistent but MI(A) is inconsistent. A consistent set A has $\mathrm{MI}(A) = \varnothing$, while for an inconsistent set $A', 2 \leq \mathrm{card}(\mathrm{MI}(A')) \leq \mathrm{card}(A') - 2$. We expect U_t to change in time, as E changes, as well as how E supports T. The identity of theories in $\mathcal{T}_T$ is vital for PUA, as at $t_1 > t_0$, one or more of them will replace T as newly accepted theories in D, once new evidence E_1 is acquired at t_1. We correlate the "seriousness" of the alternatives to T with the $\mathrm{MI}(\mathcal{T}_T \wedge T)$, but not with the $\mathrm{MI}(\mathcal{T}_T)$. In fact, we impose the condition:

$$T \in \mathrm{MI}(\mathcal{T}_T \wedge T)$$

This means that T has to be incompatible with members of $\mathcal{T}_T$, whereas the inconsistency of $\mathcal{T}_T$ is not relevant here.

An unconceived alternative to T in D is a member of the set $\mathcal{T}_T = U_t(T)$, all empirically supported by E at t. Members of $\varUpsilon_E$, the set of known and accepted theories in D is supposed to be as consistent as possible, whereas $\overline{\varUpsilon}_E$ is not, and neither is $\Theta_E = \varUpsilon_E \cup \overline{\varUpsilon}_E$. Crucially, Stanford thinks that $\mathcal{T}_T$ and T have a non-trivial degree of logical inconsistency, albeit they are at t supported by the same empirical evidence. However, the definition above is not sufficient for the present argument. The framework in which the present argument is couched includes these extra components:

- Delineation of the 'inconceivability of alternatives' $\mathcal{I}_T$, given T;
- Simplified versions for the (future) unificatory posit of two theories T_1 and T_2 using the "syntactic view".
- A requirement for consistency and the delineation of this requirement from the lax presence of inconsistencies among sets Υ_E and $\overline{\Upsilon}_E$.

The result is an eliminative inference that weakens the *PUA* by discarding classes of alternatives to *T*. The eliminative inference proposed here removes parts of *UAT* from consideration and restricts the attention to its significant subspaces. The assumption used here is a 'unificatory' posit, which qualifies in Stanford as "substantive assumption" (Stanford 2006, 40). The unification posit relates two theories from *D*, rather than a single theory (as does, for example, the parsimony posit), and for the present goal, it would weaken in some cases *PUA*. Unification is more akin to a normative and prescriptive posit and not as an *actual* and *effective* inter-theoretical relation: statistically, we have robust reasons to believe that contemporary scientists adopt it as an ideal (Mizrahi 2022; Schindler 2022). Instead of focusing on one theory and its alternatives and successors, the present approach relates two different theories simultaneously, mostly known or potentially known to the scientific community. The unification is the hypothesis that in the future they could be unified by a simple process of identification of theoretical terms, and it probably should be called the "austere unification" which is in this approach expected, sought-for, or just hoped-for.

Two caveats are in order here. First, we use a simplified version of scientific unification, decoupled from explanation, prediction, or understanding.[8] We proceed this way to better relate to Stanford's claim

[8] In this sense, the present approach does not follow the more orthodox approaches to unification of Friedman, Kitcher, Morrison, or Schupbach (Friedman 2001, 1974; Philip Kitcher 1981; P. Kitcher 1999; Morrison 2000; Schupbach 2005). Unification as theoretical

that his new induction is about scientists, their goals, ideals, and practices of science, not about scientific theories *per se*. It appears that neither explanation (including inference to the best explanation), nor prediction plays a central role in *PUA*. What is central in *PUA* is being realist about the theoretical terms of our current accepted theories.

Second, we use a basic version of the "syntactic view" of scientific theories. There are two main reasons for employing the syntactic view: first, it seems that Stanford's own argument is couched more in terms of the syntactic view (*aka* "Received View"), and less in terms of the more popular "Semantic View" of scientific theories. *PUA* is about how scientists conceive and formulate their theories as collections of statements about the world, and not as models. Second, even if this present approach is provisional and probably simplistic, it is unclear whether a semantic view approach would fundamentally change the eliminative inference argument against *PUA*.[9] The present choice reflects a general trend to view the syntactic and semantic views as alternative descriptions of scientific theories, rather than opposing ones. Some authors talk about the peaceful coexistence of these two views (Lutz 2017, 2014), while others dismiss the semantic view altogether and favor the syntactic view (Halvorson 2013). Without further ado, a syntactic approach based on first-order logic is assumed to be sufficient for the present purpose. Our aim no is to formalize scientific theories, but to provide an eliminative inference that weakens *PUA* in specific cases. An alternative approach, based on the semantic view of science (models or partial structures rather than theories), or any other alternatives to the syntactic view, may or may not illuminate interesting aspects, but is not followed here.

---

virtue is discussed more recently in (Kao 2019; Patrick 2018; Roche and Sober 2017; Schindler 2022). Pluralism about scientific theories and the 'dis-unity posit' in science is disucssed in (Cartwright 1999; Dupré 1993; Hartmann 2001).

[9] This can be an interesting venue for research, not addressed here. We use the idea of conceptual spaces, similar in spirit to the syntactic view, as discussed in (Gärdenfors 2000). An even more attractive option, belonging to the semantic view, is to use the formalism of "partial structures" and quasi-truth, advanced by French and da Costa (Costa and French 2003). See a comparison in (Bueno 2015).

