Are metaphysical claims testable?

Chrysovalantis Stergiou

The American College of Greece: Deree

cstergiou@acg.edu

[DRAFT]

Abstract

To consider metaphysical claims *a priori* and devoid of empirical content, is a rather commonplace received opinion. This paper attempts an exploration of a contemporary philosophical heresy: it is possible to test metaphysical claims if they play an indispensable role in producing empirical success, i.e. novel predictions. To do so one, firstly, needs to express the metaphysical claims employed in the logico-mathematical language of a scientific theory, i.e. to explicate them. Secondly, they should have an understanding of what it is to test and to verify or to falsify a metaphysical claim. Finally, they also need to consider the philosophical practice of testing a metaphysical claim. These three aspects are introduced in this paper and they are illustrated by means of the metaphysical concept of common cause and the principle of the common cause.

0. Introduction.

This paper concerns scientific metaphysics, the branch of metaphysics that deals with specific fields of knowledge and discourse and explores possible domains of existence as determined by scientific theories. In particular, I intend to investigate the testability of metaphysical claims when considered in the context provided by theories of empirical sciences and to explore the

prospects of what Hawley (2006) and French (2017:51) identified as *optimism* regarding the possibility of informing metaphysics by physics. Namely, the view that "…[t]here are actual cases in which the involvement of a metaphysical claim in an empirical successful scientific theory provides some reasons to think that the claim is true." Furthermore, source of inspiration for this work has been A. Shimony's (1984) and M. Redhead's (1987) quite provocative use of the term "*experimental metaphysics*" to refer to the philosophical implications of the experimentally confirmed violation of Bell-type inequalities in quantum mechanics, for our world picture.

I urge, firstly, that metaphysical claims are *empirically testable* only if they contain terms that can be *adequately* explicated in the language of a scientific theory. The adequacy conditions that guarantee the testability of a claim will be discussed and illustrated in terms of the principle of the common cause. Secondly, I suggest conditions for the verification and falsification of a testable metaphysical claim. Thirdly, issues regarding the philosophical practice of explicating and testing metaphysical claims are discussed. Among them the proliferation of explications of metaphysical concepts is examined with respect to two different attitudes: one that aims to devise new explications to circumscribe falsifying instances and evade conflicts with evidence, and the one I propose which, in the light of a falsifying instance for a given explication, aims at suggesting new explications to attain conclusive falsification of a metaphysical claim.

1. **On Explication**

Explication, according to Carnap (1950:3), is the transformation of an inexact, possibly prescientific concept, the *explicandum*, into a new exact concept, the *explicatum*, that obeys explicitly stated rules for its use. By means of this transformation a concept of ordinary discourse or a metaphysical concept may be incorporated into a well-structured body of logico-mathematical or empirical concepts. Explication has a long history as a philosophical method that, in a wide sense, may be traced back even to Plato's investigations on definitions. Strictly speaking, however, Carnap borrowed the term "*Explikation*" from Kant and Husserl while Frege

may be considered his precursor in this method of philosophical analysis[1] and Goodman, Quine and Strawson among his prominent intellectual inheritors. Various attempts to explicate "explication"[2], have found their way within the analytic tradition. In this paper, I have chosen to refer to a recent account, due to Cordes (2017), which, generally, is considered compatible with the Carnapian view despite some differences[3].

Cordes suggests that an explication (Df.3-9) is a six-tuple

$$E = \langle L_1(E),\ L_2(E),\ EX1(E),\ EX2(E),\ CEA(E),\ EI(E) \rangle.\ (1)$$

$L_1(E)$ is the *explicandum* language of E, i.e. the language in which the inexact concept or expression that one aims to explicate occurs. Ordinary English or the idiom of a particular philosophical current or even the scientific jargon of some discipline may be taken as examples of *explicandum* languages. $L_2(E)$ is the *explicatum* language of E i.e. the linguistic framework in which the expression that explicates the inexact concept belongs. In general, $L_2(E)$ is more precise than $L_1(E)$ and it can also be just "a more exact part of" $L_1(E)$ (Carnap 1963:935). $EX1(E)$ is an atomic expression of $L_1(E)$, the *explicandum* of E, while $EX2(E)$ is an atomic expression of $L_2(E)$, the *explicatum* of E. For example, the explicatum of "triangle" in a first order language may be a name (that a Platonist about mathematics would consider the name of an abstract object) or a unary predicate representing the property of being triangular. Thus, different representations of the *explicatum* in $L_2(E)$ may express different metaphysical commitments. The explicative introduction $EI(E)$ of E is a formula of $L_2(E)$ that contains the *explicatum*, $EX2(E)$, as an atomic subexpression. It includes the sentences that introduce the *explicatum* in $L_2(E)$, be they axioms regulating the use of the *explicatum* or definitions

---

[1] Beaney (2004) points out the absence of a systematic connection between Carnap and Kant or Husserl regarding explication and he discusses the influence of Frege's "Logic in Mathematics" lectures on Carnap's conception of it.

[2] "An explication of 'explication'" is the title of Hanna's article (1968) and of section 5 in (Cordes 2017). For a general presentation of the notion of explication, see (Cordes and Siegwart 2019).

[3] One major difference is that Cordes discusses the explication of expressions and considers the explication of concepts only subsequently as built upon his account of explication (2017: n.5) while Carnap discusses the explication of concepts. This difference, however, plays no important role for our discussion in this paper. For this reason, sometimes I talk about concepts without making any distinction.

introducing it explicitly or anything similar. $CEA(E)$ is a list of formulas, statements or sentences which constitute the criteria of explicative adequacy of E. The criteria of explicative adequacy include all sentences of $L_2(E)$ that one would desire to follow from the introduction of the explicated expression. They usually stem from pre-explicative intuitions associated with the explicandum and their satisfaction may be taken to express Carnap's requirement of similarity of the *explicatum* to the *explicandum* (1950:7). In the case of an explication of a metaphysical concept, metaphysical claims that are expected to be satisfied are included in the criteria of explicative adequacy. Since in a successful explication the criteria of adequacy should be satisfied given the *explicandum* introduction in $L_2(E)$, Cordes deem them a measure of success of the explication. Thus, he suggested that an explication E is *successful* if and only if for all $\psi \in CEA(E)$, $\{EI(E)\} \vdash_{L_2(E)} \psi$.(Df.10)

