dialectica

International Journal of Philosophy

Contents

EUROPEAN SOCIETY FOR ANALYTIC PHILOSOPHY

Vol 76 : 1
March 2022

dialectica

International Journal of Philosophy
Official Organ of the European Society of Analytic Philosophy

founded in 1947 by Gaston Bachelard, Paul Bernays and Ferdinand Gonseth dialectica

vol 76
issue 1

March 2022

Contents

Actual Causation

Holger Andreas & Mario Günther

We put forth an analysis of actual causation. The analysis centers on the notion of a causal model that provides only partial information as to which events occur. The basic idea is this: *c* causes *e* only if there is a causal model that is uninformative on *e* and in which *e* will occur if *c* does. We show that our analysis captures more causal scenarios than any account that tests for counterfactual dependence under certain contingencies.

We analyse causation between token events. Here is the gist of the analysis: an event *c* is a cause of another event *e* only if both events occur, and—after taking out the information whether or not *e* occurs—*e* will occur if *c* does. We will show that the analysis successfully captures a wide range of causal scenarios, including overdetermination, preemption, switches, and scenarios of double prevention. This set of scenarios troubles counterfactual accounts of actual causation. Even sophisticated counterfactual accounts still fail to deal with all of its members. And they fail for a principled reason: to solve overdetermination and preemption, they rely on a strategy which gives the wrong results for switches and a scenario of double prevention. Our analysis, by contrast, is not susceptible to this principled problem.

Counterfactual accounts try to analyse actual causation in terms of counterfactual dependence. An event *e* counterfactually depends on an event *c* if and only if (iff), were *c* not to occur, *e* would not occur. Among the accounts in the tradition of Lewis (1973), counterfactual dependence between two occurring events is taken to be sufficient for causation.[1] That is, an occurring event *c* is a cause of a distinct occurring event *e* if, were *c* not to occur, *e* would not occur. Counterfactual accounts thus ask, "What would happen if the putative cause were absent?" Under this counterfactual assumption, they claim causation if the presumed effect is absent as well.

1 See Lewis (1973, 2000), Ramachandran (1997), Hitchcock (2001), Yablo (2002), Woodward (2003), Hall (2004, 2007), Halpern and Pearl (2005), Halpern (2015), and many others.

Overdetermination is troublesome for counterfactual accounts. Consider the scenario depicted in figure 1.

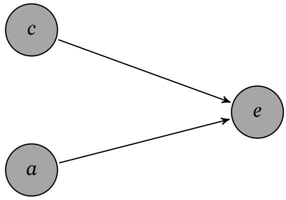


Figure 1: CAPTION NEEDED

Neuron c and neuron a fire. The firing of each of c and a alone suffices to excite neuron e. Hence, the common firing of c and a overdetermines e to fire. Arguably, the firing of c is a cause of e's excitation, and so is the firing of a.

What would have happened had c not fired? If c had not fired, e would have been excited anyways. After all, a would still have fired. Hence, as is well known, c is not a cause of e on Lewis's (1973) account. More sophisticated accounts solve the scenario of overdetermination as follows: c's excitation is a cause of e's firing because e's firing counterfactually depends on c's excitation if a were not to fire. The non-actual contingency that a does not fire reveals a hidden counterfactual dependence of the effect e on its cause c. The general strategy is to test for counterfactual dependence under certain contingencies, be they actual or non-actual. We call counterfactual accounts relying on this strategy 'sophisticated'.[2]

Numerous sophisticated accounts analyse causation relative to a causal model. A causal model represents a causal scenario by specifying which events occur and how certain events depend on others. Formally, a causal model $\langle M, V \rangle$ is given by a variable assignment V and a set M of structural equations. For the above scenario of overdetermination, V may be given by the set $\{c, a, e\}$, which says that all neurons fire. M is given by $\{e = c \vee a\}$, which says that e fires iff c or a does. In this causal model, we may set the variable c to $\neg c$, a to $\neg a$ and propagate forward the changes effected by these interventions. Given that $\neg c$ and $\neg a$, the structural equation determines that $\neg e$. The equation tells us that e would not have fired, if c had not fired under the contingency that a had

2 Sophisticated counterfactual accounts are, for example, provided by Ramachandran (1997), Hitchcock (2001), Yablo (2002), Woodward (2003, chap. 2.7), Halpern and Pearl (2005), Hall (2007), and Halpern (2015).

not fired. Hence, the above solution of overdetermination can be adopted: c is a cause of e (relative to the causal model) because e counterfactually depends on c if $\neg a$ is set by intervention.[3]

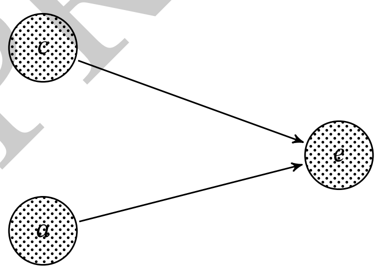
We solve the problem of overdetermination in a different way. The idea is this: remove enough information about which events occur so that there is no information on whether or not a putative effect occurs; an event c is then a cause of this effect only if—after the removal of information—the effect will occur if c does.

We use causal models to implement the idea. The result of the information removal is given by a causal model $\langle M, V' \rangle$ that provides only partial information as to which events occur, but complete information about the dependences between the events. To outline the preliminary analysis: c is a cause of e relative to a causal model $\langle M, V \rangle$ iff

1. c and e are true in $\langle M, V \rangle$, and
2. there is $V' \subset V$ such that $\langle M, V' \rangle$ contains no information as to whether e is true, but in which e will become true if c does.

By these conditions, we test whether an event brings about another event in a causal scenario. Causation is here actual production.

Why is c's excitation a cause of e's firing in the overdetermination scenario? Take the causal model $\langle M, V' \rangle$ that contains no information about whether or not the effect e occurs:

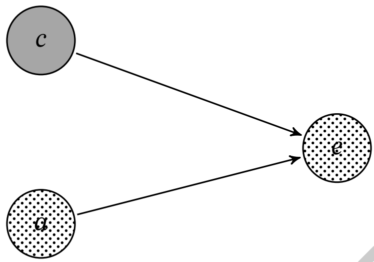


Here, a neuron is dotted iff V' contains no information as to whether the neuron fires or not. Since all neurons are dotted, the causal model contains no information on which neurons fire. But it still contains all the information about dependences among the neurons, as encoded by the structural equation

3 Sophisticated accounts that rely on causal models are, for example, provided by Hitchcock (2001), Woodward (2003, chap. 2.7), Halpern and Pearl (2005), Hall (2007), and Halpern (2015).

doi: 10.48106/dial.v76.i1

of the overdetermination scenario. Let us now intervene such that *c* becomes excited:



The structural equation is triggered and determines *e* to fire. Hence, *c*'s excitation is a cause of *e*'s firing on our analysis. The overdetermination scenario is solved without counterfactually assuming the absence of the cause and without invoking any contingency.

It should be noted that the recent counterfactual theories of Gallow (2021) and Andreas and Günther (2021a) are not sophisticated in our sense: they do not test for counterfactual dependence under certain contingencies. And so they are not susceptible to the principled problem. Indeed, both theories solve the set of scenarios that troubles sophisticated accounts. The analysis of Andreas and Günther (2021a) relies on a removal of information just like the analysis proposed here, and can thus be seen as its counterfactual counterpart. We will briefly and favourably compare our analysis to its counterfactual counterpart in the Conclusion.

In what follows, we refine our analysis, apply it to causal scenarios, and compare it to counterfactual accounts. In section 1, we introduce our account of causal models. In section 2, we state a preliminary version of our analysis and explain its rationale. We apply this analysis to various causal scenarios in section 3. In response to certain switching scenarios, we amend our preliminary analysis by a condition of weak difference making. In section 4, we state the final version of our analysis. In section 5, we compare our analysis to the extant counterfactual accounts. Section 6 concludes the paper.

1 Causal Models

In this section, we explain the basic concepts of causal models. Our account parallels the account of causal models in Halpern (2000). Unlike Halpern, we introduce structural equations as formulas and not as functions. Another difference is that our account is confined to binary variables, the values of which are represented by literals.[4] We will see shortly that these modelling choices allow us to define causal models in a straightforward way, in particular causal models that carry only partial information as to which events occur. In the appendix, we supplement the explanations of the core concepts of causal models with precise definitions.

Our causal models have two components: a set M of structural equations and a consistent set V of literals. Where p is a propositional variable, p is a positive literal and $\neg p$ a negative literal. We give literals a semantic role. The literals in V denote which events occur and which do not, that is, which events and absences are actual. $p \in V$ means that the event corresponding to p occurs. $\neg p \in V$, by contrast, means that no token event p of the relevant type occurs. Since the set of literals is consistent, it cannot be that both p and $\neg p$ are in V. Arguably, an event cannot both occur and not occur at the same time.

A structural equation denotes whether an event would occur if some other events were or were not to occur. Where p is a propositional variable and ϕ a propositional formula, we say that

$$p = \phi$$

is a structural equation. Each logical symbol of ϕ is either a negation, a disjunction, or a conjunction. ϕ can be seen as a truth function whose arguments represent occurrences and non-occurrences of events. The truth value of ϕ determines whether p or $\neg p$.

Consider the scenario of overdetermination depicted in figure 1. There are arrows from the neurons c and a to the neuron e. The arrows represent that the propositional variable e is determined by the propositional variables c and a. The specific structural equation of the overdetermination scenario is $e = c \vee a$. This equation says that e occurs iff c or a does. A set of structural equations describes dependences between actual and possible token events.

For readability, we will represent causal models in two-layered boxes. The causal model of the overdetermination scenario, for example, is given by $\langle\{e = c \vee a\}, \{c, a, e\}\rangle$. We will depict such causal models $\langle M, V\rangle$ in a box, where the upper layer shows the set M of structural equations and the lower layer the set V of actual literals. For the overdetermination scenario, we obtain:

4 With a few modifications, both the framework and the analysis can be extended to non-binary variables.

doi: 10.48106/dial.v76.i1

$e = c \vee a$
c, a, e

We say that a set V of literals satisfies a structural equation $p = \phi$ just in case both sides of the equation have the same truth value when plugging in the literals in V. In the case of overdetermination, the actual set of literals satisfies the structural equation. By contrast, the set of literals $\{c, a, \neg e\}$ does not satisfy $e = c \vee a$. When plugging in the literals, the truth values of e and $c \vee a$ do not match. We say that a set V of literals satisfies a set M iff V satisfies each member of M.

The structural equations and the literals determine which events occur and which do not occur in a causal model. This determination can be expressed by a relation of satisfaction between a causal model and a propositional formula.

Definition 1 ($\langle M, V\rangle$ satisfies ϕ). $\langle M, V\rangle$ satisfies ϕ iff ϕ is true in all complete sets V^c of literals that extend V and satisfy M. A set V^c of literals is complete iff each propositional variable (in the language of M) is assigned to a truth value by V^c.

If V is complete, this definition boils down to: $\langle M, V\rangle$ satisfies ϕ iff V satisfies ϕ, or V does not satisfy M. Provided V is complete, $\langle M, V\rangle$ satisfies at least one of ϕ and $\neg\phi$ for any formula ϕ.

Our analysis relies on causal models that contain no information as to whether or not an effect occurs. We say that a causal model $\langle M, V\rangle$ is *uninformative* about a formula ϕ iff $\langle M, V\rangle$ satisfies none of ϕ and $\neg\phi$. Note that $\langle M, V\rangle$ cannot be uninformative on any formula if V is complete.