Practitioners of science in D are employing a first-order language $\mathcal{L}$, formalized as a vocabulary $\mathcal{V}$, containing logical terms such as some quantifiers $(\exists, \forall, \dots)$, connectives $\rightarrow \leftrightarrow \vee \wedge, \dots$, the identity symbol (= used between constants or variables), and a signature Σ which includes the constants $c_1, c_2, c_3 \dots$ (representing theoretical or unobservable terms), predicates $P_1, P_2 \dots . P_n$ with arbitrary arity, and functions $F_1 \dots F_p$, each with an arbitrary number of arguments $F_i(x_1, x_2 \dots .)$.[10] To all these we can add rules of entailment for $\mathcal{L}$ accepted by scientists in D.[11] We assume that the signature Σ of $\mathcal{V}$ has a *finite* number of theoretical terms, predicates, and functions. This model's UAT space comprises all possible combinations between logical terms, constants, predicates, and functions as defined by T and all theories in $\mathcal{T}_T$. In this sense, the space of possible combinations is countable, given the infinite number of combinations among logical terms (even if the number of constants, predicates, and functions is finite). We focus here exclusively on the countable case, in which the signatures of our vocabulary are countable. The PUA idea is that even if the number of possible combinations is finite, scientists cannot grasp at t the combinations between, let us say, the theoretical objects of their theories and the possible (but conceivable) set of predicates.

In this simplified view, a scientific theory consists of a set of theoretical terms and predicates that can be attributed to these terms, along with all the logical consequences that can be inferred. For example, a theory T can quantify over two theoretical terms c_1 and c_2 with several predicates of any n-arity: $\{P_1, P_2, \dots P_n\}$ will have a simple signature $\sigma_T = \langle \{c_1, c_2\}, \{P_1, P_2 \dots P_n\} \rangle$.[12]

[10] The signature is the part of the vocabulary that contains all the constants c, predicates P, and functions F. In our approach the scientific domain D uses the same vocabulary, but theories may have different signatures $\sigma_{T_1}, \sigma_{T_2}, \dots$.

[11] We take here a syntactic view about entailment in D and assume simple forms of entailment in $\mathcal{L}$ from a theory T such as: $T \nvdash_D \bot$ for 'T is true' and $T \vdash_D \bot$ for 'T is not true.' This implies that the language of D comes with rules of entailment.

[12] Here, we ignore the observational terms of T as they do not play a clear role in PUA. We also ignore the correspondence functions that relate theoretical terms to observational terms, unless otherwise stated. We characterize the relation of a theory with evidence in Bayesian probabilistic terms.

3. Inconceivability of theories: three scenarios

The main question is: how do we define the inconceivability of T over the space spanned by $\mathcal{T}_T$? The space-inspired accounts (B. van Fraassen, P. Gärdenfors, F. Zenker, *i.a.*) are helpful to our approach: the space of $\mathcal{T}_T$ is larger than what the scientific community can conceive.[13] This paper presents several scenarios for the inconceivability of theories, ranked from strong to weak, and applies the eliminative inference to a weak form of inconceivability.

3.1. **Scenario 1: *terminological & ideological* (full) inconceivability**

The strongest model for what Sklar and Stanford might imply about our (recurrent) inability to exhaust the space of alternatives to a theory is based on our failure to exhaust the space of theoretical terms c_i, the space of the predicates P_j, and that of functions F_k. Given E at t, the scientists fail to connect the correct theoretical terms with the appropriate predicates or functions. In this scenario, the scientist(s) may have limited access to the space of theoretical terms. They cannot imagine enough relevant, serious, and meaningful alternatives to T in $\mathcal{T}_T$ because the alternatives to a given theory may use different theoretical entities.

In a very simplified version, if a theory T with a signature $\sigma_T = \langle c_1, P_1 \rangle$ is composed of a constant c_1 and a predicate P_1 that applies to c_1, then one of its alternative $T' \in \mathcal{T}_T$ has a different constant c_2, a different predicate P_2 and another signature $\sigma_{T'} = \langle c_2, P_2 \rangle$. Theory T may claim that $P_1(c_1)$ is true, whereas T' claims that $P_2(c_2)$ is true. T and T' are consistent in this simple case as they have disjunctive signatures and any combination of them can be conducive to truth. But the conceivability of T' is a difficult epistemic process. Scientists need to "jump" in the conceptual space from the point c_1 with property P_1 to a different point c_2

[13] Van Fraassen's and Gärdenfors' 'state-space' and 'conceptual space' are different and may serve different purposes, but we use both approaches here. See (Bueno 2015; Gardenfors 1990; Gärdenfors 2014; Van Fraassen 2008).

with a new property P_2. The inconceivability of T' relative to T means that given the posits of D at t and various constraints on the epistemic reach of the scientists practicing D, this 'jump' is unlikely at t. Together with P. Gärdenfors, we assume that points in conceptual spaces have properties, and regions of points with the same property form a subspace. The distance between areas of the extensions of P_1 and that of P_2 is a central concept in this approach, as it correlates with the probability of T′ being conceivable (although not accepted yet) coming from the space of T. We also need to postulate that T and T' are empirically equivalent in that they are supported by the same empirical evidence E.[14] Let us call this the "terminological&ideological inconceivability".[15] We can assume that this scenario is a case of a serious UAT for Stanford.

3.2. Scenario 2: *terminological* inconceivability

A weaker inconceivability scenario occurs when scientists currently use a set of predicates and functions, but lack the correct theoretical term(s) to be predicated of. Scientists can still conceive the relevant predicates P_i or functions F_i, although they are predicated about the improper theoretical terms. This scenario can be called a "term inconceivability." For example, if T is the wrong theory at present, with signature $\sigma_T = \langle c_1, P_1 \rangle$ and it claims $P_1(c_1)$ (that c_1 is in the extension of P_1), then the correct, alternative theory T' with $\sigma_{T'} = \langle \{c_1, c_2\}, P_1 \rangle$ will state correctly that a different term c_2 is in the extension of the predicate P_1: $P_1(c_2)$ and state that $\sim P_1(c_1)$. T and $T'$are inconsistent: $T \wedge T' \vdash_D \perp$, although T cannot assign truth values to $P_1(c_2)$: $T \wedge P_1(c_2) \nvdash_D \perp$ and $\sim (T \wedge P_1(c_2) \nvdash_D \perp$.