To illustrate his account, Cordes presents Carnap's explication of the term "successor" (1934:58-59) in terms of the following six-tuple:

$$\langle \text{English}, \text{Language I}, \text{"successor"}, \text{"nf"}, \{"nf(0) = 1"\}, "nf(x) = x^| \wedge 1 = 0^|" \rangle, \quad (2)$$

In this explication, English is the *explicandum* language and "successor" the *explicandum*. In ordinary English the term "successor" applies to an object and designates the next object in an ordered series. Language I, the *explicatum* language, is Carnap's language of definite number properties. This language includes that part of arithmetic of natural numbers in which the possession or non-possession of any property by a number can be determined in a finite number of steps using a fixed method (1934:11). This language does not designate objects in the universe of discourse by proper names but with positional coordinates; it is a coordinate language. Thus, the symbol "0" designates the initial position in a one-dimensional series with a definite direction; the symbol "$^|$" applied to any position designates the next position in this series; and "x" is a numerical variable. All these are logical primitive symbols of Language I. "nf" is the *explicatum* of the successor (*Nachfolger*) which is a function (functor) introduced by explicit definition D1: "$nf(x) = x^|$". The explicative introduction is a conjunction of D1 and

4

definition D7.1: ""$1 = 0^|$"". Thereby, the criterion of explicative adequacy of the explication "$nf(0) = 1$" is derived and the explication is deemed successful.

To examine the relation between the explication E of a metaphysical concept to a scientific theory T, I view the scientific theory as a *collection of formulas* in the *explicatum* language $L_2(E)$ (Cordes 2017: Df.12). In addition, it is required that the formulas of T are not making any use of the *explicatum*, since I purport to obtain a clear idea of what the process of explication adds, in terms of content, to theory T. The concept of background theory to an explication captures this *desideratum*: T is a background theory to E if and only if E is an explication and T is a theory in $L_2(E)$, and for all $\psi \in T$, the *explicatum* $EX2(E)$ is not an atomic subexpression of ψ (Cordes 2017: Df.13). Finally, to ascertain that the scientific theory has empirical content and it is not just a formal scaffold, let me point out that its formulas make claims about the physical world that are empirically testable.

The success of an explication E can be relativized to a background scientific theory T. As a consequence, an explication that is not successful *simpliciter*, i.e. there is a $\psi \in CEA(E)$, $\{EI(E)\} \nvdash_{L_2(E)} \psi$, can be rendered successful relative to a scientific theory since the theory may provide additional premises that enhance the inferential capacity of a given explicative introduction $EI(E)$. Thus, E is a successful explication for T if and only if T is a background theory to E and for all $\psi \in CEA(E)$, $T \cup \{EI(E)\} \vdash_{L_2(E)} \psi$ (Df.14). Moreover, E is a non-trivially successful explication for T if and only if E is a successful explication for T , and there is $\psi \in CEA(E)$ such that $T \nvdash_{L_2(E)} \psi$. (Df.15).

Cordes, in (2017), does not restrict the elements of $CEA(E)$ to sentences which contain the *explicatum* $EX2(E)$ as an atomic subexpression. This means that a non-trivially successful explication may be a creative act, which does not only bind some formulas from the theory's language to the act of making precise an inexact concept, but also produces new theoretical content, which can be formulated without using the *explicatum* although it cannot be inferred from T without the explicative introduction, $EI(E)$. Suppes (1957:154) provides an example of a creative introduction of a new symbol by means of a "definition". In his example, the theory T

consists of all algebraic structures $< G,\circ>$, where G is a set and $\circ$ an associative binary operation in G. The new symbol to be introduced as a distinguished element of G is the (right) identity or unit element e, which satisfies the condition: $\forall x \in G, x \circ e = x$. From this proposition, one may infer that $\exists y \in G, \forall x \in G, x \circ y = x$; eliminating, thus, any reference to the newly introduced symbol. Moreover, the latter proposition cannot be inferred by T alone. Hence, the introduction of the new symbol amounts to a creative act producing novel theoretical content. Such a "definition", for Suppes, is inadequate and should not be used for the introduction of a new symbol since the act of defining should be non-creative and regulated by the criterion of non-creativity. For the discussion in this paper, however, quite the contrary is the case. Namely, to test a metaphysical claim, its explication should create more content, testable consequences, than the background theory considered.

Moreover, the account of explication adopted here, allows, in principle, all metaphysical concepts to be explicated in the formal language of an empirical theory as long as the latter is rich enough in means of expression, i.e. as long as it has the adequate expressive capacity to represent the universe of discourse. Take, for instance, accounts in combinatorial ontology which makes a logic the basis of ontology (see Jacquette 2002:51f) and any empirical theory formulated in the language of that logic. Then, in principle, all celebrated metaphysical concepts, such as "object", "entity", "predication", "existence", "property" "possibility", may be explicated in the formal language of that empirical theory.

2. **Explication and Testability**

Consider an explication E of a metaphysical term $EX1(E)$ and a scientific theory T formulated in language $L_2(E)$, which is a background theory to E. Let ψ be a sentence of $L_2(E)$ that intends to express a metaphysical claim. The intent is to have an explication E that would consider ψ a condition of explicative adequacy, $\psi \in CEA(E)$, entailed by the explicative introduction $EI(E)$ together with the scientific theory T, while not entailed by T, if taken in isolation:

$$T \cup \{EI(E)\} \vdash_{L_2(E)} \psi \quad \text{(3a)}$$

$$\text{and } T \nvdash_{L_2(E)} \psi, \text{(3b)}$$

Notice that a more stringent condition than (3a) and (3b) would be to demand that E is a non-trivially successful explication for T.