In the scenario of overdetermination, the causal model $\langle M, V\rangle$ is uninformative on e for $V = \emptyset$. There are four complete extensions that satisfy $M = \{e = c \vee a\}$. One of these is $\{\neg c, \neg a, \neg e\}$. Hence, $\langle M, V\rangle$ does not satisfy e. Similarly, $\langle M, V\rangle$ does not satisfy $\neg e$. There is a complete extension of V that satisfies M but fails to satisfy $\neg e$. The actual set $\{c, a, e\}$ of literals, for example, but also the sets $\{c, \neg a, e\}$ and $\{\neg c, a, e\}$. The structural equation constrains the overdetermination scenario to four possible cases. These cases are expressed by the complete sets of literals which satisfy M.

Why is $\langle M, V\rangle$ not uninformative on e for $V = \{a\}$? Well, there is no complete extension of V that satisfies the structural equation in M but fails to satisfy e. There are only two such complete extensions: $\{c, a, e\}$ and $\{\neg c, a, e\}$. If a remains in the set V of literals, e is determined independent of whether or not c occurs.

It remains to introduce interventions. Recall that a structural equation $p = \phi$ determines the truth value of the variable p if certain variables q occurring in ϕ are given truth values by the literals in V. To represent an intervention that sets p to one of the truth values, we replace the equation $p = \phi$ by the corresponding literal p or $\neg p$. We implement such interventions by the notion of a submodel. M_I is a submodel of M relative to a consistent set I of literals just in case M_I contains the literals in I and the structural equations of M for the variables which do not occur in I. In symbols,

$$M_I = \{(p = \phi) \in M \mid p \notin I \text{ and } \neg p \notin I\} \cup I.$$

We denote interventions by an operator $[\cdot]$ that takes a model M and a consistent set of literals I, and returns a submodel. In symbols, $M[I] = M_I$. In the overdetermination scenario, for instance, we may intervene on $M = \{e = c \lor a\}$ by $\{\neg a\}$. This yields: $M[\{\neg a\}] = \{\neg a, e = c \lor a\}$. The causal model $\langle M_{\{\neg a\}}, \varnothing\rangle$ satisfies $\neg a$, and $\langle M_{\{\neg a\}}[\{\neg c\}], \varnothing\rangle$ satisfies $\neg e$. If $\neg c$ were actual under the contingency that $\neg a$, $\neg e$ would be actual.

Finally, note that the above definition of satisfaction applies to causal models and causal submodels. The definition does not only capture the relation of a causal model $\langle M, V\rangle$ satisfying a formula ϕ, but also the relation of a causal submodel $\langle M_I, V\rangle$ satisfying such a formula. This is explained further in the appendix.

2 The Analysis

We are now in a position to spell out our analysis in a more precise way. The key idea is as follows: for c to be a cause of e, there must be a causal model $\langle M, V'\rangle$ that is uninformative about e, while intervening by c determines e to be true. The latter condition must be preserved under all interventions by a set A of actual events. In more formal terms:

Definition 2 (Actual Cause, preliminary). Let $\langle M, V\rangle$ be a causal model such that V satisfies M. c is an actual cause of e relative to $\langle M, V\rangle$ iff

(C1) $\langle M, V\rangle$ satisfies c and e, and
(C2) there is $V' \subset V$ such that $\langle M, V'\rangle$ is uninformative on e, while for all $A \subseteq V$, $\langle M_A[\{c\}], V'\rangle$ satisfies e.

The rationale behind our analysis is straightforward: there must be a way in which a genuine cause actually brings about its effect. This production of the

doi: 10.48106/dial.v76.i1

effect can be reconstructed by means of a causal model $\langle M, V' \rangle$ that contains some information of the original causal model $\langle M, V \rangle$, but no information about whether the effect is actual. Or so requires condition (C2).

Furthermore, (C2) says production of an effect must respect actuality. The idea is that the causal process initiated by a genuine cause must respect what actually happened. A genuine cause cannot produce its effect via non-actual events and absences. The process from cause to effect must come about as it actually happened. This idea requires that a genuine cause must bring about its effect by events and absences that are actual. We implemented this requirement as follows: intervening upon the uninformative model $\langle M, V' \rangle$ by any subset of the actual events and absences V must preserve that e will become actual if c does. Thereby, it is ensured that a genuine cause cannot bring about its effect by events or absences that are not actual. If c is a genuine cause, there can be no subset A of the actual literals V that interferes with the determination of e by c in the respective uninformative model. We describe this feature of (C2) as *intervention by actuality*.

3 Scenarios

In this section, we test our analysis of actual causation against causal scenarios, and compare the results to the counterfactual accounts due to Lewis (1973), Hitchcock (2001), Halpern and Pearl (2005), and Halpern (2015). We follow Paul and Hall (2013, 10) in laying out the structure of causal scenarios by neuron diagrams. "Neuron diagrams earn their keep," they write, "by representing a complex situation clearly and forcefully, allowing the reader to take in at a glance its central causal characteristics."[5] We introduce simple neuron diagrams for which there is always a corresponding causal model. Our causal models, however, can capture more causal scenarios than simple neuron diagrams.

A neuron diagram is a graph-like representation that comes with different types of arrows and different types of nodes. Any node stands for a neuron, which fires or else does not. The firing of a neuron is visualized by a gray-shaded node, the non-firing by a white node. For the scenarios to be considered, we need two types of arrows. Each arrow with a head represents a stimulatory connection between two neurons, each arrow ending with a black

5 This being quoted, there are some shortcomings of neuron diagrams. For details, see Hitchcock (2007b).

dot an inhibitory connection. Furthermore, we distinguish between *normal* neurons that become excited if stimulated by another and *stubborn* neurons whose excitation requires two stimulations. Normal neurons are visualized by circles, stubborn neurons by thicker circles. A neuron diagram obeys four rules. First, the temporal order of events is left to right. Second, a normal neuron will fire if it is stimulated by at least one and inhibited by none. Third, a stubborn neuron will fire if it is stimulated by at least two and inhibited by none. Fourth, a neuron will not fire if it is inhibited by at least one.

Typically, neuron diagrams are used to represent events and absences. The firing of a neuron indicates the occurrence of some event and the non-firing indicates its non-occurrence. Recall that we analyse causation between token events relative to a causal model $\langle M, V \rangle$, where the causal model represents the causal scenario under consideration. We thus need a correspondence between neuron diagrams and causal models.

Here is a recipe to translate an arbitrary neuron diagram, as detailed here, into a causal model. Given a neuron diagram, the corresponding causal model can be constructed in a step-wise fashion. For each neuron n of the neuron diagram:

i. Assign n a propositional variable p.
ii. If n fires, add the positive literal p to the set V of literals.
iii. If n does not fire, add the negative literal $\neg p$ to V.
iv. If n has an incoming arrow, write on the right-hand side of p's structural equation a propositional formula ϕ such that ϕ is true iff n fires.[6]

This recipe adds a positive literal p to the set V of literals for each neuron that fires, and a negative literal $\neg p$ for each neuron that dos not fire. Then the neuron rules are translated into structural equations. One can thus read off a neuron diagram its corresponding causal model: if a neuron is shaded gray,

6 The structural equations can be explicitly constructed from the rules governing neuron diagrams. That is, the catch-all condition (iv) can be replaced by the following clauses. (v) For each stimulatory arrow ending in a normal neuron n, add disjunctively to the right side of p's structural equation the variable that corresponds to the neuron where the arrow originates. (vi) For each pair of stimulatory arrows ending in a stubborn neuron n, add disjunctively to the right side of p's structural equation the conjunction of the two variables that correspond to the two neurons where the arrows originate. (vii) For each inhibitory arrow ending in n, add conjunctively to the right side of p's structural equation the negation of the variable that corresponds to the neuron where the arrow originates. This translation shows that there is a principled transition from simple neuron diagrams to our causal models.

doi: 10.48106/dial.v76.i1

p is in the set V of literals of the corresponding causal model; if a neuron is white, $\neg p$ is in V.

We have already added a feature to neuron diagrams in the introduction. Recall that dotted nodes represent neurons about which there is no information as to whether or not they fire. In more formal terms, if $p \notin V$ and $\neg p \notin V$, the corresponding neuron will be dotted. We portray now how our analysis solves the problems posed by overdetermination, conjunctive causes, early and late preemption, switches, prevention, and two scenarios of double prevention.

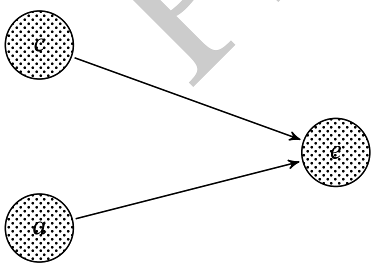
3.1 *Overdetermination*

Scenarios of overdetermination are commonly represented by the neuron diagram depicted in figure 1. Here is a story that fits the structure of overdetermination: A prisoner is shot by two soldiers at the same time (c and a), and each of the bullets is fatal without any temporal precedence. Arguably, both shots should qualify as causes of the death of the prisoner (e).

Our recipe translates the neuron diagram of figure 1 into the following causal model $\langle M, V\rangle$:

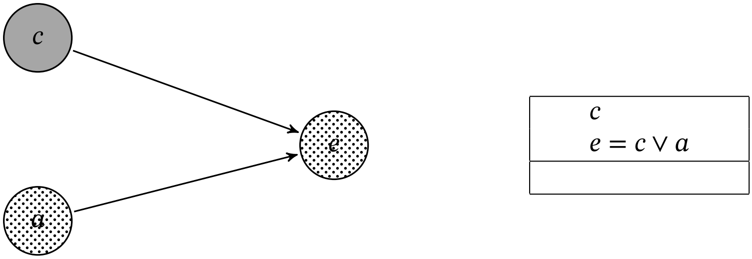
$e = c \vee a$
c, a, e

Relative to $\langle M, V\rangle$, c is a cause of e. For this to be seen, consider the following causal model $\langle M, V'\rangle$ that is uninformative on e.



$e = c \vee a$

Intervening by $\{c\}$ yields:



Obviously, this causal model determines e to be true. In more formal terms, $\langle M_{\{c\}}, V' \rangle$ satisfies e. And intervening by any subset of actual events does not undo the determination.[7] In more detail, any intervention by a subset of $\{c, a, e\}$ yields a causal model that determines e to be true. Due to the symmetry of the scenario, a is a cause of e.[8]

Overdetermination is trouble for the counterfactual account of Lewis (1973). There, Lewis defines actual causation as the transitive closure of counterfactual dependence between occurring events. Let c and e be distinct events. c is a cause of e iff c and e occur, and there is a sequence $\langle c, d_1, ..., d_n, e \rangle$ of distinct events and absences such that each element in the sequence (except the first) counterfactually depends on its predecessor in a non-backtracking way.[9] Recall that e counterfactually depends on c just in case if c were not to occur, e would not occur. Lewis insists that each counterfactual in the series of counterfactual dependences is non-backtracking.[10] A backtracking counterfactual retraces some past causes from an effect: if the effect e were not to occur, its past causes c and a must have been absent. Intuitively, this backtracking counterfactual is true in the confines of the overdetermination scenario. Yet Lewis does not allow such backtracking counterfactuals to figure in the series of counterfactual dependences.

7 We will not always explicitly mention this intervention by actuality in the scenarios to come.

8 The final analysis of section 4 counts the set $\{c, a\}$ as a cause of e.

9 Lewis (1986b, 189) says that an absence $\neg a$ is the non-occurrence of any event of type A. If the absence $\neg a$ had not been, some token event a of type A would have been. Counterfactual dependence between occurring events is thus only a special case of counterfactual dependence between actual events and absences. The latter is still sufficient for causation, or so argues Lewis.