[14] We do not state here what it means to be empirically supported by evidence. Still, in general, this can be couched in terms of correspondence functions between theoretical and observational terms. It is possible that T and T' need to share a set of observational terms and have their own correspondence functions.

[15] We use ideology here in a restrictive sense inspired by Quine: the ideology of a theory is the list of n-place predicates used by that theory (Quine 1951).

3.3. Scenario 3: *ideological* inconceivability

There is another way in which scientists cannot conceive an alternative to T by operating with a different set of predicates and functions on the same set of theoretical terms. Here, scientists do not have the whole ideology available when considering all possible alternatives to a theory, although the same set of theoretical terms is used. This is a form of "ideological inconceivability." For example, if $\sigma_T = \langle c_1, P_1 \rangle$, then an alternative theory T′ in $\mathcal{T}_T$ uses the same theoretical term c_1 with a different predicate P_2: $T' = \langle c_1, P_2 \rangle$. In this case, the truth value of $P_2(c_1)$ is true which, according to PUA is equally supported by E at t and will be accepted and known at at $t_1 > t$.

Think of a Kuhnian example: the properties of mass in Newtons and Einsteins theories of gravity. The same theoretical term has different properties, but these meanings can be compared and contrasted with one another. Although the geometric and topological properties of the Newtonian and relativistic spacetimes overlap at the lower velocity limit, some of their properties are fundamentally different, the most obvious being the mass dependence on the velocity. As before, the theories T and T' can be empirically equivalent, for a given set of data (in this case, for low velocities compared to the speed of light).

3.4. Scenario 4: *predication* inconceivability

Last, the weaker form of inconceivability is determined by the inability to predicate a known (conceivable in principle) property about a known theoretical entity. The scientists possess the proper theoretical terms c_i, the proper predicates P_j and functions F_k, but cannot make the appropriate predication. A community of cognitive agents may have limited ability to relate the predicates to theoretical terms correctly. In this sense, this is an inconceivability of alternatives due to the incomplete set of possible predications, when the theoretical terms are the right ones, as well as the predicates and functions. One potential way of expanding the extension of a given predicate is by conjecturing that two different theories refer to the same theoretical term, which therefore obeys the same set of rules of inference. Imagine we have a set of theoretical terms and a set of predicates and two theories that operate on them, but attribute different truth values to the same claims: $\sigma_T = \langle c_1, c_2, P_1, P_2 \rangle$, such that T states that $P_1(c_1) \wedge P_2(c_2)$ is true or in syntactic notation: $P_1(c_1) \wedge P_2(c_2) \nvdash_D \bot$ and one of its alternatives T' with the same signature, $\sigma_{T'} = \langle c_1, c_2, P_1, P_2 \rangle$ states that, syntactically, $P_1(c_2) \wedge P_2(c_1) \nvdash_D \bot$. Scientists in D are unable to conceive (or understand) $P_1(c_2)$ and $P_2(c_1)$, so they do not attribute meaning or truth values to these statements.

We believe that all the scenarios above accurately characterize *PUA*, but the following section focuses on a version of scenario 4, specifically the weak "predication inconceivability." To expand their predictive capabilities, scientists can identify the theoretical terms used by two theories and envision possible unification, whether ideal or real. However, a quick remark about scenarios 1-3 is in order. The cases in which new terms or predicates are needed are probably too strong for the eliminative inference proffered here. This eliminative inference does not apply in cases of major conceptual and nomological revolutions in science. This paper suggests that certain situations, including those discussed by Stanford, are susceptible to eliminative inferences and are closer to scenario 4 than to stronger forms of inconceivability, such as scenarios 1, 2, or 3.

4. *PUA* in the intertheoretical mill: unification, fragmentation, effective fields, and inconsistency

Unification is the virtue of a new scientific theory, T_σ (or hypothesis), to represent multiple phenomena that seemed unrelated before the introduction of T_σ. This new theory, T_σ, is created by combining two existing theories, T_1 and T_2, which differ in their theoretical terms, predicates, functions, and empirical support, E_1 and E_2, and each has its own U function that creates two sets of UAT: $\mathcal{T}_{T_1}$ and $\mathcal{T}_{T_2}$.

The possibility that two different theories can be unified acted both as a theoretical posit and as a concrete accomplishment in the history of science after the Scientific Revolution. One can read the history of science as a partial history of successive unifications. Still, the story of physics in the 20th and 21st centuries can hardly be told without stressing the desire for unification: Einstein's unified field theory, various Grand Unified Theories (*GUT*), Supersymmetry, Superstring Theory, Canonical Quantum Gravity, and many more. Other theories, such as statistical mechanics, quantum mechanics, or quantum field theory, also have unification as one of their motivations.[16]

There are cases of unification akin to scenario 4 in biology. Foremost, the Modern Synthesis in biology brought together previously fragmented biological subfields into a coherent framework centered on the concept of evolution. Previous areas, such as genetics, paleontology, systematics, and embryology, operated largely independently, with conflicting theories about how life evolved and radically different empirical support. The progress between the 1930s and 1950s integrated Mendelian genetics with Darwinian evolution, solving disputes about the mechanisms of inheritance and natural selection.[17] The modern synthesis has eliminated contradictions between genetics and evolution, demonstrating that mutations and recombination provide the raw material for natural selection, and has linked microevolution to macroevolution, thereby bridging the gaps between genetics and paleontology. As B. Mayr suggested, the synthesis straightened out conflicts and disagreements between genetics and evolution, so "a united picture of evolution emerged" (Smocovitis 1992). This unified framework remains the foundation of modern evolutionary biology.