To test a metaphysical claim ψ, one needs to infer a prediction ψ_{obs}, an observation statement, which can be verified or falsified by experiment.[4] The premises of this inference, apart from ψ, stem from the scientific theory T and they possibly include further auxiliary assumptions[5], hence,

$$T \cup \{\psi\} \vdash_{L_2(E)} \psi_{obs}. \quad \text{(4)}$$

Furthermore, to distribute blame or praise to ψ in terms of the outcome of the test for ψ_{obs}, one needs to show that ψ's contribution to the generation of the prediction is indispensable. This notion has been discussed in (Psillos 1999:110)[6] and it has been used by Hawley to "spot involvement" of metaphysical claims in generating empirical success (2006). In addition to the foregoing entailment, along Psillos' line of reasoning, one would require that T alone cannot yield ψ_{obs},

$$T \nvdash_{L_2(E)} \psi_{obs} \quad \text{(5)}$$

and there is no other hypothesis ψ' that together with T entails ψ_{obs}, which is consistent with T, and does not entail ψ, if taken together with T:

$$T \cup \{\psi'\} \nvdash_{L_2(E)} \psi, \quad \text{(6)}$$

---

[4] Here, I assume, tacitly, that there is a way to circumscribe the observation sentences in $L_2(E)$.

[5] In subsequent discussion, without loss of generality, I suppress reference to auxiliary assumptions.

[6] "Suppose", writes Psillos, "that H together with another set of hypotheses H' (and some auxiliaries A) entail a prediction P. H indispensably contributes to the generation of P if H' and A alone cannot yield P and no other available hypothesis H^* which is consistent with H' and A can replace H without loss in the relevant derivation of P." (Psillos 1999:110)

$$T \cup \{\psi'\} \nvdash_{L_2(E)} \bot, \quad (7)$$

$$T \cup \{\psi'\} \vdash_{L_2(E)} \psi_{obs}. \quad (8)$$

The indispensability of ψ for the generation of ψ_{obs} would be undermined if (8) were satisfied, unless the inference of ψ_{obs} were trivial, stemming from an inconsistent set of premises that entail anything, (7), or it was mediated by ψ, (6). In the latter case, ψ is still indispensable for the generation of ψ_{obs} as a minimum requirement, although the inference may use an alternative hypothesis as a premise. This condition, (6), deals effectively with the well-known problem of producing trivial alternative hypotheses from a given one, simply by "tacking on" it any other claim.[7]

What would be the source of such an alternative hypothesis, ψ'? One way to understand ψ' would be as expressing in the language $L_2(E)$ a *different* metaphysical claim, couched in terms of a different metaphysical concept, explicated in terms of another explication E'. A second way to understand ψ' is in terms of alternative explications of one and the same metaphysical expression. According to Cordes, E is an explication alternative to E' if and only if E and E' are explications, $E \neq E'$, $L_1(E) = L_1(E')$ and $EX1(E) = EX1(E')$ (2017: Df. 22). Thus, ψ' may be understood as expressing the *same* metaphysical claim in the language $L_2(E) = L_2(E')$, in terms of a *different* explication of the same concept.

Finally, notice that conditions (3a), (4), entail that

$$T \cup \{EI(E)\} \vdash_{L_2(E)} \psi_{obs}. \quad (9)$$

---

[7] Conditions (6) and (7) are not required by Psillos, although he discussed the problem of inferring the prediction from trivial hypotheses. He notes that "there are senses in which all theoretical assertions are eliminable, if, for instance we take the Craig-transform of a theory, or we "cook up" a hypothesis H^* [it would be ψ' in our discussion] by writing P [ψ_{obs}] into it. But if we impose some natural epistemic constraints on the potential replacement- if, for instance, we require that the replacement be independently motivated, non-ad hoc, potentially explanatory, etc.- then it is not certain at all that a suitable replacement can always be found" (1999:110)

According to what has been already discussed, $EI(E)$ is not a good candidate for being an alternative hypothesis ψ' for the generation of ψ_{obs}, since, given (3a) it violates (6). However, (9) together with (5) manifest the creative nature of the act of explication with respect to a given theory T since I believe that it is not far-fetched or risky to assume that the prediction ψ_{obs} does not contain the *explicatum EX2(E)*, as an atomic subexpression.

To sum up, I suggest that a metaphysical claim is testable if and only if there is an admissible test associated with that metaphysical claim. An admissible test for a metaphysical claim is a quadruple,

$$Test = \langle E, T, \psi, \psi_{obs} \rangle ,$$

that satisfies the following conditions:

a) E is an explication of a metaphysical concept or expression that occurs in that claim,

$$E = \langle L_1(E), L_2(E), EX1(E), EX2(E), CEA(E), EI(E) \rangle;$$

b) T is a background theory for E;

c) ψ is a formula in $L_2(E)$, corresponding to the metaphysical claim, such that,

$$\psi \in CEA(E),$$

$$T \cup \{EI(E)\} \vdash_{L_2(E)} \psi ,$$

$$\text{and } T \nvdash_{L_2(E)} \psi.$$

d) ψ_{obs} is a formula in $L_2(E)$, an observation sentence, for the generation of which ψ has an indispensable contribution, i.e.,

$$T \cup \{\psi\} \vdash_{L_2(E)} \psi_{obs} ,$$

and there is no other hypothesis ψ', such that

$$T \cup \{\psi'\} \nvdash_{L_2(E)} \psi,$$

$$T \cup \{\psi'\} \nvdash_{L_2(E)} \bot,$$

$$T \cup \{\psi'\} \vdash_{L_2(E)} \psi_{obs}.$$

3. An Illustration of the Method.

To illustrate this approach to deriving testable consequences of metaphysical claims explicated in the context of a background scientific theory, I am examining the metaphysical concept of common cause and the principle of the common cause. I will present three alternative explications of the concept and equinumerous formulations of the principle, while I will consider a Bell-type inequality as the observation sentence to be experimentally tested.

To begin with, consider an explication $E1$ of the metaphysical term,

$$EX1(E1) = \text{"common cause"},$$

which occurs in

$$L_1(E1) = \text{English},$$

in a language $L_2(E1)$ rich enough to accommodate the mathematical theory of classical probability spaces, $T,$ which is the background theory to E.