10 See Lewis (1986b, 201; 1973, 566; 1979, 456–459). Lewis (1979, 456) characterises reasoning by backtracking as follows: "We know that present conditions have their past causes. [...] if the present were different then these past causes would have to be different." The exclusion of backtracking counterfactuals plays a crucial role in Lewis's analysis of causation. For subtle details regarding backtracking counterfactuals see Lewis (1979).

doi: 10.48106/dial.v76.i1

It follows from Lewis's account that non-backtracking counterfactual dependence between occurring events is sufficient for causation. As soon as c and e occur, there is a sequence $\langle c, e \rangle$. If, in addition, e counterfactually depends on c in a non-backtracking way, c is a cause of e. In the scenario of overdetermination, c is not a cause of e on this account.[11] There is no suitable series of counterfactual dependences. If c had not fired, e would have been excited all the same. After all, a would still have fired and excited e. Due to the symmetry of the scenario, a is not a cause of e either. But then, what caused the death of the prisoner? Surely, we do not want to say that the death is uncaused.

The counterfactual accounts of causation due to Hitchcock (2001) and Halpern and Pearl (2005) solve the scenario of overdetermination as follows: c is a cause of e because e counterfactually depends on c if $\neg a$ is set by intervention. Their tests for causation allow for non-actual contingencies, that is, to set variables to non-actual values and to keep them fixed at these non-actual values. We will see that this feature is problematic in switching scenarios and extended double prevention.

Halpern (2015) modifies the Halpern and Pearl (2005) definition of actual causation. The main difference is that the modified definition admits only actual contingencies for the counterfactual test. Hence, the modified definition fails to recognize the individual overdeterminers as actual causes, while it counts the set $\{c, a\}$ of overdeterminers to be an actual cause of e.[12] It has troubles to handle overdetermination, as already pointed out by Andreas and Günther (2021b). This indicates that overdetermination haunts counterfactual accounts to date.

3.2 *Conjunctive Causes*

In a scenario of conjunctive causes, an effect occurs only if two causes obtain. The neuron diagram in figure 2 depicts a scenario of conjunctive causes:

The neurons c and a fire. Together they bring the stubborn neuron e to fire. Had one of c and a not fired, e would not have been excited. Hence, the firing of both neurons is necessary for e's excitation.

11 Lewis (2000) refines his earlier account. There, the idea to hold certain events fixed while altering others surfaces (2000, 191). However, he does not advocate to keep certain merely possible events or absences fixed. Hence, his refined account does not solve overdetermination either.

12 This being said, Halpern (2015) calls each conjunct of an actual cause *part of a cause*.

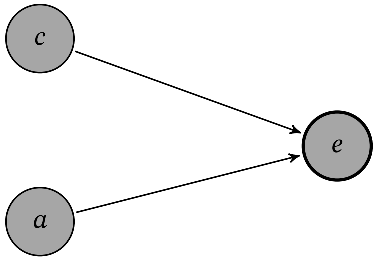


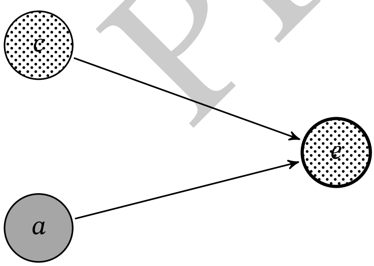
Figure 2: CAPTION NEEDED (Conjunctive causes?)

Our recipe translates the neuron diagram of figure 2 into the following causal model $\langle M, V \rangle$:

$e = c \wedge a$
c, a, e

The scenario of conjunctive causes differs from the scenario of overdetermination only in the structural equation for e. While the structural equation is *disjunctive* in the scenario of overdetermination, here the equation is *conjunctive*. The occurrence of both events, c and a, is necessary for e to occur.

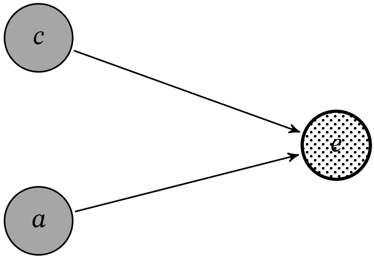
Relative to $\langle M, V \rangle$, c is a cause of e. For this to be seen, consider the following causal model $\langle M, V' \rangle$ that is uninformative on e.



$e = c \wedge a$
a

Intervening by $\{c\}$ yields:

doi: 10.48106/dial.v76.i1



c $e = c \wedge a$
a

Obviously, this causal model determines e to be true. In more formal terms, $\langle M_{\{c\}}, V' \rangle$ satisfies e. Again, due to the symmetry of the scenario, a is a cause of e.[13]

At first sight, conjunctive causes seem to be no problem for counterfactual accounts. If c had not fired, e would not have fired. Hence, on the counterfactual accounts, c is a cause of e. And by the symmetry of the scenario, a is a cause of e. However, the accounts due to Lewis (1973) and Hitchcock (2001) do not allow sets of events to be causes, unlike the definitions of actual causation provided by Halpern and Pearl (2005) and Halpern (2015). Yet the latter definitions still do not count the set containing c *and* a as an actual cause of e in this scenario of *conjunctive* causes. Hence, none of these counterfactual accounts counts the set containing the two individual causes as a cause of the effect. This is peculiar for reasons worked out by Andreas and Günther (2021b).

3.3 *Early Preemption*

Preemption scenarios are about backup processes: there is an event c that, intuitively, causes e. But even if c had not occurred, there is a backup event a that would have brought about e. Paul and Hall (2013, 75) take the following neuron diagram as canonical example of early preemption:

c's firing excites neuron d, which in turn leads to an excitation of neuron e. At the same time, c's firing inhibits the excitation of b. Had c not fired, however, a would have excited b, which in turn would have led to an excitation of e. The actual cause c preempts the mere potential cause a.[14]

13 The final analysis of section 4 counts the set $\{c, a\}$ as a cause of e.

14 Following Halpern and Pearl (2005, 861–862), we take the model of symmetric overdetermination in section 3.1 to be inappropriate for representing the structure of the early preemption scenario.

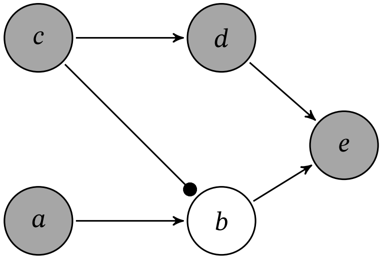
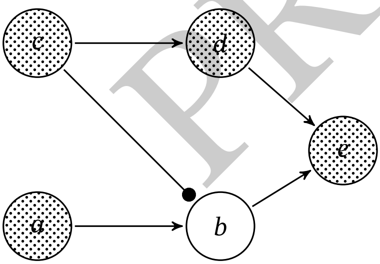


Figure 3: CAPTION NEEDED (Preemption?)

Our recipe translates the neuron diagram of early preemption into the following causal model $\langle M, V \rangle$:

$$\begin{array}{|l|}\hline d = c \\ b = a \wedge \neg c \\ e = d \vee b \\ \hline c, a, d, \neg b, e \\ \hline \end{array}$$

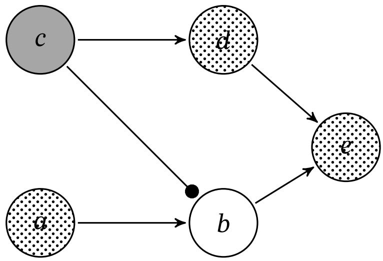
Relative to $\langle M, V \rangle$, c is a cause of e. For this to be seen, consider the following causal model $\langle M, V' \rangle$ that is uninformative on e.



$$\begin{array}{|l|}\hline d = c \\ b = a \wedge \neg c \\ e = d \vee b \\ \hline \neg b \\ \hline \end{array}$$

Intervening by $\{c\}$ yields:

doi: 10.48106/dial.v76.i1



$$c$$
$$d = c$$
$$b = a \wedge \neg c$$
$$e = d \vee b$$
$$\neg b$$

Obviously, this causal model determines e to be true. In more formal terms, $\langle M_{\{c\}}, V' \rangle$ satisfies e.

Relative to $\langle M, V \rangle$, a is not a cause of e. The reason is that actuality intervenes. The causal model $\langle M, V' \rangle$ is uninformative on e only for $V' = \emptyset$ or $V' = \{\neg b\}$. Intervening on $\langle M, V' \rangle$ by $\{\neg b\}$ yields a causal model in which a does not produce e, independently of the choice of V'. In more formal terms, $\langle M_{\{\neg b\}}[\{a\}], V' \rangle$ does not satisfy e. For each choice of V', there is a complete extension that satisfies the structural equations $a, \neg b, d = c$, and $e = d \vee b$ but does not satisfy e. This extension of V' is $\{a, \neg b, \neg c, \neg d, \neg e\}$. Intuitively, a is not a genuine cause of e since a would produce e only via an event b that did not actually occur. Hence, a is not a cause of e because a does not *actually* produce e.

Lewis's (1973) account solves early preemption. In figure 3, c is a cause of e. Both occur, and there is a sequence $\langle c, d, e \rangle$ such that e counterfactually depends in a non-backtracking way on d, and d does so on c. The counterfactual 'If d had not fired, its cause c would have to have not fired' is backtracking. Barring backtracking, we do not obtain that b would have fired because c did not, and thus b would not be inhibited. Hence, if d had not fired, b would still not have fired. And so 'If d had not fired, e would not have fired' comes out true under the non-backtracking requirement. a is not a cause of e. For there is no sequence of events and absences from a to e where each counterfactually depends on its predecessor in a non-backtracking way. If b had fired, e would still have fired.

The solution to early preemption by Hitchcock (2001) and Halpern and Pearl (2005) is analogous to their solution for overdetermination. c is a cause of e because e counterfactually depends on c under the contingency that $\neg b$. By contrast to their solution for overdetermination, the contingency is actual in cases of early preemption. Hence, Halpern's (2015) account solves early preemption as well.

3.4 *Late Preemption*

Lewis (1986b, 200) subdivides preemption into early and late. We have discussed early preemption in the previous section: a backup process is cut off before the process started by the preempting cause brings about the effect. In scenarios of late preemption, by contrast, the backup process is cut off only because the genuine cause brings about the effect before the preempted cause could do so. Lewis (2000, 184) provides the following story for late preemption:

> Billy and Suzy throw rocks at a bottle. Suzy throws first, or maybe she throws harder. Her rock arrives first. The bottle shatters. When Billy's rock gets to where the bottle used to be, there is nothing there but flying shards of glass. Without Suzy's throw, the impact of Billy's rock on the intact bottle would have been one of the final steps in the causal chain from Billy's throw to the shattering of the bottle. But, thanks to Suzy's preempting throw, that impact never happens.

Crucially, the backup process initiated by Billy's throw is cut off only by Suzy's rock impacting the bottle. Until her rock impacts the bottle, there is always a backup process that would bring about the shattering of the bottle an instant later.[15]

Halpern and Pearl (2005, 861–862) propose a causal model for late preemption, which corresponds to the following neuron diagram:

Suzy throws her rock (c) and Billy his (a). Suzy's rock impacts the bottle (d), and so the bottle shatters (e). Suzy's rock impacting the bottle (d) prevents Billy's rock from impacting the bottle ($\neg b$). (The "inhibitory signal" from d takes "no time" to arrive at b.)