In physics, even given possible troubles in the paradise of unification, most physicists would endorse an architectonic representation of known interactions that can be read as a progressive history toward unification. After confirming the existence of four fundamental physical

---

[16] A book-length analysis of unification in physics and biology is (Morrison 2000). Recent analyses based on explanation, theories of truth, and Bayesianism are: (Bangu 2017; Blanchard 2018; Patrick 2018; Schupbach 2005).

[17] Key figures like Th. Dobzhansky, E. Mayr, J. Huxley, and G. Simpson played crucial roles in demonstrating how genetic variation and selection drive evolutionary change. See (Morrison 2000; Plutynski 2005; Smocovitis 1992).

forces—all the other forces being merely apparent or derivative from these: electromagnetism (being already unified), gravity, the strong nuclear force, and the weak nuclear force, in the first half of the 20th century and developing accurate theories of these forces for each of them: "the aim of physics is now to produce theories which unify these forces, which show, ultimately, that there is at base only one fundamental force in the universe, which has come to display itself as if it were many different forces" (Maudlin 1996, 129). This is the intuition that disparaged empirical phenomena E_1 and E_2 may be explained by a common 'structure' for which scientists strive to find a representation within theory T_U.

Within theoretical physics itself, unification can be understood in several ways. For example, some unificatory programs were designed to unify fundamental fields, while others aimed to unify matter with fields, and yet others were premised on even stronger assumptions and endeavored to unify gravity with all the other known quantifiable fields. One can see successful unifications in physics and biology. S. Glashow suggested that in the 1950s, after the massive success of quantum field theories, physics was "patchy":

> The study of elementary particles was like a patchwork quilt. Electrodynamics, weak interactions, and strong interactions were clearly separate disciplines, separately taught and separately studied. There was no coherent theory that described them all. Developments such as the observation of parity-violation, the successes of quantum electrodynamics, the discovery of hadron resonances and the appearance of strangeness were well-defined parts of the picture, but they could not be easily fitted together (Glashow 1980, 539).

However, is unification a general principle in science? Today, enthusiasm for unification is less common among biologists and chemists, where fragmentation in specialized fields may be more pronounced in the sciences. There is significant fragmentation in specific disciplines, such as molecular genetics or oxidative metabolism, partly because the same processes do not operate uniformly across all orders of life or in the same manner. Nonetheless, some believe that biology has reached a level at which a steady consolidation process will replace the fragmentation process. The most enthusiastic scientists see 'consolidation' as a sign of unification:

> Scientific progress is based ultimately on unification rather than fragmentation of knowledge. At the threshold of what is widely regarded as the century of biology, the life sciences are undergoing a profound transformation. They have long existed as a collection of narrow, even parochial, disciplines with well-defined territories. Now they are undergoing consolidation, forming two major domains: one extending from the molecule to the organism, the other bringing together population biology, biodiversity studies, and ecology. Kept separate, these domains, no matter how fruitful, cannot hope to deliver on the full promise of modern biology. They cannot lead to an appreciation of life in its full complexity, from the molecule to the biosphere, nor to the generation of maximal benefits to medicine, industry, agriculture, or conservation biology (Kafatos and Eisner 2004, 1257).

4.1. The 'austere unification' by term identification

Pairs of theories in Υ_E with two sets of empirical support (*E1* and *E2*), are potential candidates for unification, reduction, equivalence relation, approximation, and so forth, even if this is not yet accomplished. But each of these known theories has their own set of $\mathcal{T}_{T_1} = U_t(T_1)$ and $\mathcal{T}_{T_2} = U_t(T_2)$ raise an important question: what if in the future T_1 will enter into a inter-theoretical relation? Following the line of F. C. Kafatos and T. Eisner, the current proposal investigates the potential future unification of two unrelated theories, especially when no new predicates are needed to account for alternatives to a theory. When two theories are unified the new theory contributes to scientific progress, even when the unification is not fully realized (P. Kitcher 1999).

Assume that in *D*, there are two known and accepted theories with these signatures $\sigma_{T_1} = \langle c_1, P_1 \rangle$ and $\sigma_{T_2} = \langle c_2, P_2 \rangle$, with their two UAT of unconceived alternatives spaces $\mathcal{T}_1 = U_t(T_1)$ and $\mathcal{T}_2 = U_t(T_2)$. Within scenario 4, scientists can conceive the predicates needed for an alternative, but they cannot link them to the appropriate theoretical terms. Unification through identification is a mechanism that relates two different domains of inquiry when the theoretical term c_1, quantified by T_1, is essentially the same as c_2, quantified by T_2:$c_1 = c_2$. In this way, the predicate P_1 from T_1 can now be used by T_2.

Originally, before we identified c_1 with c_2, T_1 and T_2 differed in their theoretical claims. T_1 states that $P_1(c_1)$ and T_2 states that $P_2(c_2)$. In this sense, both P_1 and P_2 are conceivable ideologies, and c_1 and c_2 are conceivable terms, but what is not conceivable are $P_1(c_2)$ and $P_2(c_1)$.

In this toy example, the number of alternatives is limited. It is also essential to see that other alternatives to T_1 are: $T_1' = \langle c_2, P_1 \rangle$ stating that $P_1(c_2)$ and $T_1^{"} = \langle c_2, P_1 \rangle$ stating that $\sim P_1(c_2)$.