For a probability space $< X, \mathcal{L}, p >$, the *explicatum of $E1$*, is a function $Common1$[8],

$$EX2(E1) = Common1 \colon \mathcal{L} \times \mathcal{L} \to \mathcal{L} \ \text{such that} \ (A, B) \longmapsto C$$

where $C \in \mathcal{L}$ is the common cause of $A, B \in \mathcal{L}$.

---

[8] I would like to thank the anonymous referee for their suggestion to explicate the common cause in terms of a function.

The explicative introduction of the common cause, $EI(E1)$, consists in the conjunction of the two following sentences:

$$\forall A, B \in \mathcal{L}, \exists C \in \mathcal{L}, \; Common1(A, B) = C$$

if and only if

$$Corr(A, B) \wedge R_{ind}(A, B) \Longrightarrow (10a) \wedge (10b) \wedge (10c) \wedge (10d)$$

$$(10)$$

$$p(A \cap B | C) = p(A|C)p(B|C) \quad (10a)$$

$$p(A \cap B | C^\perp) = p(A|C^\perp)p(B|C^\perp) \quad (10b)$$

$$p(A|C) > p(A|C^\perp) \quad (10c)$$

$$p(B|C) > p(B|C^\perp) \quad (10d)$$

$$T \vdash_{L_2(E)} R_{ind}(A, B) \equiv L_{ind}(A, B) \qquad (11)$$

where $p(Y|Z)$ denotes the conditional probability of Y on condition Z, which is assumed well-defined, and $Y^\perp$ denotes the complement of a set $Y \in \mathcal{L}$. Moreover, $Corr(A, B) = "p(A|B) > p(A)"$, and, $R_{ind}(A, B)$ denotes a causal independence condition that should be satisfied by A, B in order for their common cause to exist. This condition would stipulate that neither A is the cause of B nor B the cause of A. (11) specifies that given the theory of probability spaces T, R_{ind} is equivalent to the logical independence L_{ind} relation: two events A, B are logically independent if and only if $A \cap B$; $A \cap B^\perp$; $A^\perp \cap B$; and $A^\perp \cap B^\perp$ are nonempty sets (Hofer-Szabó *et al.* 1999: 12). In order for the postulation of $Common1$ to be an adequate definition, Suppes would demand that the formula that follows the expression "if-and-only-if" in (10) is derivable from T (1957:158). However, if this were the case, the explicative introduction would not create novel content with respect to the background theory T. Yet, as we have already explained novel content is required for an explication of a metaphysical concept to yield testable consequences in combination with T.

Furthermore, $\psi_1 \in CEA(E1)$ may express in $L_2(E1)$ the following formulation of the principle of the common cause: "every correlation is either due to a direct logical dependence of the correlated events or to an indirect connection via a common cause of the correlated events."

$$\forall A, B \in \mathcal{L}, \exists C \in \mathcal{L}, \text{ such that } Corr(A, B) \implies \sim L_{ind}(A, B) \vee ((10a) \wedge (10b) \wedge (10c) \wedge (10d)). \quad (12)$$

Or, one may weaken the requested criterion of explicative adequacy for E by letting ψ_1' denote the following claim:

$$\forall A, B \in \mathcal{L}, \exists C \in \mathcal{L}, \text{ such that } Corr(A, B) \implies \sim L_{ind}^+(A, B) \vee ((10a) \wedge (10b) \wedge (10c) \wedge (10d)), \quad (12a)$$

where L_{ind}^+ is the logical independence modulo measure zero events relation: two events A, B are logically independent modulo measure zero events if and only if $p(A \cap B)$; $p(A \cap B^\perp)$; $p(A^\perp \cap B)$; and $p(A^\perp \cap B^\perp)$ are positive numbers. Since every pair of events that are logically independent modulo measure zero events are logically independent events, (12a) can be deduced from (10) and (11) together with the theory of probability spaces T. (Hofer-Szabó *et al.* 2013: 12)

Explication $E1$ of the common cause is just one of the eight possible explications suggested by Hofer-Szabó, Rédei and Szabó (2013:147), in terms of which one may provide alternative formulations of the principle of the common cause. To show how one can derive an observable consequence of the principle of the common cause, in line with (4), I will utilize $E2$, an explication alternative to $E1$, and the respective principle.

Thus, consider explication $E2$ of the metaphysical term,

$$EX1(E2) = \text{"common cause"},$$

which occurs in

$$L_1(E2) = \text{English},$$

in a language $L_2(E2)$ rich enough to accommodate the mathematical theory of classical probability spaces, T, which is the background theory to $E2$.

For a probability space $< X, \mathcal{L}, p >$, the *explicatum of E2, is*

$$EX2(E2) = Common2: \mathcal{L} \times \mathcal{L} \to \mathcal{L} \ such \ that \ (A, B) \mapsto C,$$

where $C \in \mathcal{L}$ is the common cause of $A, B \in \mathcal{L}$.

Analogously to $EI(E1)$, one may introduce the notion of a common cause by means of $EI(E2)$, defined as the conjunction of the following sentences:

$$\forall \mathcal{F} \subseteq \mathcal{L} \times \mathcal{L}, \ 1 \leq |\mathcal{F}| < \infty, \ \exists C \in \mathcal{L}, \ \forall (A, B) \in \mathcal{F}, \ Common2(A, B) = C$$

if and only if

$$Corr(A, B) \wedge R_{ind}(A, B) \Longrightarrow (10a) \wedge (10b) \wedge (10c) \wedge (10d)$$

$$(13)$$

$$T \vdash R_{ind}(A, B) \equiv L_{ind}(A, B). \ (14)$$

$EI(E2)$ postulates common causes for finite families $\mathcal{F}$ ($|\mathcal{F}| < \infty$), of positively correlated causally independent pairs of events and not only for pairs of events, as $EI(E1)$. The generalization is reflected on the reversal of the order of the universal and the existential quantifier as applied, respectively, to the correlated pair and the common cause in $EI(E1)$, and on the essential reference in $EI(E2)$ to a finite family of positively correlated events $\mathcal{F}$ which entails the association of a single common cause with that family.