Our recipe translates the neuron diagram of late preemption into the following causal model $\langle M, V\rangle$:

15 The problem posed by late preemption can be solved by fine-grained individuation conditions for events. According to these conditions, the shattering of the bottle and the shattering of the bottle an instant later are two different events. By adopting this strategy counterfactual accounts run into the trouble of spurious causation: they identify causal relations where, intuitively, there are none. See, for instance, Lewis (1986b, 204–205), Collins, Hall and Paul (2004, 45–48) and Paul and Hall (2013, chap. 3.4.2).

doi: 10.48106/dial.v76.i1

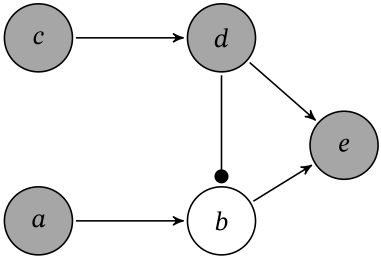
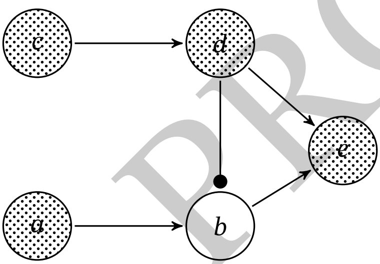


Figure 4: CAPTION NEEDED

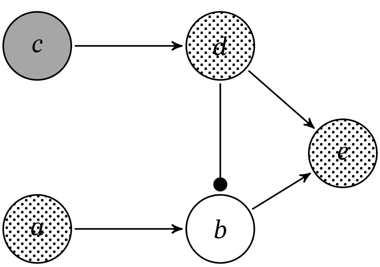
$d = c$
$b = a \wedge \neg d$
$e = d \vee b$
$c, a, d, \neg b, e$

Relative to $\langle M, V \rangle$, c is a cause of e. For this to be seen, consider the following causal model $\langle M, V' \rangle$ that is uninformative on e.



$d = c$
$b = a \wedge \neg d$
$e = d \vee b$
$\neg b$

Intervening by $\{c\}$ yields:



c
$d = c$
$b = a \wedge \neg d$
$e = d \vee b$
$\neg b$

Obviously, this causal model determines e to be true. In more formal terms, $\langle M_{\{c\}}, V' \rangle$ satisfies e.

Relative to $\langle M, V \rangle$, a is not a cause of e. The intuitive reason is that Billy's rock did not actually impact the bottle. The formal reasoning is perfectly analogous to the one for the scenario of early preemption in the previous section. Our analysis solves early and late preemption in a uniform manner.

Lewis's (1973) account does not solve late preemption. Suzy's throw (c) is not a cause of the bottle shattering (e). There is no sequence $\langle c, \ldots, e \rangle$ of events and absences such that each event (except c) counterfactually depends on its predecessor in a non-backtracking way. There is, of course, the sequence $\langle c, d, e \rangle$, and if Suzy had not thrown ($\neg c$), her rock would not have impacted the bottle ($\neg d$). However, if Suzy's rock had not impacted the bottle ($\neg d$), the bottle would have shattered anyways (e). The reason is that—on a non-backtracking reading—if Suzy's rock had not impacted the bottle ($\neg d$), Billy's rock would have (b). But if Billy's rock had impacted the bottle (b), it would have shattered (e). By contrast to scenarios of early preemption, there is no chain of stepwise dependences that run from cause to effect: there is no sequence of non-backtracking counterfactual dependences that links Suzy's throw and the bottle's shattering.[16]

The counterfactual accounts of causation due to Hitchcock (2001), Halpern and Pearl (2005), and Halpern (2015) solve the scenario of late preemption analogous to early preemption. c is a cause of e because e counterfactually depends on c under the contingency that $\neg b$.

3.5 *Simple Switch*

In switching scenarios, some event f helps to determine the causal path by which some event e is brought about (Hall 2000, 205). The following neuron diagram represents a simple version of a switching scenario:

The firing of neuron f excites r's firing, which in turn excites neuron e. At the same time, f's firing inhibits the excitation of l. The neuron l is a little special: it would have been excited in case f had not fired. f determines which one of l and r is firing, and thus determines the causal path by which e is excited. We say f acts like a switch as to e.

16 Lewis (2000) claims to solve late preemption. This claim is highly controversial. See, for instance, Paul (1998).

doi: 10.48106/dial.v76.i1

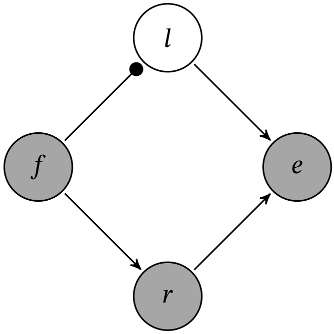


Figure 5: CAPTION NEEDED (Switching?)

Let us supplement our neuron diagram by a story due to Hall (2007, 28). Flipper is standing by a switch in the railroad tracks. A train approaches in the distance. She flips the switch (f), so that the train travels down the right track (r), instead of the left (l). Since the tracks reconverge up ahead, the train arrives at its destination all the same (e). We agree with Hall that flipping the switch is not a cause of the train's arrival. The story assumes that flipping the switch makes no difference to the train's arrival: "the train arrives at its destination all the same." The flipping merely switches the causal path by which the train arrives.[17]

Our recipe translates the neuron diagram of the switching scenario into the following causal model $\langle M, V \rangle$:

$$\begin{array}{|c|} \hline l = \neg f \\ r = f \\ e = l \lor r \\ \hline f, \neg l, r, e \\ \hline \end{array}$$

Relative to $\langle M, V \rangle$, f is not a cause of e. The reason is that there exists no causal model $\langle M, V' \rangle$ uninformative on e. Any complete extension of the empty set V'

17 There is a noteworthy difference between switching scenarios and scenarios of preemption. If the non-actual switch position $\neg f$ were actual, $\neg f$ would help bring about e. By contrast, if it were actual that the genuine cause does not occur in scenarios of preemption, it's absence would not help bring about the effect. If Suzy were not to throw her rock, her not throwing would not help to bring about the bottle's shattering.

of literals that satisfies the structural equations of M contains e. In fact, there are only two complete extensions that satisfy the structural equations, viz. the actual $\{f, \neg l, r, e\}$ and the non-actual $\{\neg f, l, \neg r, e\}$. The structural equations in M determine e no matter what.[18]

Our analysis requires for c to be a cause of e that there must be a causal model uninformative about e in which c brings about e. The idea is that, for an event to be caused, it must arguably be possible that the event does not occur. However, in the switching scenario, there is no causal model uninformative on e in the first place. Hence, f is not a cause of e in the simple switch.

A simplistic counterfactual analysis says that an event c is a cause of a distinct event e just in case both events occur, and e would not occur if c had not occurred. This suggests that the switching scenario is no challenge for counterfactual accounts, because e would occur even if f had not. And yet it turns out that cases like the switching scenario continue to be troublesome for counterfactual accounts.

Recall that Lewis (1973) defines actual causation to be the transitive closure of non-backtracking counterfactual dependence between occurring events. In the switching scenario, f, r, and e occur, and both r counterfactually depends on f in a non-backtracking way and e does so on r. Barring backtracking, if r had not fired, e would not have fired. By the transitive closure imposed on the one-step causal dependences, Lewis (1973) is forced to say that f is a cause of e.[19]

The sufficiency of (non-backtracking) counterfactual dependence for causation is widely shared among the accounts in the tradition of Lewis, for instance by Hitchcock (2001), Woodward (2003), Hall (2004, 2007), and Halpern and Pearl (2005). However, the counterfactual accounts based on structural equations reject the transitivity of causation. Still, Hitchcock (2001) counts f to be a cause of e. The reason is that there is an active causal path from f over r to e and keeping the off-path variable l fixed at its actual value induces a counterfactual dependence of e on f. Similarly, Halpern and Pearl (2005) and Halpern (2015) count f to be a cause of e, since e counterfactually depends

18 Hall (2007, 118) writes that the "basic" switch in Paul and Hall (2013, 232) has "the obvious causal model": $M = \{b = a, l = b \wedge f, r = b \wedge \neg f, e = l \vee r\}, V = \{a, b, f, l, \neg r, e\}$. Relative to this causal model, our analysis says that f is not a cause of e, as desired. Relative to the causal scenario, where the equation for e is replaced by $e = l$, our analysis says that f is a cause of e, as desired (Paul and Hall 2013, 235).

19 Lewis (2000, 194–195) still imposes transitivity on his refined analysis of causation. As a result, the refined anlysis is also forced to say that f is a cause of e in the switching scenario.

doi: 10.48106/dial.v76.i1

on f under the actual contingency that $\neg l$. Hence, even the contemporary counterfactual accounts misclassify f to be a cause of e.[20] Allowing for actual contingencies solved preemption, but leads to trouble in switching scenarios. Without allowing for actual contingencies, it is unclear how the counterfactual accounts solve preemption. It seems as if the sophisticated counterfactual accounts have no choice here but to take one hit.

3.6 *Realistic Switch*

The representation of switching scenarios is controversial. Some authors criticize the simple switch in figure 5 from the previous section because they believe that any "real-world" event has more than one causal influence (e.g., Hitchcock 2009, 396). The idea is that the train can only pass on the right track because nothing blocks the track, it is in good conditions, and so on. These critics insist on "realistic" scenarios in which there is always more than just one event that causally affects another. The simple switch is thus inappropriate because there must be another neuron whose firing is necessary for the excitation of l. Some authors then quickly point out that the causal model of the resulting switch is indistinguishable from the one of early preemption (e.g., Beckers and Vennekens 2018, 848–851). And this is a problem for any account of causation that only relies on causal models. For c should intuitively be a cause of e in early preemption, but f should not be a cause in a "realistic" switching scenario.[21]

It is too quick to point out that switches and early preemption are structurally indistinguishable. After all, the critics who insist on "realistic" scenarios are bound to say that there should also be another neuron whose firing is necessary for the excitation of r. This restores the symmetry between l and r which seems to be essential to switching scenarios. The following neuron diagram depicts our realistic switch:

The joint firing of neurons f and h excites r's firing, which in turn excites neuron e. At the same time, f's firing inhibits the excitation of l. Had f not fired, the firing of g would have excited l, which in turn would have excited e. In the actual circumstances, f determines which one of l and r is firing, and thus acts like a switch as to e.

20 Halpern (2015) uses normality considerations to solve the present switching scenario. See Blanchard and Schaffer (2017) for a criticism of this strategy.

21 The problem posed by structurally indistinguishable causal models where our intuitive causal judgments differ is further discussed in section 5.1.

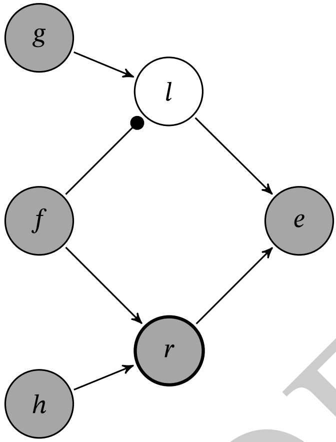
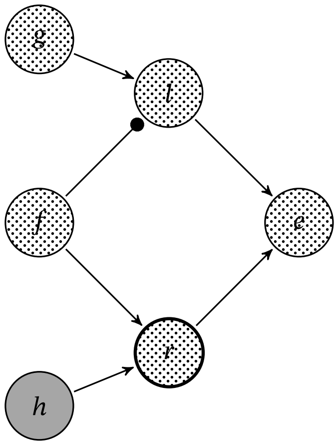


Figure 6: CAPTION NEEDED (Realistic switch?)