The eliminative inference is based on the idea that the unification achieved by identifying c_1 with c_2 eliminates alternatives to T_1 and T_2. Remember that both terms c_1 and c_2 are conceivable at this moment, and by this identification, we enlarge the space of our predication with these two new sentences: $P_1(c_2)$ and $P_2(c_1)$. However, we now have an inconsistency between pairs T_1' and T_2'. This inconsistency indicates that the space of alternatives to T_1 and T_2 is reduced after unification by identification.

4.2. PUA, UAT, the 'lush unification,' and minimal inconsistency with evidence

As *PUA* depends on the existence and the 'size' of *PUA* space $\mathcal{T}_T$ the antirealist can point to the inconsistency of the T with any (all?) of its alternatives: $\exists T'(T' \in \mathcal{T}_T \wedge \sim (T \wedge T')$. As we have seen in the previous examples, consistency in itself is not enough to restrain *PUA*, even when T and its alternatives are supported empirically by the same evidence. The case in which $\Upsilon_E$includes two known theories, which potentially can be unified, is worth exploring further. However, it is the role of empirical evidence within expected unification posit that can diminish the strength of the *PUA*. In a more idealistic case, the scientists hope a new theory T_σ, different from T_1 and T_2 will reduce the inconsistency between them and unify their empirical bases $E_1 \wedge E_2$ better than them taken individually:

$$pr(T_\sigma|E_1 \wedge E_2) > pr(T_1|E_1) \times pr(T_2|E_2)$$

In the least problematic case, the empirical evidence of the unifying theory T_σ should be the union of E_1 and E_2, and we do not assume that unification is triggered or conditioned by the occurrence of new empirical data beyond E_1 and E_2. Unification is the emergence of T_σ with its own new signature.

This is referred to here as the 'lush unification' of T_1 and T_2, when scientists hope that T_σ will subsume vastly different phenomena under the same theoretical framework. Although T_σ may or may not contain identification of theoretical terms, it is expected, hoped for, and deemed as a new theory. It comes with its own range of unconceived alternatives $\mathcal{T}_{T_\sigma}$, but this does not preclude us from comparing the $\mathcal{T}_{\mathcal{T}_1}$ and $\mathcal{T}_{\mathcal{T}_1}$ before and after the emergence of T_σ.

In this case of "lush unification," empirical evidence "cross-pollinates" into the space of alternative theories of the pair T_1 and T_2. Let us put this in the form used before: we can conditioned the minimal set of T_1 by evidence E_2 because of T_σ, and compare it with the minimal set of T1 without evidence E_2. This is a result that we do not demonstrate here:

$$\mathrm{MI}\left(T_1, \mathcal{T}_{T_1}, E_2\right) \supset \mathrm{MI}\left(T_1, \mathcal{T}_{T_1}\right) \text{and } \mathrm{MI}\left(T_2, \mathcal{T}_{T_2}, E_1\right) \supset \mathrm{MI}\left(T_2, \mathcal{T}_{T_2}\right)$$

therefore, the cardinality of the minimal inconsistency set of any of the two theories is reduced by the present of the evidence used by the other theory, when the two are unified by T_σ. The cardinality of the $\mathrm{MI}(\{T, \mathcal{T}_T\})$ is taken here as being strongly correlated with what Stanford means by "serious" UAT. A high $card\left(\mathrm{MI}(\{T, \mathcal{T}_T\})\right)$ means that given E, there is a significant number of unconceived alternatives to T. The "lush unification" creates a new unification theory T_σ that brings together previously unrelated theories and use the evidence of all these theories "collectively."

4.3. Austere and lush unification in search of some case studies

This paper is particularly focused on the unificatory ideal of science, arguing that *when* and *where* the unificatory ideal operates in a community of scientists, the semantic eliminative inference is more powerful. As stated here, the unificatory posit is not the same as the logical positivist "unity of science" as we operate in *PUA* at the local level in D, rather than as a whole science. We know that any *D* has several theories that share some common features and some empirical evidence. The claim is that a community of scientists who adopts a local unificatory posit, as local as limited as it may be, is less prone to the problem of unconceived alternatives. A community of scientists who see every theory in D as insular and isolated will be more vulnerable to the issue of unconceived alternatives. Excessive semantic pluralism is conducive to stronger antirealism instrumentalism and a stronger *PUA* (Ruhmkorff 2019).

The paradigmatic case that comes to mind is the unification of the theory of light and the theory of electromagnetic waves. One way to unify these two domains is to postulate an identity between the light wave and the electromagnetic wave, a hypothesis advanced by Maxwell through the introduction of the displacement current, a theoretical term. A step further, Maxwell could identify two other theoretical terms: the luminiferous ether and the electromagnetic ether. This introduced new "ideologies" to the previously unrelated theories of electromagnetic waves and optics, in the sense that optical concepts were applied to electromagnetic waves and vice versa. More importantly, none of the previous theories were surveyed unscathed: Ampere's law was modified, and some aspects of interference in optics were adjusted accordingly. This is a case of non-reductive unification, in which neither of the two theories reduces the other (Morrison 2000, 78). We can imagine that this process eliminates some alternatives to theories in optics and electromagnetism, such as theories about the transfer of energy in the two ethers, the speed of propagation in the two media, and ultimately the very idea of polarization of light and EM waves. In all these cases, constraints from optics "shaved off" alternatives in electromagnetism and the other way around, such that the *UAT* space of the new theory was reduced.