In this case, the principle of the common cause, $\psi_2 \in CEA(E2)$, acquires the following form: "every correlation in any finite family of events is either due to a direct logical dependence of the correlated events or to an indirect connection via a common cause of the correlated events associated with that family",

$$\forall \mathcal{F} \subseteq \mathcal{L} \times \mathcal{L}, \ \exists C \in \mathcal{L}, \forall (A, B) \in \mathcal{F}, \text{ such that } Corr(A, B) \Longrightarrow \ \sim L_{ind}(A, B) \vee ((10a) \wedge (10b) \wedge (10c) \wedge$$

$$(10d)).$$

$$(15)$$

By applying this principle to a quadruple of positively correlated pairs of logically independent events (A_1, B_1), (A_1, B_2), (A_2, B_1), $(A_2, B_2) \in \mathcal{F}$ one may derive the following Bell-type inequality (Hofer-Szabó *et al.* 1999; Stergiou 2012):

$$-1 \leq p(A_1 \cap B_1) + p(A_1 \cap B_2) + p(A_2 \cap B_1) - p(A_2 \cap B_2) - p(A_1) - p(B_1) \leq 0 \quad (16)$$

Since 1972, Bell-type inequalities have been tested experimentally several times with the most recent test to have been performed in 2018 (see, Rauch *et al.* 2018) and the most notable one, by Alain Aspect and his team, in 1982. Hence, the derivation of these inequalities amounts to the derivation of observation statement ψ_{obs} from the theory of probability spaces, T, along with the principle of the common cause ψ_2, (15), as required by (4).

What remains to be examined is the indispensability of the principle of the common cause for the derivation of the Bell-type inequality, so as to characterize the metaphysical claim testable. Things, however, do not always follow our desires! Bell-type inequalities may be inferred from the alternative assumption of the existence of *a common cause system* for a quadruple of positively correlated pairs of events in $\mathcal{F}$.[9] A common cause system of size n is a generalization of the common cause for a finite family $\mathcal{F}$ of pairs of events determined in terms of a partition of the event space $\mathcal{L}$ in n parts, with $1 < n \leq \infty$. In what follows I refer to finite size common cause systems $(n < \infty)$ only and I consider them to be a third alternative account of common cause (Hofer-Szabó *et al.* 2013:147) introduced by explication $E3$:

$$EX1(E3) = \text{"common cause",}$$

which occurs in

$$L_1(E3) = \text{English,}$$

in a language $L_2(E3)$ rich enough to accommodate the mathematical theory of classical probability spaces, T, which is the background theory to E.

---

[9] A common cause system can be defined for a single pair of events (A, B). In this case, $\mathcal{F} = \{(A, B)\}$.

For a probability space $< X, \mathcal{L}, p >$, the *explicatum of E3*, is

$$EX2(E3) = Common3 \colon \mathcal{L} \times \mathcal{L} \to \overbrace{\mathcal{L} \times ... \times \mathcal{L}}^{n} \text{ such that } (A,B) \mapsto (C_1, ..., C_n)$$

where $\{C_1, ..., C_n\} \subseteq \mathcal{L}$ is the common cause of $A, B \in \mathcal{L}$.

The explicative introduction of the common cause, $EI(E3)$, consists in the conjunction of the following two sentences:

$$\forall \mathcal{F} \subseteq \mathcal{L} \times \mathcal{L}, \quad 1 \leq |\mathcal{F}| < \infty, \quad \exists n > 1, \quad \exists \{C_1, ..., C_n\} \subseteq \mathcal{L}, \quad \forall (A,B) \in \mathcal{F},$$

$$Common3(A,B) = (C_1, ..., C_n) \text{ if and only if}$$

$$Corr(A,B) \wedge R_{ind}(A,B) \Longrightarrow (17a) \wedge (17b) \wedge (17c) \qquad (17)$$

$$C_i \cap C_j = \emptyset, i,j = 1, ..., n, \ i \neq j \text{ and } \bigcup_{i=1}^{n} C_i = X \quad (17a)$$

$$p(A \cap B | C_i) = p(A|C_i)p(B|C_i), \ i = 1, ..., n \qquad (17b)$$

$$[p(A|C_i) - p(A|C_j)][p(B|C_i) - p(B|C_j)] > 0, \ i,j = 1, ..., n, \ i \neq j \qquad (17c)$$

$$T \vdash R_{ind}(A,B) \equiv L_{ind}(A,B). \quad (18)$$

In this case, the alternative metaphysical claim $\psi_3 \in CEA(E3)$ is the following: "every correlation in a finite family of events is either due to a direct logical dependence of the correlated events or to an indirect connection via a *collection of partial common causes* of the correlated events associated with that family."[10]

$$\forall \mathcal{F} \subseteq \mathcal{L} \times \mathcal{L}, \ 1 \leq |F| < \infty, \ \exists n > 1, \ \exists \{C_1, ..., C_n\} \subseteq \mathcal{L}, \ \forall (A,B) \in \mathcal{F}, \text{ such that}$$

---

[10] In (Hofer-Szabó *et al.* 2013:80) one may read: "Another natural idea is to suspect that the correlation between A and B is not due to a *single* factor but may be the cumulative result of a (possibly large) number of different 'partial common causes,' none of which can in and by itself yield a complete common-cause-type explanation of the correlation, all of which, taken together, can however account for the entire correlations".

$$Corr(A, B) \implies \sim L_{ind}(A, B) \vee \big((17a) \wedge (17b) \wedge (17c)\big).$$

Kitajima (2017) has shown (16), for every quadruple of positively correlated pairs of events in $\mathcal{F}$; thus, (8) is satisfied. Moreover, since a common cause system for a finite family of correlated pairs of events is not a common cause for that family, the existence of the former does not entail the existence of the latter; hence, (6). Thereby, ψ_3 qualifies as an alternative hypothesis for the derivation of Bell-type inequalities, ψ_{obs}; consequently, the principle of the common cause ψ is not indispensable for the derivation of ψ_{obs}, and Bell-type inequalities do not qualify as evidence for or against that metaphysical claim.