Our recipe translates the neuron diagram of our realistic switch into the following causal model $\langle M, V \rangle$:

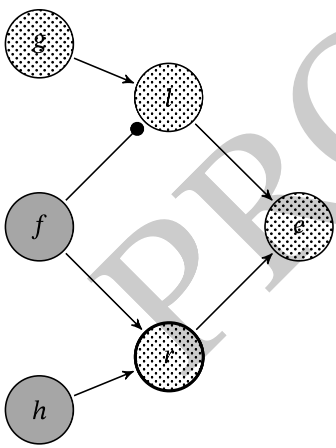
$l = g \land \neg f$ $r = f \land h$ $e = l \lor r$
$g, f, h, \neg l, r, e$

Relative to $\langle M, V \rangle$, f is a cause of e according to our preliminary analysis. For this to be seen, consider the following causal model $\langle M, V' \rangle$ that is uninformative on e.



$$l = g \wedge \neg f$$
$$r = f \wedge h$$
$$e = l \vee r$$
$$h$$

Intervening by $\{f\}$ yields:



$$f$$
$$l = g \wedge \neg f$$
$$r = f \wedge h$$
$$e = l \vee r$$
$$h$$

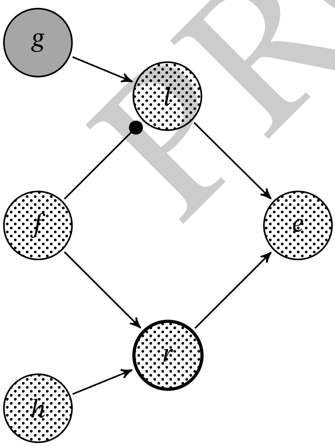
Obviously, this causal model determines e to be true. In more formal terms, $\langle M_{\{f\}}, V' \rangle$ satisfies e. Our preliminary analysis wrongly counts the "realistic switch" f as a cause of e.

It is time to amend our preliminary analysis by a condition of *weak difference making*. The idea is this: if some event c is a cause of an event e, then it is not the case that $\neg c$ would be a cause of the same event e. Sartorio (2006, 75) convinces us that this principle of weak difference making is a condition "the true analysis of causation (if there is such a thing) would have to meet."[22] But this condition is violated by "realistic switches": f helps to bring about an effect e, and so would the non-actual $\neg f$. So a "realistic switch" is not a cause if we demand of any genuine cause c of some effect e that $\neg c$ would not also bring about e. We demand that $\neg c$ would not also bring about e by the following condition:

(C3) There is no $V'' \subset V \setminus \{c\}$ such that $\langle M, V'' \rangle$ is uninformative on e and $\langle M[\{\neg c\}], V'' \rangle$ satisfies e.

(C3) demands that there is no causal model uninformative on e in which e is actual if $\neg c$ is. The condition ensures that a cause is a difference maker in the weak sense that its presence and its absence could not bring about the same effect. This implies Sartorio's principle of weak difference making: if c is a cause of e, then $\neg c$ would not also be a cause of e. And note that our condition of difference making is weaker than the difference making requirement of (sophisticated) counterfactual accounts of causation. Unlike them, we do not require that $\neg e$ is actual under the supposition that $\neg c$ is actual (given certain contingencies).

(C3) ensures that f is not a cause of e in the realistic switch. For this to be seen, consider the following causal model $\langle M, V'' \rangle$ that is uninformative on e.

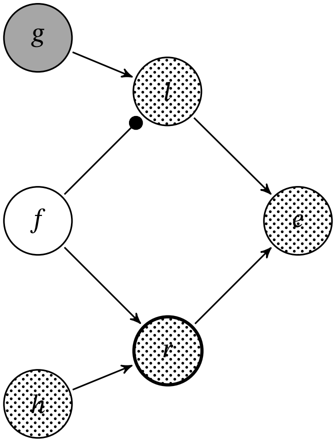


$l = g \wedge \neg f$ $r = f \wedge h$ $e = l \vee r$
g

Intervening by $\{\neg f\}$ yields:

22 For more details, see Andreas and Günther (2020, 1584, 1590).

doi: 10.48106/dial.v76.i1



$$
\begin{array}{|l|}
\hline
f \\
l = g \land \neg f \\
r = f \land h \\
e = l \lor r \\
\hline
h \\
\hline
\end{array}
$$

Obviously, this causal model determines e to be true. In more formal terms, $\langle M_{\{f\}}, V'' \rangle$ satisfies e. Our preliminary analysis amended by (C3) says that the "realistic switch" f is not a cause of e, as desired.[23] We will leave it as an exercise for the reader to check that (C3) does not undo any causes our preliminary definition identifies in this paper, except for the "realistic switches."

Lewis's (1973) account misclassifies f as a cause of e in our realistic switch. As in the simple switch, there is a causal chain running from f to e: the sequence $\langle f, r, e \rangle$ of actual events such that each event (except f) counterfactually depends on its predecessor in a non-backtracking way. Similarly, Hitchcock (2001), Halpern and Pearl (2005), and Halpern (2015) all misclassify f as a cause of e. The reasons are analogous to the reasons in the simple switch. Roughly, e counterfactually depends on f when l is fixed at its actual value.

23 Hitchcock (2009, 395–396) modifies Paul and Hall's (2013) "basic" switch of footnote 18. The modified switch has the "obvious causal model": $M = \{b = a, l = g \land b \land f, r = b \land h \land \neg f, e = l \lor r\}$, $V = \{a, g, b, h, f, l, \neg r, e\}$. Relative to this causal model, (C3) rules out that f is a cause of e, as desired. Halpern and Hitchcock (2010, 16) and Halpern (2016, 72) propose to model the train scenario by the following causal model: $M = \{e = (f \land \neg lb) \lor (\neg f \land \neg rb)\}$, $V = \{f, \neg lb, \neg rb, e\}$. The variables rb and lb indicate whether or not the right and left track are blocked, respectively. Relative to this causal model, (C3) rules out that f is a cause of e, as desired.

3.7 *Prevention*

To prepare ourselves for a discussion of double prevention, let us take a look at simple prevention first. Paul and Hall (2013, 174) represent the basic scenario of prevention by the following neuron diagram:

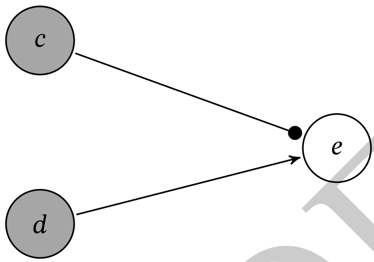


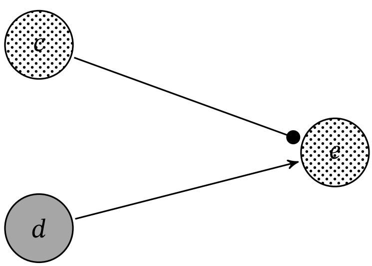
Figure 7: CAPTION NEEDED (Prevention?)

Neuron c fires and thereby inhibits that neuron e gets excited. e would have been excited by d if the inhibitory signal from c were absent. But as it is, c prevents e from firing. That is, c causes $\neg e$ by prevention.

Our recipe translates the neuron diagram of prevention into the following causal model $\langle M, V \rangle$:

$e = \neg c \wedge d$
$c, d, \neg e$

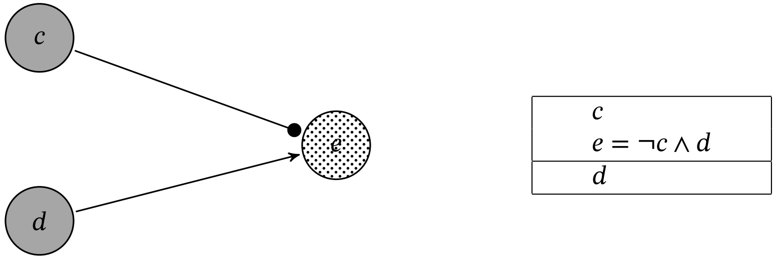
Relative to $\langle M, V \rangle$, c is a cause of $\neg e$. For this to be seen, consider the following causal model $\langle M, V' \rangle$ that is uninformative on $\neg e$.



$e = \neg c \wedge d$
d

Intervening by $\{c\}$ yields:

doi: 10.48106/dial.v76.i1



Obviously, this causal model determines $\neg e$ to be true. In more formal terms, $\langle M_{\{c\}}, V' \rangle$ satisfies $\neg e$. Moreover, d is not a cause of $\neg e$ relative to $\langle M, V \rangle$. Any causal model $\langle M, V' \rangle$ uninformative on $\neg e$ must be uninformative on c as well. Intervening by d in $\langle M, V' \rangle$ does not determine $\neg e$.

Counterfactual accounts face no challenge here. If c had not fired, e would have fired. Counterfactual dependence between actual events and absences is sufficient for causation. Hence, c is a cause of $\neg e$. If d had not fired, e would not have fired, even under the contingency that c did not fire. Hence, d is not a cause of $\neg e$.

3.8 *Double Prevention*

Double prevention can be characterized as follows. c is said to double prevent e if c prevents an event that, had it occurred, would have prevented e. In other words, c double prevents e if c cancels a threat for e's occurrence. Paul and Hall (2013, 154, 175) represent an example of double prevention by the following neuron diagram:

c's firing prevents d's firing, which would have prevented e's firing. The example of double prevention exhibits a counterfactual dependence: given that b fires, e's firing counterfactually depends on c's firing. If c did not fire, d would fire, and thereby prevent e from firing. Hence, c's firing double prevents e's firing in figure 8. In other words, c's firing cancels a threat for e's firing, viz. the threat originating from b's firing.

Paul and Hall (2013) say that c is a cause of e in the scenario of figure 8. They thereby confirm that there is causation by double prevention. e counterfactually depends on c. Hence, the accounts of causation due to Lewis (1973, 2000), Hitchcock (2001), Halpern and Pearl (2005), and Halpern (2015) agree with Paul and Hall in counting c a cause of e. How does our account fare?

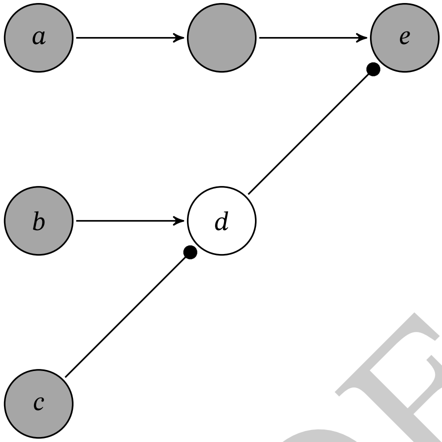


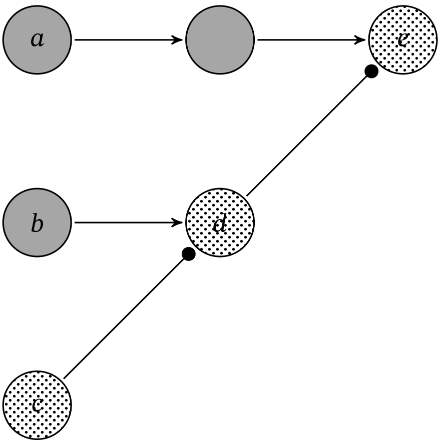
Figure 8: CAPTION NEEDED (Double prevention?)