The second example comes from the debate about variation in a biological population. Before the 1900s, there were mainly two theories, each with its own followers. The Darwinians, a group that included A. Weismann and Fr. Galton, believed that selection alone produced the change from one generation to the next. The Mendelians (a much smaller group, represented by W. Bateson and his competitor W.F.R. Weldon) believed that something other than selection was the leading cause of this change: mutation was a possible candidate. However, heritable traits and natural selection, as theoretical terms, were considered. Staunch Darwinians were gradualists and rejected the idea that mutation can play any role in evolution. The synthesis of these two theories was made possible much later, when people began to consider the genetic basis of evolution. This likely occurred with the work of K. Pearson, S. Wright, and R. Fisher in the late 1920s. The unification posit was to conceive that at least a slight selection pressure and heredity could contribute *together* as explanatory factors of change in population. Therefore, more recent theories of inheritance would attribute predicates such as generation, inheritance, growth, and development, which are all present in the Darwinian pangenesis theory, to a different theoretical term: a shared germinal source, or "hereditary particles," and not to the development of an organism's tissues (Stanford 2006, 68).

More importantly, once Weldon started to disagree with Bateson on the foundations of Mendelianism, he built his alternative on Francis Galton's 'ancestral heredity,' in which hereditary information from a distant ancestor is reduced by half with each generation and mixed during mating. Unlike the discrete, binary Mendelian traits, Weldonian traits vary continuously and follow a normal distribution. This constituted a 'conceived but unaccepted' alternative to Batesons dominant view of genetic determinism. Even more enticing is to consider Weldon's theory, following his untimely death in 1906, as an unexplored alternative to Bateson's theory. Gr. Radick has recently explored this counterfactual history (Radick 2022). Had Weldon lived, he might have produced a different synthesis of evolution and genetics. The fictional "Weldonian genetics" would have been more unifying than Bateson's genetics because it would eliminate alternatives to Mendel's theory that are more incompatible with evolution.

Finally, one can summon a third example of 'clipping' the *UAT* space from the quantum gravity program. As both quantum physics and general relativity lack the necessary generality, physicists frequently conjecture that the two theories will eventually be unified in an integrated quantum gravity theory. This unificatory posit is more often marshalled than the instrumentalist view about the theories or the sheer search for their 'corrective' alternatives. One plausible venue here is scenario 1, where a new conceptual basis and ideology will be needed. The discrete geometry of space, the separation of space and time, the emergence of spacetime, dualities, and holographic principles could potentially be components of such a new theory. However, one can also speculate that we are closer to scenario 4 in current physics, which endorses more conservative unification and a less disruptive revolution. In such cases, the theoretical terms of quantum mechanics and general relativity "work together" but are predicated on different principles. Therefore, one can think that quantum and relativity elements would generalize well in a subsequent program and explore how those elements function together to generate the whole structure of the new theory: "Next one can explore the more general structures that can be obtained by loosening the constraints imposed in the current theory on one or another of the components that goes to make up the theoretical framework" (Lawrence Sklar 2000, 112).

If scenario 4 applies to quantum gravity, then one can see that, based on constraints from the other domain, unconceived alternatives, such as quantum theory, are constrained by requirements from general relativistic considerations, and vice versa. When quantum physicists can identify theoretical terms from quantum theories with theoretical terms from general relativity (e.g., entropy, energy, information, etc.), the quantum alternatives are constrained and restricted, even if quantum alternatives are still unconceived.

4.4. Coda on language dependence and conceptual spaces

There are several loose ends to this argument. First, scenarios 1, 2, and 3 are more frequently encountered in the history of science. Second, one can only speculate that the more pluralist fragmentation posit, somewhat the opposite of the unificatory posit, if dominant in an epoch, would enlarge the number of possible alternatives to a theory and make science more vulnerable to Stanford's PUA. Whether we live in the fragmentation of the scientific epoch or a more unificatory one is an empirical question, tricky to address here. This paper has no issue with the fragmentation of science, but it shows that *NI* is weakened when other perspectives dominate a scientific discipline. If the current science is dominated by the disunity posit, the *PUA* space is augumented by fragmentation.

Second, another problem with this approach is its dependence on the first-order language, including its vocabulary and semantics. Moving to a semantic view may solve this problem, as models are not linguistic units of analysis. The 'conceptual spaces' mentioned before and their recent incarnations are promising candidates in this respect (Gärdenfors 2000, 2014; Zenker and Gärdenfors 2015). Theoretical terms are not only regions of conceptual space but they can be equipped with a geometry (convexity) and a metric. Geometric, non-linguistic representation can represent knowledge and the inconceivability of alternatives to theories. Gärdenfors and his collaborators claim that the qualities of objects (mainly their theoretical terms) can be represented without presuming an internal language. As a prospective alternative to the semantic approach, the conceptual space approach enables a more robust evaluation of what it means to embed and, ultimately, unify two theories within a larger and richer theoretical structure.

5. Conclusion

The present proposal attempts to weaken K. Stanford's problem of unconceived alternatives (*PUA*) by showing that some posits (called in the antirealist literature "standards" by M. Massimi or "substantive assumptions" by Stanford), such as unification and consistency, when adopted even tacitly in domain *D*, reduce the relevance and number of alternatives to *D*'s accepted theories. To do so, we need to think more holistically and see a theory T_1 as part of a field of theories in its unconceived alternatives enter in an intertheoretical relationship based on how T_1 relates to other accepted theory (T_2) in *D*. In this proposal, alternatives to T_1 are associated with alternatives to T_2, at least in the case of the most conservative case where T_1 and T_2 share the vocabulary but have different signatures (scenario 4 above). In the case of a "lush unification," scientists posit the hypothesis that the two theories, T_1 and T_2, will be unified and replaced with a new theory T_σ that will bring the empirical support E_1 and E_2 together. This will reduce the space of alternatives to both T_1 and T_2. The overall goal of the present argument is to ease the PUA for cases where scientists in domain *D* endorse some (normative) 'expected posits' such as unification, consilience, or parsimony of future theories in *D*.