In this section, I considered E_1, E_2, E_3 to be alternative explications of the concept of common cause and ψ_1, ψ_2, ψ_3 to be alternative formulations of the principle of the common cause. This view, however, is not unobjectionable. The metaphysical concept explicated by E_2 is also called, in the literature, *common-common cause* of a family of positively correlated pairs of events and that introduced by E_3, a *common-common cause system*. This fact, however, would not change much of our discussion. Instead of considering E_1, E_2, E_3 alternative explications of a single concept, one should take them as explications of different concepts and ψ_1, ψ_2, ψ_3, respectively, as formulations of different hypotheses.[11]

---

[11] At this point I would like to comment on a concern raised by an anonymous referee about the metaphysical character of the concept of common cause and of the principle of common cause. The referee seems to consider common cause, in its various formal versions, a scientific concept and the respective formulations of the principle, physical claims. It is commonplace to treat the principle of common cause as a general metaphysical claim about the denial of pre-established harmony and in favor of causal action by contact with a probabilistic and a spacetime aspect. To examine whether this general claim is compatible with, validated or falsified by, a scientific theory, it should be given an exact formulation in the logico-mathematical framework of that theory by explicating the concept of common cause. Thus, emerged the different probabilistic explications I examined in this paper and not as the result of an elaboration on a scientific theory. Moreover, no major currently accepted physical theory uses the notion of common cause or the principle of common cause with possible sole exception the discussion of foundational issues. Yet, these are pragmatic grounds and the referee is right to raise the issue of an in principle justification of the distinction between an explicated metaphysical claim and any other, say, physical claim: if both claims are formulated in the logico-mathematical context of a physical theory and their consequences are, in principle, empirically testable, as I attempt to argue in this paper, what is the justification for their being considered differently?

4. Multiple Explications, Multiple Tests and Prospects of Verification.

In this section, I am going to delve into the idea of multiple tests associated with different explications of the same metaphysical concept. To begin with, let me define the class of admissible tests associated with a given metaphysical claim and a given theory T as the collection of quadruples,

$$\text{TEST} = \{Test_{ik} = \langle E_i, T, \psi_i, \psi_{obs\ ik} \rangle, i \in I, k \in K\},$$
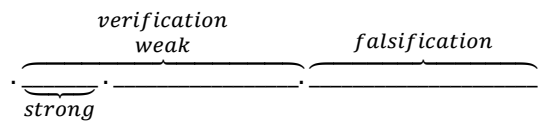
that satisfy the following condition:

a) For every $i \in I, k \in K, Test_{ik} = \langle E_i, T, \psi_i, \psi_{obs\ ik} \rangle$ is a test, according to the aforementioned conditions.

b) For every $i, j \in I, L_2(E_i) = L_2(E_j) = L(T)$, where $L(T)$ is the language of T.

c) For every $i, j \in I, \ i \neq j$, and $k, l \in K, Test_{ik} = \langle E_i, T, \psi_i, \psi_{obs\ ik} \rangle, Test_{jl} = \langle E_j, T, \psi_j, \psi_{obs\ jl} \rangle$, E_i, E_j are explication alternatives.

d) If $E_i = E_j$ then $\psi_i = \psi_j$.

Condition (a) and (b) demand that all elements of TEST are tests defined in terms of explications that share a common *explicatum* language, which is also the language of T. Condition (c) requires that all explications E_i, are explications of a single metaphysical concept originally expressed in the same language. The fourth condition, (d), stipulates that each test, in TEST, identifies one explication with one sentence in $L_2(E_i)$ which expresses the original metaphysical claim. As Cordes commented (personal communication), at this point one needs to be clear that, in general, no explication uniquely defines a statement *simpliciter*; however, the set TEST is constructed in such a way that each explication is identified with a single sentence.

In terms of a set of admissible tests, TEST, I suggest conditions for the verification and the falsification of a metaphysical claim. Namely, a metaphysical claim associated with a set of admissible tests, TEST, is said to be

- *weakly verified* if and only if, there is $i \in I$, such that for every $k \in K$, $\psi_{obs\,ik}$ is verified by experiment.

- *strongly verified* if and only if, for every $i \in I$, and for every $k \in K$, $\psi_{obs\,ik}$ is verified by experiment.

- *falsified* if and only if, for every $i \in I$, there is a $k \in K$, such that $\psi_{obs\,ik}$ is falsified by experiment.

Weak verification amounts to rendering a metaphysical claim more likely true, given the verification of the observation sentences $\psi_{obs\,ik}$, $k \in K$, while strong verification verifies the metaphysical claim conclusively for a given theory T. On the other hand, to falsify a metaphysical claim, one needs to find at least one falsification instance $\psi_{obs\,ik}$, *for every* explication E_i of the metaphysical concept. Weak verification is contradictory to falsification and a necessary, but not sufficient, condition for strong verification; the diagram below illustrates these logical relations:

The distinction between strong and weak verification alludes to Ayer's distinction which purported to introduce a relaxation of the verifiability principle (1936). Nevertheless, there is an important difference between the two pairs of terms employed: for Ayer strong verification is an asymptotic condition, satisfied only if an infinite number of verification instances for a hypothesis is obtained, while each verification instance weakly verifies that hypothesis. In our discussion, weak verification as well may be an asymptotic condition since it may require the

examination of an infinite number of observation sentences $\psi_{obs\ ik}, k \in K$, associated with a single explication E_i, for some $i \in I$. In particular, the set of indices K, which serves to particularize the tests for a given explication E_i, may have infinite cardinality, since for a given explication E_i and a corresponding ψ_i one may infer in terms of the same scientific theory T *different* predictions referring, say, to local tests taking place at different spacetime points. Thus, weak verification of a metaphysical claim may be considered similar to Ayer's strong verification of a scientific hypothesis. Moreover, strong verification of a metaphysical claim builds upon weak verification demanding further the metaphysical claim to be weakly verified for every explication E_i, $i \in I$.[12]

Last but not least, Popper's (1935:19) celebrated *asymmetry* between verification and falsification of scientific hypotheses seems not to be correct, in general, with regard to metaphysical claims. It can be traced *on the condition* that an infinite number of observation instances is required for the weak verification of a metaphysical claim while a finite number of observation instances is sufficient for its falsification. However, this condition is not trivially satisfied (see note 11).