Our recipe translates the neuron diagram of double prevention into the following causal model $\langle M, V \rangle$:

$d = b \wedge \neg c$ $e = a \wedge \neg d$
$a, b, c, \neg d, e$

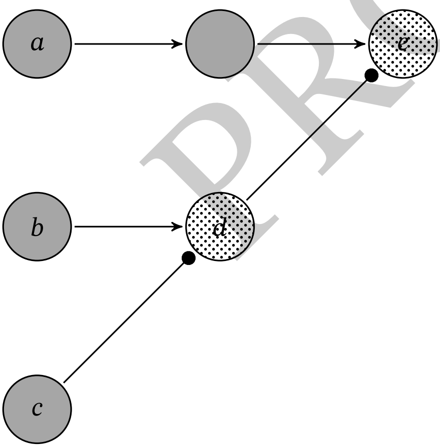
Relative to $\langle M, V \rangle$, c is a cause of e. For this to be seen, consider the following causal model $\langle M, V' \rangle$ that is uninformative on e.

doi: 10.48106/dial.v76.i1



$$d = b \wedge \neg c$$
$$e = a \wedge \neg d$$
$$a, b$$

Intervening by $\{c\}$ yields:



$$c$$
$$d = b \wedge \neg c$$
$$e = a \wedge \neg d$$
$$a, b$$

Obviously, this causal model determines $\neg d$ and so e to be true. In more formal terms, $\langle M_{\{c\}}, V' \rangle$ satisfies e.

3.9 *Extended Double Prevention*

Hall (2004, 247) presents an extension of the scenario depicted in figure 8. The extended double prevention scenario fits the structure of the following neuron diagram:

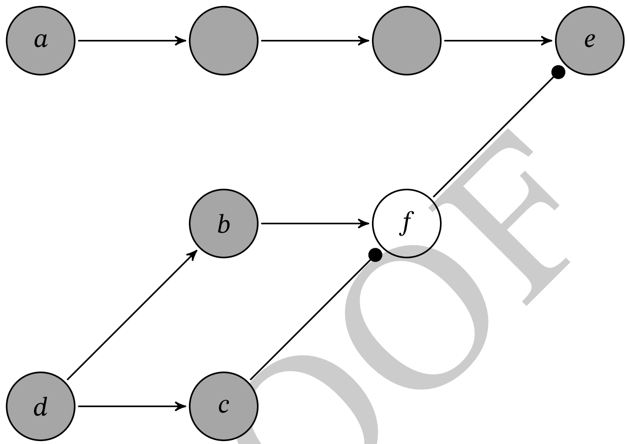


Figure 9: CAPTION NEEDED

Figure 9 extends figure 8 by neuron *d*, which figures as a common cause of *b* and *c*. *d* starts a process via *b* that threatens to prevent *e*. At the same time, *d* initiates another process via *c* that prevents the threat. *d* cancels its *own* threat—the threat via *b*—to prevent *e*. In the example of the previous section, the threat originated independent of its preventer. Here, by contrast, *d* creates and cancels the threat to prevent *e*. This difference is sufficient for *d* not to be a cause of *e*, or so argue for instance Paul and Hall (2013, 216). Observe that the structure characteristic of double prevention is embedded in figure 9. The firing of neuron *c* inhibits *f*'s firing that, had it fired, would have inhibited *e*'s firing. Nevertheless, this scenario of double prevention exhibits an important difference to its relative of the previous section: *e* does not counterfactually depend on *d*. If *d* had not fired, *e* would still have fired.

Hitchcock (2001, 276) provides a story that matches the structure of the scenario. A hiker is on a beautiful hike (*a*). A boulder is dislodged (*d*) and rolls toward the hiker (*b*). The hiker sees the boulder coming and ducks (*c*),

doi: 10.48106/dial.v76.i1

so that he does not get hit by the boulder ($\neg f$). If the hiker had not ducked, the boulder would have hit him, in which case the hiker would not have continued the hike. Since, however, he was clever enough to duck, the hiker continues the hike (e).

Hall (2007, 36) calls the subgraph $d - b - c - f$ a *short circuit* with respect to e: the boulder threatens to prevent the continuation of the hike, but provokes an action that prevents this threat from being effective. Like switching scenarios, the scenario seems to show that there are cases where causation is not transitive: the dislodged boulder d produces the ducking of the hiker c, which in turn enables the hiker to continue the hike e. But it is counterintuitive to say that the dislodging of the boulder d causes the continuation of the hike e. After all, the dislodgement of the boulder is similar to a switch as to the hiker *not* getting hit by the boulder: d helps to bring about $\neg f$, and if $\neg d$ were actual, $\neg d$ would also help to bring about $\neg f$. In this sense, d is causally inert.

Our recipe translates the neuron diagram of the boulder scenario into the following causal model $\langle M, V \rangle$:

$b = d$ $c = d$ $f = b \wedge \neg c$ $e = a \wedge \neg f$
$a, d, b, c, \neg f, e$

Relative to $\langle M, V \rangle$, d is not a cause of e. The reason is that the causal model $\langle M, V' \rangle$ is only uninformative on e if a is not in V'. But then $\langle M_{\{d\}}, V' \rangle$ does not satisfy e.

In words, the causal model $\langle M, V' \rangle$ is uninformative about e only if a is not in the set V' of literals. But then intervening with d does not make e true. After all, a is necessary for determining e. If we were to keep a in the literals, the model would not be uninformative. There is no complete extension of $V' = \{a\}$ that satisfies all the structural equations of M but fails to satisfy e.

On Lewis's (1973) account, d is a cause of e. There is a sequence $\langle d, c, \neg f, e \rangle$ of events and absences such that each element (except d) counterfactually depends on its predecessor in a non-backtracking way. The structural equation accounts of Hitchcock (2001), Halpern and Pearl (2005), and Halpern (2015) classify d as a cause of e. The reason is that e counterfactually depends on d under the contingency that b.

The situation is bad for the sophisticated counterfactual accounts. While their general strategy to allow for possibly non-actual contingencies solves overdetermination and preemption, it is the very same strategy that is at fault for the unintuitive results in the switching scenario and extended double prevention. The backfiring of their general strategy casts doubt on whether it was well motivated in the first place. If the general strategy is merely motivated by solving overdetermination, it turns out that overdetermination still haunts the sophisticated accounts of causation. By contrast to these counterfactual accounts, our analysis of actual causation solves overdetermination without further ado. Our analysis has thus a major advantage over the sophisticated counterfactual accounts.

4 Final Analysis

In section 1, we stated a preliminary version of our analysis and amended it in section 3.6 by condition (C3). The amended version is still preliminary because it assumes that both the cause and the effect are single events. This assumption is violated in certain causal scenarios. Recall, for instance, the scenario of conjunctive causes from section 3.2. There, two events are necessary for an effect to occur, and so the set containing the two events should count as a cause of said effect. To give an example, lightning resulted in a forest fire only because of a preceding drought. Here, it seems plausible that lightning together with the preceding drought is an—if not *the*—cause of the forest fire.[24]

We lift the restriction of cause and effect to single literals as follows. A cause is a set of literals C, an effect an arbitrary Boolean formula. Where C is a set of literals, $\bigwedge C$ stands for the conjunction of all literals in C and $\neg C$ for the negation of all literals in C. Our final analysis of actual causation can now be stated.

Definition 3 (Actual Cause). Let $\langle M, V\rangle$ be a causal model such that V satisfies M. C is a set of literals and ε a formula. C is an actual cause of ε relative to $\langle M, V\rangle$ iff

(C1*) $\langle M, V\rangle$ satisfies $\bigwedge C \wedge \varepsilon$, and

24 Andreas and Günther (2021b, 608–610) argue that it is desirable if an account of causation can count sets of events as causes.

(C2*) there is $V' \subset V$ such that $\langle M, V' \rangle$ is uninformative on ε, while for all $A \subseteq V$ and all non-empty $C' \subseteq C$, $\langle M_A[C'], V' \rangle$ satisfies ε; and

(C3*) there is no $V'' \subset V \setminus C$ such that $\langle M, V'' \rangle$ is uninformative on ε and $\langle M[\neg C], V'' \rangle$ satisfies ε.

In this more general analysis, clause (C2*) contains a minimality condition ensuring that any cause contains only causally relevant literals. For this to be seen, suppose there is a set $C' \subset C$ whose members are causally irrelevant for ε. That is, intervening by C' in any partial model uninformative on ε does not make ε true (under all interventions by actuality). Then, by the minimality condition, C would not be a cause, contrary to our assumption. Thanks to this condition, causally irrelevant factors cannot simply be added to genuine causes.[25]

How fare the counterfactual accounts with respect to sets of causes? Let us consider the scenario of overdetermination. As explained in section 4, Halpern's (2015) account counts only the set of individual causes as a genuine cause. The other counterfactual accounts do not count this set as a cause. We think it is reasonable to recognize both the individual causes and the set of these causes as a proper cause. We would say that, for instance, two soldiers shooting a prisoner, where each bullet is fatal without any temporal precedence, is a perfectly fine cause for the death of the prisoner. The shooting of the two soldiers brings about the death of the prisoner.

The account of Hitchcock (2001) does not admit causes that are sets of variables. Hence, the set containing the two individual causes does not count as a cause in the scenarios of overdetermination and conjunctive causes. Unlike Hitchcock's account, the accounts due to Halpern and Pearl (2005) and Halpern (2015) admit causes to be sets of variables. Still, these accounts do not recognize the set containing the two individual causes as a cause in the scenario of conjunctive causes. The accounts share the same minimality condition according to which a strict superset of a cause cannot be a cause. Hence, they are forced to say that, for instance, the drought together with the lightning is not a cause of the forest fire *because* one of these events (and indeed both) already counts as a cause for this effect. This reason for why the set is not a cause is a little odd.

25 If one wants cause and effect to be distinct, one should amend definition 3 by a clause like this: no element of C occurs in ε.

5 Comparison

In this section, we compare our analysis to the considered counterfactual accounts. First, we focus on the results of the different accounts. Then we compare—on a conceptual level—our analysis to the counterfactual accounts that rely on causal models.

5.1 *Results*

The results of our analysis and of the considered counterfactual accounts are summarized in the following table. We abbreviate the accounts of Lewis (1973), Hitchcock (2001), Halpern and Pearl (2005), and Halpern (2015) by $\mathcal{L}$'73, Hitch'01, HP'05, and H'15, respectively.

Causes of e or $\neg e$	$\mathcal{L}$'73	Hitch'01	HP'05	H'15	Author(s)
Overdetermination	–	c, a	c, a	$\{c, a\}$	$c, a, \{c, a\}$
Conjunctive Causes	c, a	c, a	c, a	c, a	$c, a, \{c, a\}$
Early Preemption	c	c	c	c	c
Late Preemption	–	c	c	c	c
Switches	f	f	f	f	–
Prevention	c	c	c	c	c
Double Prevention	c	c	c	c	c
E. Double Prevention	d	d	d	d	–

None of the counterfactual accounts listed in the table provides the intuitively correct results for the simple and "realistic" switching scenarios and extended double prevention. Lewis's (1973) account misclassifies f and d as causes of e, respectively, because of the transitive closure he imposes on the step-wise and non-backtracking counterfactual dependences. And without imposing transitivity, his analysis of causation cannot solve early preemption. For Halpern (2015), Hitchcock (2001) and Halpern and Pearl (2005), the reason for the misclassification is that they allow for actual contingencies. And if they were not to allow for such, their accounts would fail to solve preemption. The counterfactual accounts due to Hitchcock (2001) and Halpern and Pearl (2005) solve overdetermination, but only by allowing for even non-actual contingencies.