6. References

Bangu, S. (2017). Scientific Explanation and Understanding: Unificationism Reconsidered. *European Journal for Philosophy of Science* 7 (1): 103–26. https://doi.org/10.1007/s13194-016-0148-y.

Benferhat, S., Dubois, D., & Prade, H. (1997). Some Syntactic Approaches to the Handling of Inconsistent Knowledge Bases: A Comparative Study Part 1: The Flat Case. *Studia Logica* 58 (1): 17–45. https://doi.org/10.1023/A:1004987830832.

Bhakthavatsalam, S., and Kidd, I. J. (2019). Science, Realism, and Unconceived Alternatives: Introduction to the Special Issue on Unconceived Alternatives. *Synthese* 196 (10): 3911–13. https://doi.org/10.1007/s11229-019-02353-z.

Blanchard, T. (2018). Bayesianism and Explanatory Unification: A Compatibilist Account. *Philosophy of Science* 85 (4): 682–703. https://doi.org/10.1086/699157.

Bueno, O. (2015). Belief Systems and Partial Spaces. *Foundations of Science* 21 (1): 225–36. https://doi.org/10.1007/s10699-015-9416-0.

Carnap, R. (1995/1966). *An Introduction to the Philosophy of Science*. New York: Dover Publications.

Cartwright, N. (1999). *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.

Chakravartty, A. (2008). What You Don't Know Can't Hurt You: Realism and the Unconceived. *Philosophical Studies* 137 (1): 149–58. https://doi.org/10.1007/s11098-007-9173-1.

Da Costa, N. C., & French, S. (2003). *Science and Partial Truth: A Unitary Approach to Models and Scientific Reasoning*. Oxford University Press. 273 pp.

Dellsén, F. (2019). Should Scientific Realists Embrace Theoretical Conservatism? *Studies in History and Philosophy of Science Part A* 76 (August): 30–38. https://doi.org/10.1016/j.shpsa.2018.09.005.

Devitt, M. (2011). Are Unconceived Alternatives a Problem for Scientific Realism? *Journal for General Philosophy of Science* 42 (2): 285–93. https://doi.org/10.1007/s10838-011-9166-9.

Dupré, J. (1993). *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*. Cambridge, Mass.: Harvard University Press.

Forber, P. (2007). Forever Beyond Our Grasp?: Review of P. Kyle Stanford (2006), Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives. *Biology & Philosophy* 23 (1): 135–41. https://doi.org/10.1007/s10539-007-9074-x.

Friedman, M. (1974). Explanation and Scientific Understanding. *Journal of Philosophy* 71 (17): 5–19.

Friedman, M. (2001). *Dynamics of Reason*. Center for the Study of Language and Inf.

Gardenfors, P. (1990). Induction, Conceptual Spaces, and AI. *Philosophy of Science* 57 (1): 78–95.

Gärdenfors, P. (2000). *Conceptual Spaces the Geometry of Thought*. A Bradford Book. Cambridge, Mass: MIT Press.

Gärdenfors, P. (2014). *The Geometry of Meaning: Semantics Based on Conceptual Spaces*. The MIT Press. https://doi.org/10.7551/mitpress/9629.001.0001.

Glashow, S. L. (1980). Towards a Unified Theory: Threads in a Tapestry. *Reviews of Modern Physics* 52 (3): 539. http://dx.doi.org/10.1103/RevModPhys.52.539.

Godfrey-Smith, P. (2007). Recurrent Transient Underdetermination and the Glass Half Full. *Philosophical Studies* 137 (1): 141–48. https://doi.org/10.1007/s11098-007-9172-2.

Halvorson, H. (2013). The Semantic View, If Plausible, Is Syntactic. *Philosophy of Science* 80 (3): 475–78. https://doi.org/10.1086/671077.

Hartmann, S. (2001). Effective Field Theories, Reductionism and Scientific Explanation. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 32 (2): 267–304. https://doi.org/10.1016/S1355-2198(01)00005-3.

Hempel, C. (1966). *Philosophy of Natural Science*. 1st ed. Pearson.

Kafatos, F. C., and T. Eisner. (2004). Unification in the Century of Biology. *Science (New York, N.Y.)* 303 (5662): 1257.

Kao, M. (2019). Unification Beyond Justification: A Strategy for Theory Development. *Synthese* 196 (8): 3263–78. https://doi.org/10.1007/s11229-017-1515-8.

Kitcher, P. (1999). Unification as a Regulative Ideal. *Perspectives on Science* 7: 337–48.

Kitcher, P. (1981). Explanatory Unification. *Philosophy of Science* 48 (December): 507–31.

Kitcher, P. (1995). *The Advancement of Science: Science Without Legend, Objectivity Without Illusions*. Oxford University Press.

Kukla, A. (2010). A. Kukla's Review of 'Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives', by P. Kyle Stanford. *Mind* 119 (473): 243–46.

Lutz, S. (2014). What's Right with a Syntactic Approach to Theories and Models? *Erkenntnis* 79 (S8): 1475–92. https://doi.org/10.1007/s10670-013-9578-5.

Lutz, S. (2017). What Was the Syntax-Semantics Debate in the Philosophy of Science About? *Philosophy and Phenomenological Research* 95 (2): pp. 319–352. https://www.jstor.org/stable/48578792.

Magnus, P. D. (2006). What's New about the New Induction? *Synthese* 148 (2): 295–301. https://doi.org/10.2307/20118691.

Massimi, M. (2021). Realism, Perspectivism, and Disagreement in Science. *Synthese* 198 (S25): 6115–41. https://doi.org/10.1007/s11229-019-02500-6.