5. Discussion

The first point to notice is that testability of a metaphysical claim is relativized to a scientific theory. This is expressed in the definition of weak (strong) verifiability and falsifiability of a metaphysical claim *in association with* a set of admissible tests, TEST, defined relative to a single theory T. For a set of admissible tests, TEST1, defined in terms of a different theory T_1, the same

---

[12] In an earlier version of this paper, I stipulated the set of indices I that counts different explications to be a finite set. The rationale for this condition was the following: the act of explicating a metaphysical term or concept is a work in progress and the number of available explications may become all the greater. No matter, however, how big the number may be, it remains always finite, at any given time. However, Ruetsche (personal communication) observed that one may consider metaphysical concepts whose explications could conceivably associate them with points of space; as an example, she suggested the concept "the navel of the earth". In these cases, explications are as many as points of space, hence, infinite. In view of her counterexample I dropped this requirement: available explications of a metaphysical concept may be finitely or infinitely many.

explication E may not yield a statement that indispensably contributes to the generation of an observation sentence; hence the metaphysical claim may not be testable. Things can get worse if there is no explication alternative E' that does the job and the metaphysical claim may not be testable in terms of a scientific theory, although it is testable in terms of another. Finally, one cannot exclude *a priori* the worst-case scenario, of a metaphysical claim being verified with respect to a theory T and falsified with respect to another theory T_1.

The second remark relates to philosophical practice. Philosophers explicate metaphysical concepts to render them precise and to avoid ambiguity. For this reason, the *explicatum* language is taken to be more precise than the *explicandum* language. A formal language, the language of first or second order logic, would certainly be a good example of such a more precise language, however, there is no need to think of our scientific theories as formalized exclusively in these languages to have a good *explicatum* language. The mathematical apparatus of a scientific theory, be that the theory of groups, the theory of probability spaces or the theory of topological algebras, would certainly work quite well as *explicatum* language and although these theories can be formulated in the language of first or second order logic I find no reason to restrict my attention to an uninterpreted linguistic system.

My third remark also concerns the philosophical practice of testing a metaphysical claim. Whenever a philosopher explicates a metaphysical concept and formulates a claim in the *explicatum* language, the aim is to examine, firstly, its compatibility with the axioms, or, more generally, with well-established statements of the scientific theory; secondly, to examine the testability of this claim, i.e. the existence of an admissible test associated with the metaphysical claim and defined in terms of the explication under examination; and thirdly, to see whether the claim is verified (weakly) or not, for a subset of admissible tests defined in terms of that same explication.

Let us take for granted the compatibility with the axioms requirement, which constitutes a rather obvious demand, and focus on the testability of the metaphysical claim requirement. Assume, as previously, that the philosopher works in the framework of a given explication E_i. Is

it correct to say that they face the following dilemma: either $\psi_i \in CEA(E_i)$ induces an extension, in terms of $\psi_{obs\,ik}, k \in K$, of the observational content of T, or the metaphysical claim is devoid of factual content and can be deemed honorific? Obviously not! This is a case of a false dilemma, since a failure in finding a number of tests associated with a given explication, $Test_{ik} = \langle E_i, T, \psi_i, \psi_{obs\,ik} \rangle, k \in K$, does not decide for the testability of the metaphysical claim *per se*. It only, forces the philosopher to consider a different explication E_j. Thus, a testability failure is ensued by devising a new explication. In a similar vein, assume that the philosopher working in the context of a given explication finds the metaphysical claim testable and proceeds with testing it. At this stage their dilemma is different: either the metaphysical claim is weakly verified, or it is falsified. A failure to verify the metaphysical claim leads to devising a new explication. The moral of the foregoing discussion is that if philosophers fail in anyone of the three goals set upon explicating a metaphysical concept, they suggest a different explication and they start all over again. This is how explications proliferate and explication alternatives are produced.[13]

To illustrate the practice of producing novel explications of a concept upon reaching a falsifying consequence, I consider anew the violation of Bell-type inequalities by experiment and the concept of common cause. In an attempt to explain the existence of superluminal correlations between spatially distant quantum systems, philosophers sought to employ local causal principles, such as the principle of the common cause, and to infer the existence of an underlying common cause which would be responsible for the observed correlations. However, as Butterfield (1989) and van Fraassen (1989) have shown, the evocation of the principle of the common cause would entail the Bell-type inequalities that were experimentally falsified. In view of this fact, Hofer-Szabó, Rédei and Szabó (1999) re-explicated the concept of common cause so as to avoid the falsified observable consequence and to leave open the possibility of the

---

[13] Notice that in the process of revising an explication so as to produce an alternative, the revised explication is not in general independent from the original one. This practice has been conceptually grasped by Cordes' definition of a more demanding explication: for all explication alternatives E, E^*, E is a more demanding explication than E^* if and only if for all $\psi \in CEA(E^*)$ it holds that $CEA(E) \vdash_{L_2(E)} \psi$; and there is a $\psi \in CEA(E)$ such that $CEA(E^*) \nvdash_{L_2(E)} \psi$ (2017: Df.30).

common cause explanation of the superluminal correlation. They distinguished between two conceptions of the common cause: the Reichenbachian common cause, as presented by explication $E1$ in section 3, and the common-common cause as introduced by explication $E2$ in that same section. Only in the second understanding of the common cause one may infer the falsified Bell-type inequalities. Hence, by interpreting the common cause principle in terms of the first explication provided, ($E1$), as a Reichenbachian common cause, one renders it safe from the violation of the Bell-type inequalities and compatible with the existence of superluminal correlations. "In short", says Rédei, "under the present specification of Bell's inequality and of the concept of Reichenbachian common cause, it is impossible to give meaning to the claim "Bell's inequality is implied by Reichenbach's common cause principle"; hence, on the present interpretation, violation of Bell's inequality does not imply the impossibility of Reichenbachian common causes of superluminal correlations" (1998:224).