We have thus shown that the sophisticated counterfactual accounts fail to capture the set of overdetermination, preemption, switches, and extended double prevention. And they fail for a principled reason: they can solve overdetermination and preemption only if they allow for contingencies. But, by allowing for contingencies, they fail to solve the switching scenario and extended double prevention. If they were not to allow for contingencies, they would solve the switching scenario and extended double prevention, but it would be unclear how they could solve overdetermination and preemption. Our analysis, by contrast, does not fall prey to such a principled problem.

Let us summarize the verdicts about the results, where ✓, × and ! stand for correct, false, and partially correct, respectively.

Causes of e or $\neg e$	$\mathcal{L}$'73	Hitch'01	HP'05	H'15	Author(s)
Overdetermination	×	✓	✓	!	✓
Conjunctive Causes	!	!	!	!	✓
Early Preemption	✓	✓	✓	✓	✓
Late Preemption	×	✓	✓	✓	✓
Switch	×	×	×	×	✓
Prevention	✓	✓	✓	✓	✓
Double Prevention	✓	✓	✓	✓	✓
E. Double Prevention	×	×	×	×	✓

There remains another problem to be solved. The problem concerns any account that relies on simple causal models which only factor in structural equations and values of variables (or our sets of literals). Such accounts face pairs of scenarios for which our causal judgments differ, but which are structurally indistinguishable. Overdetermination, for instance, is isomorphic to bogus prevention. In bogus prevention, an event p would prevent another event d. But, as it is, there is no event c present that would bring about d in the first place. Hence, the preventer p and the absence of c overdetermine that d does not occur. By contrast to overdetermination, however, the preventer p is intuitively not a cause of the absence $\neg d$. Since the accounts of Hitchcock (2001) and Halpern and Pearl (2005) consider only structural equations and the values of variables, they cannot distinguish between p and one of the causes in overdetermination. The former must be falsely classified to be a cause if the latter is correctly classified so.[26] And our analysis has the same problem.[27]

Hitchcock (2007a), Hall (2007), Halpern (2008), Halpern and Hitchcock (2015), and Halpern (2015) all aim to solve the problem of isomorphism by taking into account default or normality considerations. This additional factor gives considerable leeway to solve some of the isomorphic pairs. However, actual causation does not seem to be default-relative, as pointed out by Blanchard and Schaffer (2017). They also show that the accounts amended by a notion of default still face counterexamples and even invite new ones. Nevertheless, the problem of isomorphism suggests that simple causal models ignore a factor that impacts our intuitive causal judgments. We think this ignored factor are not default considerations, but a meaningful distinction between events that occur and events that do not. After all, a distinction between events and absences seems to be part of the structure of causation. Yet current accounts relying on causal models are blind to such a distinction.

Our analysis of causation is thus incomplete. We need to amend it by a meaningful distinction between events and absences, which allows us to tackle the problem of isomorphism. More generally, we miss an account of what constitutes an appropriate causal model. That is, an account that tells us which causal models are appropriate for a given causal scenario. For now, we have just assumed that the causal models obtained from simple neuron diagrams are appropriate. This assumption already smuggled in certain metaphysical assumptions about events. We will elaborate these underpinnings of our analysis elsewhere.

5.2 *Conceptual Differences*

Let us compare—on a more conceptual level—our analysis to the counterfactual accounts that likewise rely on causal models. As we have seen, these

26 As pointed out by Hiddleston (2005, 32) and Hall (2007, 44), Hitchcock's (2001), and Halpern and Pearl's (2005) allowance of non-actual contingencies solves the overdetermination scenario, but it leads to the intuitively wrong results in *bogus* cases of both prevention and double prevention. From this perspective, the non-actual contingencies, as opposed to merely actual contingencies, are thus even more bad news.

27 This being said, the causal model of bogus prevention is: $M = \{d = \neg p \wedge c\}, V = \{\neg c, p, \neg d\}$. Blanchard and Schaffer (2017, 200–202) argue that this causal model is inappropriate for bogus prevention and propose to model the bogus scenario by a model isomorphic to early preemption. If they are right, our analysis would give the correct verdict for bogus prevention. We would like to thank an anonymous referee for this observation.

doi: 10.48106/dial.v76.i1

sophisticated counterfactual accounts analyse actual causation in terms of contingent counterfactual dependence relative to a causal model. Hitchcock (2001), Halpern and Pearl (2005), and Halpern (2015), for instance, have put forth such accounts. All of these accounts have in common that the respective causal model provides full information about what actually happens, and what would happen if the state of affairs were different. Hence, causal models allow them to test for counterfactual dependence: provided *c* and *e* are actual in a causal model, would ¬*e* be actual if ¬*c* were? If so, *e* counterfactually depends on *c*; if not, not.

The mentioned accounts put forth more elaborate notions of counterfactual dependence. These notions specify which variables other than *c* and *e* are to be kept fixed by intervention when testing for counterfactual dependence. The accounts ask a test question for contingent counterfactual dependence: relative to a causal model, where *c* and *e* are actual, would ¬*e* be actual if ¬*c* were under the contingency that certain other variables are kept fixed at certain values? If so, *e* counterfactually depends on *c* under the contingency; if not, not. To figure out whether *c* is a cause of *e*, counterfactual accounts propagate forward—possibly under certain contingencies—the effects of the counterfactual assumption that a putative cause were absent.

We analyse, by contrast, actual causation in terms of production relative to a causal model that provides only partial information. More specifically, our analysis relies on models that carry no information with respect to a presumed effect *e*: they are uninformative as to whether or not the event or absence *e* is actual. Such uninformative models allow us to test whether an actual event or absence is actually produced by another. The test question goes as follows: in a model uninformative on *e*, will *e* become actual if *c* does? If so, *c* is a producer of *e*; if not, not. And a producer *c* is then a cause of *e* if ¬*c* would not also be a producer of *e*.

Our test has no need that ¬*e* becomes actual if ¬*c* were actual. Instead the question is whether, in an uninformative model, an actual event produces (and makes a weak difference to) another in accordance with what actually happened. The novelty of our account is not so much to consider actual production, but to consider production in a causal model that is uninformative on the presumed effect. As a consequence, when testing for causation, we never intervene on a causal model, where the set of actual literals is complete. This stands in stark contrast to counterfactual accounts which always intervene on causal models, where each variable is assigned a value.

On our analysis, *c* is a cause of *e* only if *c* produces *e* under *all* interventions by actuality. There is a mentionable symmetry to Halpern's (2015) account which allows only for actual contingencies. On this account, *c* is a cause of *e* if *there is* an intervention by actuality such that the actual *e* counterfactually depends on the actual *c*.[28] Production under all interventions by actuality is *necessary* for causation on our account, whereas counterfactual dependence between actual events under some intervention by actuality is *sufficient* on Halpern's.

Counterfactual notions of causation generally say that a cause is necessary for an effect: without the cause, no effect. By contrast, our notion of causation says that a cause is sufficient for its effect given certain background conditions. The background conditions are given by the partial set of literals of the causal model that is uninformative on the effect. That is, these conditions are jointly not sufficient for the effect given the structural equations. However, together with a genuine cause these conditions are jointly sufficient for the effect (given the same structural equations). Relative to the causal model uninformative on the effect, a cause is thus necessary and sufficient for its effect.[29]

6 Conclusion

We have put forth an analysis of actual causation. In essence, *c* is a cause of *e* just in case *c* and *e* are actual, and there is a causal model uninformative on *e* in which *c* actually produces *e*, and there is no such uninformative causal model in which ¬*c* would produce *e*. Our analysis successfully captures various causal scenarios, including overdetermination, preemption, switches, and extended double prevention. All extant sophisticated counterfactual accounts of causation fail to capture at least two of the causal scenarios considered.

28 The intervention by actuality on Halpern's (2015) account can just be the intervention by the empty set.

29 Perhaps, our analysis bears more resemblance to regularity analyses of causation than to counterfactual accounts. The core idea behind regularity analyses can be glossed as follows: *c* is a cause of *e* just in case, given the laws of nature, *c* together with a minimal set of background conditions is jointly sufficient for *e*. Indeed, our analysis of causation can be seen as a regularity theory when one replaces 'laws of nature' by 'structural equations' and 'minimal set of background conditions' by 'partial set of actual literals'. In a causal model uninformative on *e*, intervening by a cause *c* is sufficient to bring about the effect *e*. In a very specific sense, this says that the 'laws' and 'minimal background conditions' imply that *c* is sufficient for *e*. However, we are not aware of any regularity theory that employs an equivalent to our uninformative models.

With respect to this set, our analysis is strictly more comprehensive than those accounts.

The sophisticated counterfactual accounts, which rely on causal models, run into problems for a principled reason. They fail to solve the switching scenario and extended double prevention because they allow for possibly non-actual contingencies when testing for counterfactual dependence. Such contingencies are needed to solve the problems of overdetermination and preemption. Our analysis, by contrast, is neither premised on counterfactuals of the form 'If $\neg c$, then $\neg e$', nor on considering possibly non-actual contingencies. Hence, our analysis escapes the principled problem to which the sophisticated counterfactual accounts are susceptible.

The present analysis of causation has a counterfactual counterpart due to Andreas and Günther (2021a). The counterfactual analysis likewise relies on an information removal and uninformative causal models. The gist is this: an event c is a cause of another event e just in case both events occur, and—after removing the information whether or not c and e occur—e would not occur if c were not to occur. This analysis does not rely on the strategy common to the sophisticated counterfactual accounts, and is therefore also not susceptible to their principled problem.

The two analyses largely come to the same verdicts. However, unlike the present preliminary analysis, the preliminary counterfactual analysis cannot identify the overdetermining causes in scenarios of symmetric overdetermination. And while the present final analysis counts the set $\{c, a\}$ as a cause in the scenario of conjunctive causes, the final counterfactual analysis does not. More importantly, the present final analysis does not count "realistic switches" as causes, whereas the final counterfactual analysis does. The present analysis has therefore a slight edge over its counterfactual counterpart.

Appendix: The Framework of Causal Models

In this appendix, we supplement the explanations of the core concepts of causal models with precise definitions. Let P be a set of propositional variables such that every member of P represents a distinct event. $\mathcal{L}_P$ is a propositional language that is defined recursively as follows: (i) Any $p \in P$ is a formula. (ii) If ϕ is a formula, then so is $\neg\phi$. (iii) If ϕ and ψ are formulas, then so are $\phi \vee \psi$ and $\phi \wedge \psi$. (iv) Nothing else is a formula.

As is well known, the semantics of a propositional language centers on the notion of a value assignment. A value assignment $v : P \mapsto \{T, F\}$ maps each propositional variable on a truth value. We can represent a value assignment, or valuation for short, in terms of literals. The set $L(v)$ yields the set of literals that represents the valuation v.

Definition 4 ($L(v)$)**.** Let $v : P \mapsto \{T, F\}$ be a valuation of the language $\mathcal{L}_P$. $L(v)$ is the set of literals of $\mathcal{L}_P$ such that, for any $p \in P$, (i) $p \in L(v)$ iff $v(p) = T$, and (ii) $\neg p \in L(v)$ iff $v(p) = F$.

We say that a set V of literals is complete—relative to $\mathcal{L}_P$—iff there is a valuation v such that $L(v) = V$. If the language is obvious from the context, we simply speak of a complete set of literals, leaving the parameter P implicit.