Maudlin, T. (1996). On the Unification of Physics. *Journal of Philosophy* 93 (3): 129–44.

Mizrahi, M. (2022). Theoretical Virtues in Scientific Practice: An Empirical Study. *British Journal for the Philosophy of Science* 73 (4): 879–902. https://doi.org/10.1086/714790.

Morrison, M. (2000). *Unifying Scientific Theories: Physical Concepts and Mathematical Structures*. Cambridge University Press.

Patrick, K. (2018). Unity As An Epistemic Virtue. *Erkenntnis* 83 (5): 983–1002. https://doi.org/10.1007/s10670-017-9923-1.

Plutynski, A. (2005). Explanatory Unification and the Early Synthesis. *The British Journal for the Philosophy of Science* 56 (3): 595–609. http://bjps.oxfordjournals.org/cgi/content/abstract/56/3/595.

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

Quine, W. V. (1951). Ontology and Ideology. *Philosophical Studies* 2 (1): 11–15. https://doi.org/10.1007/BF02198233.

Quine, W. V., and J. S. Ullian. 2007. *Ţesătura Opiniilor. (*Romanian edition of *Web of Belief)*. Translated by Mircea Dumitru. Pitești, Romania: Paralela 45.

Radick, G. (2022). *Disputed Inheritance: The Battle over Mendel and the Future of Biology*. Chicago, IL: The University of Chicago Press.

Roche, W. and Sober, E. (2017). Explanation = Unification? A New Criticism of Friedman's Theory and a Reply to an Old One. *Philosophy of Science* 84 (3): 391–413. https://doi.org/10.1086/692140.

Rowbottom, D. P. (2019). Extending the Argument from Unconceived Alternatives: Observations, Models, Predictions, Explanations, Methods, Instruments, Experiments, and Values. *Synthese* 196 (10): 3947–59. https://doi.org/10.1007/s11229-016-1132-y.

Ruhmkorff, S. (2019). Unconceived Alternatives and the Cathedral Problem. *Synthese* 196 (10): 3933–45. https://doi.org/10.1007/s11229-015-0947-2.

Saatsi, J. (2019). Historical Inductions, Old and New. *Synthese* 196 (10): 3979–93. https://doi.org/10.1007/s11229-015-0855-5.

Saatsi, J., Psillos, S., Winther, R. G., & Stanford, P. K. (2009). Grasping At Realist Straws. *Metascience* 18 (3): 355–90. https://doi.org/10.1007/s11016-009-9299-1.

Schindler, S. (2018). *Theoretical Virtues in Science: Uncovering Reality Through Theory*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108381352.

Schindler, S. (2022). Theoretical Virtues: Do Scientists Think What Philosophers Think They Ought to Think? *Philosophy of Science* 89 (3): 542–64. https://doi.org/10.1017/psa.2021.40.

Schupbach, J. N. (2005). On a Bayesian Analysis of the Virtue of Unification. *Philosophy of Science* 72 (4): 594–607.

Sklar, L. (1981). Do Unborn Hypotheses Have Rights? *Pacific Philosophical Quarterly Anc Personalist (The) Los Angeles, Cal.* 62 (1): 17–29. http://cat.inist.fr/?aModele=afficheN&cpsidt=12557219.

Sklar, L. (2000). *Theory and Truth: Philosophical Critique Within Foundational Science*. Oxford; New York: Oxford University Press.

Smocovitis, V. B. (1992). Unifying Biology: The Evolutionary Synthesis and Evolutionary Biology. *Journal of the History of Biology* 25 (1): 1–65. http://www.jstor.org/stable/4331201.

Stanford, P. K. (2006). *Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives*. New York: Oxford University Press, USA.

Stanford, P. K. (2009). Scientific Realism, the Atomic Theory, and the Catch-All Hypothesis: Can We Test Fundamental Theories Against All Serious Alternatives? *Br J Philos Sci* 60 (2): 253–69. https://doi.org/10.1093/bjps/axp003.

Stanford, P. K. (2015a). Catastrophism, Uniformitarianism, and a Scientific Realism Debate That Makes a Difference. *Philosophy of Science* 82 (5): 867. https://doi.org/10.1086/683325.

Stanford, P. K. (2015b). Unconceived Alternatives and Conservatism in Science: The Impact of Professionalization, Peer-Review, and Big Science. *Synthese* 196 (August): 3915–32. https://doi.org/10.1007/s11229-015-0856-4.

Stanford, P. K. (2021). Realism, Instrumentalism, Particularism: A Middle Path Forward in the Scientific Realism Debate. In *Contemporary Scientific Realism: The Challenge from the History of Science*, edited by Timothy D. Lyons and Peter Vickers, 0. Oxford University Press. https://doi.org/10.1093/oso/9780190946814.003.0010.

Van Fraassen, B. C. (2008). *Scientific Representation: Paradoxes of Perspective*. Oxford; New York: Clarendon Press; Oxford University Press.

Vickers, P. (2013). A Confrontation of Convergent Realism. *Philosophy of Science* 80 (2): 189–211. https://doi.org/10.1086/670297.

Votsis, I. (2007). J. Votsis's Review of 'Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives' by K. Stanford. *International Studies in the Philosophy of Science* 21 (1): 103–6.

Zamora Bonilla, J.P. (2019). Realism Versus Anti-Realism: Philosophical Problem or Scientific Concern? *Synthese* 196 (10): 3961–77. https://doi.org/10.1007/s11229-015-0988-6.

Zenker, F., & Gärdenfors, P., eds. (2015). *Applications of Conceptual Spaces The Case for Geometric Knowledge Representation*. Cham: Springer.