What is the aim of multiplying explications for a given metaphysical concept? Is it to falsify the metaphysical claim examined or to evade falsification and safeguard the metaphysical claim from its potential falsifiers? Upon yielding a falsifying instance for a given explication, does the philosopher produce a new explication to falsify conclusively the metaphysical claim in the manner I suggested in section 4 or do they propose an alternative explication to immunize the claim against that falsifying instance? Hofer-Szabó, Rédei and Szabó claimed that "[t]he history of philosophy teaches us that metaphysical claims of sweeping generality are neither verifiable nor conclusively falsifiable. One can only aim at assessing their plausibility on the basis of the best available evidence provided by the sciences - both formal and empirical sciences." (2013: 173). Since they do not believe that verification or conclusive falsification of a metaphysical claim is possible, the best they can do is to examine its conformity with the best *available* evidence. To improve compatibility and avoid contradictions they suggest new explications of the metaphysical concepts so as to immunize the respective metaphysical claims against falsifying instances. On the contrary, I believe that conclusive falsification or verification of a metaphysical claim is an attainable goal, even if it is to be accomplished asymptotically, in the limit of empirical research. And the history of philosophy can provide at most a pessimistic

inductive argument against this view; however, something being unlikely in philosophical realms does not mean that it is false. Thus, it is not futile to keep on looking for observation sentences that are verifying instances of a metaphysical claim associated with an explication or to multiply explications so as to attain conclusive falsification in case one bumps into a falsifying instance of that claim for a given explication.

6. Summary

Metaphysical claims are testable! This is the central point of this paper. Testability, however, rests on two assumptions: a) the claim to be tested has been expressed in the logico-mathematical language of a scientific theory; b) the claim plays an indispensable role in producing new empirical content along with the theory. Cordes' account of explication has been used to make condition (a) feasible. Elaborating on an idea of Psillos on the indispensable contribution of a hypothesis to the production of novel predictions led to the fulfillment of condition (b). One important feature of my account is that it makes testability of a metaphysical claim not a property of the claim *per se* but a property which is relative to a language and a scientific theory. Further, I illustrated my account of testing a metaphysical claim by means of the concept of common cause, the principle of common cause and Bell-type inequalities. I contended that testable metaphysical claims can, in principle, be verified and/or falsified, even if both verifiability and falsifiability may be attained in the limit of empirical research. Both verifiability and falsifiability of a claim depend on explication and on empirical evidence and since they apply to testable claims, they are also theory-dependent properties. Finally, two different attitudes towards metaphysical claims that meet falsifying empirical evidence were discussed: one that tends to safeguard metaphysical claims by re-explicating them, fueled by the idea that metaphysical claims are unverifiable and unfalsifiable; and, the one I suggest which tends to produce novel explications of the metaphysical claims with the aim of obtaining conclusive falsification.

References

1. Ayer, A.J., (1936 [1971]). *Language, Truth, and Logic.* London: Penguin.

2. Beaney, M., (2004). "Carnap's conception of Explication: From Frege to Husserl?". In Awodey, S. and Klein, C., (eds). *Carnap brought Home: The view from Jena.* Chicago and La Salle-Illinois: Open Court: 117-150.

3. Butterfield, J., (1989). "A space-time approach to the Bell Inequality." In Cushing, J.T, and McMullin, E., (eds.), *Philosophical Consequences of Quantum Theory: Reflections on Bell's Theorem.* Notre Dame: University of Notre Dame Press: 114-144.

4. Carnap, R., (1934 [1937]). *The Logical Syntax of Language*. London: Routledge and Kegan Paul.

5. Carnap, R., (1950). *Logical Foundations of Probability*. The University of Chicago Press-Routledge & Kegan Paul.

6. Carnap, R., (1963). "Replies and Systematic Expositions". In Schilpp, P.A., (ed.). *The Philosophy of Rudolf Carnap.* La Salle-Illinois: Open Court: 859-1016.

7. Cordes, M., (2017). "The constituents of an explication" *Synthese* https://doi.org/10.1007/s11229-017-1615-5

8. Cordes, M. and Siegwart, G., "Explication" (2019). *The Internet Encyclopedia of Philosophy*, ISSN 2161-0002, https://www.iep.utm.edu/explicat/

9. French, S., (2017). *The Structure of the World: Metaphysics and Representation.* Oxford University Press.

10. Hanna, J.F., (1968). "An Explication of 'Explication'" *Philosophy of Science* 35:28-44.

11. Hawley, K., (2006). "Science as a Guide to Metaphysics?" *Synthese* 149:451–470.

12. Hofer-Szabó, G., Rédei, M., and Szabó, L.E., (1999). "On Reichenbach's common cause principle and Reichenbach's notion of common cause." *The British Journal for the Philosophy of Science* 50(3), 377– 399.

13. Hofer-Szabó, G., Rédei, M., and Szabó, L.E., (2013). *The Principle of the Common Cause.* Cambridge: Cambridge University Press.

14. Jacquette, D., (2002). *Ontology.* McGill-Queen's University Press.

15. Kitajima, Y., (2017 v.2). "Reichenbachian common cause systems of size 3 in general probability theories." https://arxiv.org/pdf/1704.01227.pdf

16. Popper, K., (1935 [2002]). *The Logic of Scientific Discovery.* London and New York: Routledge.

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

18. Rauch, D., *et al.,* (2018). "Cosmic Bell Test Using Random Measurement Settings from High-Redshift Quasars." *Physical Review Letters*, 121(8).

19. Rédei, M., (1998). *Quantum Logic in Algebraic Approach.* Dordrecht: Kluwer Academic Publishers.

20. Redhead. M., (1987). *Incompleteness, nonlocality, and realism: a prolegomenon to the philosophy of quantum mechanics.* Oxford University Press.

21. Shimony, A., (1984). "Contextual Hidden Variables Theories and Bell's Inequalities". *British 11Journal for the Philosophy of Science* 35: 25-45.

22. Stergiou, C. (2012)." Two comments on the common cause principle in algebraic quantum field theory." In *EPSA Philosophy of Science: Amsterdam 2009*, pp. 387–402. Springer.

23. Suppes, P., (1957). *Introduction to Logic.* Mineola-New York: Dover.

24. van Fraassen, B.C., (1989). "The Charybdis of Realism: Epistemological Implications of Bell's Inequality." In Cushing, J.T, and McMullin, E., (eds.), *Philosophical Consequences of Quantum Theory: Reflections on Bell's Theorem.* Notre Dame: University of Notre Dame Press: 97-113.