The function $L(v)$ defines a one-to-one correspondence between the valuations of $\mathcal{L}_P$ and the complete sets of $\mathcal{L}_P$ literals. In more formal terms, $L(v)$ defines a bijection between the set of valuations of $\mathcal{L}_P$ and the set of the complete sets of $\mathcal{L}_P$ literals. Hence, the inverse function $L^{-1}(V)$ of $L(V)$ is well defined for complete sets V of literals. Using the inverse of $L(V)$, we can define what it is for a complete set V of literals to satisfy an $\mathcal{L}_P$ formula ϕ:

$$V \vDash \phi \text{ iff } L^{-1}(V) \vDash_C \phi, \qquad (V \vDash \phi)$$

where $\vDash_C$ stands for the satisfaction relation of classical propositional logic. In a similar vein, we define the semantics of a single structural equation:

$$V \vDash p = \phi \text{ iff, } L^{-1}(V) \vDash_C p \text{ iff } L^{-1}(V) \vDash_C \phi. \qquad (V \vDash p = \phi)$$

In simpler terms, V satisfies the structural equation $p = \phi$ iff both sides of the equation have the same truth value, on the valuation specified by V. We say that a set V of literals satisfies a set M of structural equations and literals iff V satisfies each member in M. In symbols,

$$V \vDash M \text{ iff } V \vDash \gamma \text{ for each } \gamma \in M \qquad (V \vDash M)$$

These two relations of satisfaction in place, we can say what it is for a causal model $\langle M, V \rangle$ to satisfy a Boolean formula ϕ.

Definition 5 ($\langle M, V \rangle \vDash \phi$)**.** Let $\langle M, V \rangle$ be a causal model relative to $\mathcal{L}_P$. $\langle M, V \rangle \vDash \phi$ iff $V^c \vDash \phi$ for all complete sets V^c of literals such that $V \subseteq V^c$ and $V^c \vDash M$.

The definition says that ϕ is true in $\langle M, V \rangle$ iff it is true in all complete interpretations V^c that extend V and that satisfy M. For complete models, the definition boils down to $\langle M, V \rangle \vDash \phi$ iff $V \vDash \phi$ or $V \nvDash M$.

There remains to define the notion of a submodel M_I that is obtained by an intervention I on a model M.

Definition 6 (Submodel M_I)**.** Let M be a set of structural equations of the language $\mathcal{L}_P$. Let I be a consistent set of literals. M_I is a submodel of M iff:

$$M_V = \{(p = \phi) \in M \mid p \notin I \text{ and } \neg p \notin I\} \cup I$$

A submodel M_I has two types of members. First, the structural equations of M for those variables which do not occur in I. Second, the literals in I. Hence, the syntactic form of a submodel M_I differs from the one of a model M. If I is non-empty, the submodel M_I has at least one member that is not a structural equation but a literal. The satisfaction relation $V \vDash M_I$ remains nonetheless well defined. The reason is that $V \vDash \gamma$ has been defined for both a structural equation γ and an $\mathcal{L}_P$ formula.*

Holger Andreas
0000-0002-0940-7435
University of British Columbia
holger.andreas@ubc.ca

Mario Günther
0000-0001-6208-448X
LMU Munich
Australian National University
Mario.Guenther@lmu.de

References

Andreas, Holger and Günther, Mario. 2020. "Causation in Terms of Production." *Philosophical Studies* 177(6): 1565–1591, doi:10.1007/s11098-019-01275-3.

—. 2021a. "Difference-Making Causation." *The Journal of Philosophy* 118(12): 680–701, doi:10.5840/jphil20211181243.

* The authors contributed equally. We would like to thank Frank Jackson, Philip Pettit, Katie Steele, Atoosa Kasirzadeh, Cei Maslen, Alan Hájek, Phil Dowe, and Daniel Stoljar for helpful comments on this paper. We are furthermore grateful to the anonymous reviewers for *Dialectica*. We are happy for the opportunities to present parts of this paper to the Philosophy Departmental Seminar at The Australian National University, the 2019 Annual Conference of the New Zealand Association of Philosophers, and the conference *Bayesian Epistemology: Perspectives and Challenges* at the Munich Center for Mathematical Philosophy.

—. 2021b. "A Ramsey Test Analysis of Causation for Causal Models." *The British Journal for the Philosophy of Science* 72(2): 587–615, doi:10.1093/bjps/axy074.

Beckers, Sander and Vennekens, Joost. 2018. "A Principled Approach to Defining Actual Causation." *Synthese* 195(2): 835–862, doi:10.1007/s11229-016-1247-1.

Blanchard, Thomas and Schaffer, Jonathan. 2017. "Cause without Default." in *Making a Difference: Essays on the Philosophy of Causation*, edited by Helen Beebee, Christopher R. Hitchcock, and Huw Price, pp. 175–214. Oxford: Oxford University Press, doi:10.1093/oso/9780198746911.003.0010.

Collins, John David, Hall, Ned and Paul, Laurie A. 2004. "Counterfactuals and Causation: History, Problem, and Prospects." in *Causation and Counterfactuals*, edited by John David Collins, Ned Hall, and Laurie A. Paul, pp. 1–57. Cambridge, Massachusetts: The MIT Press, doi:10.7551/mitpress/1752.003.0002.

Gallow, J. Dmitri. 2021. "A Model-Invariant Theory of Causation." *The Philosophical Review* 130(1): 45–96, doi:10.1215/00318108-8699682.

Hall, Ned. 2000. "Causation and the Price of Transitivity." *The Journal of Philosophy* 97(4): 198–222, doi:10.2307/2678390.

—. 2004. "Two Concepts of Causation." in *Causation and Counterfactuals*, edited by John David Collins, Ned Hall, and Laurie A. Paul, pp. 225–276. Cambridge, Massachusetts: The MIT Press, doi:10.7551/mitpress/1752.003.0010.

—. 2007. "Structural Equations and Causation." *Philosophical Studies* 132(1): 109–136, doi:10.1007/s11098-006-9057-9.

Halpern, Joseph Y. 2000. "Axiomatizing Causal Reasoning." *Journal of Artificial Intelligence Research* 12(1): 317–337, doi:10.1613/jair.648.

—. 2008. "Defaults and Normality in Causal Structures." in *KR2008: Principles of Knowledge Representation: Proceedings of the Eleventh International Conference*, edited by Gerhard Brewka and Jérôme Lang, pp. 198–208. Washington, D.C.: Association for the Advancement of Artificial Intelligence (AAAI), https://aaai.org/papers/kr08-020-defaults-and-normality-in-causal-structures/.

—. 2015. "A Modification of the Halpern-Pearl Definition of Causality." in *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, Buenos Aires, Argentina, 25-31 July 2015*, edited by Quiang Yang and Michael J. Wooldridge, pp. 3022–3033. Menlo Park, California: The AAAI Press, https://www.ijcai.org/Proceedings/15/Papers/427.pdf.

—. 2016. *Actual Causality*. Cambridge, Massachusetts: The MIT Press, doi:10.7551/mitpress/10809.001.0001.

Halpern, Joseph Y. and Hitchcock, Christopher R. 2010. "Actual Causation and the Art of Modeling." in *Heuristics, Probability and Causality. A Tribute to Judea Pearl*, edited by Rina Dechter, Héctor Geffner, and Joseph Y. Halpern, pp. 383–406. Tributes n. 11. London: King's College Publications, https://www.cs.cornell.edu/home/halpern/papers/festschrift.pdf.

—. 2015. "Graded Causation and Defaults." *The British Journal for the Philosophy of Science* 66(2): 413–457, doi:10.1093/bjps/axt050.

Halpern, Joseph Y. and Pearl, Judea. 2005. "Causes and Explanations: A Structural-Model Approach. Part I: Causes." *The British Journal for the Philosophy of Science* 56(4): 843–887, doi:10.1093/bjps/axi147.

Hiddleston, Eric. 2005. "Causal Powers." *The British Journal for the Philosophy of Science* 56(1): 27–59, doi:10.1093/phisci/axi102.

Hitchcock, Christopher R. 2001. "The Intransitivity of Causation Revealed by Equations and Graphs." *The Journal of Philosophy* 98(6): 273–299, doi:10.2307/2678432.

—. 2007a. "Prevention, Preemption, and the Principle of Sufficient Reason." *The Philosophical Review* 116(4): 495–532, doi:10.1215/00318108-2007-012.

—. 2007b. "What's Wrong with Neuron Diagrams?" in *Causation and Explanation*, edited by Joseph Keim Campbell, Michael O'Rourke, and Harry S. Silverstein, pp. 69–92. Topics in Contemporary Philosophy n. 3. Cambridge, Massachusetts: The MIT Press, doi:10.7551/mitpress/1753.003.0006.

—. 2009. "Structural Equations and Causation: Six Counterexamples." *Philosophical Studies* 144(3): 391–401, doi:10.1007/s11098-008-9216-2.

Lewis, David. 1973. "Causation." *The Journal of Philosophy* 70(17): 556–567. Reprinted, with a postscript (Lewis 1986b), in Lewis (1986a, 159–172), doi:10.2307/2025310.

—. 1979. "Counterfactual Dependence and Time's Arrow." *Noûs* 13(4): 455–476. Reprinted, with a postscript (Lewis 1986c), in Lewis (1986a, 32–51), doi:10.2307/2215339.

—. 1986a. *Philosophical Papers, Volume 2*. Oxford: Oxford University Press, doi:10.1093/0195036468.001.0001.

—. 1986b. "Postscript to Lewis (1973)." in *Philosophical Papers, Volume 2*, pp. 172–213. Oxford: Oxford University Press, doi:10.1093/0195036468.003.0006.

—. 1986c. "Postscript to Lewis (1979)." in *Philosophical Papers, Volume 2*, pp. 52–66. Oxford: Oxford University Press, doi:10.1093/0195036468.003.0002.

—. 2000. "Causation as Influence." *The Journal of Philosophy* 97(4): 181–197, doi:10.2307/2678389.

Paul, Laurie A. 1998. "Problems with Late Preemption." *Analysis* 58(1): 48–53, doi:10.1093/analys/58.1.48.

Paul, Laurie A. and Hall, Ned. 2013. *Causation: A User's Guide*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199673445.001.0001.

Ramachandran, Murali. 1997. "A Counterfactual Analysis of Causation." *Mind* 106(422): 263–277, doi:10.1093/mind/106.422.263.

Sartorio, Carolina. 2006. "Disjunctive Causes." *The Journal of Philosophy* 103(10): 521–538, doi:10.5840/jphil20061031029.

Woodward, James F. 2003. *Making Things Happen: A Theory of Causal Explanation*. Oxford Studies in the Philosophy of Science. Oxford: Oxford University Press, doi:10.1093/0195155270.001.0001.

Yablo, Stephen. 2002. “De Facto Dependence.” *The Journal of Philosophy* 99(3): 130–148, doi:10.2307/3655640.

A Puzzle for Realism About Ground

Olla Solomyak

On the metaphysical picture that is commonly associated with theories of grounding, reality has a hierarchical structure: there are multiple "levels" of facts, with facts at the higher levels being *grounded in*, or holding *in virtue of*, those at the lower levels. My focus in this paper is on the question of what it would take for reality to truly have such a hierarchical structure. More specifically, what would it take to be a *realist* about a worldly, metaphysical relation of ground? I'll argue that there is a tension that is implicit in the notion of ground, which makes it difficult to answer this question in a straightforward way, posing a puzzle for the grounding realist. The puzzle calls standard accounts of the metaphysics of ground into question and inspires a novel alternative approach.

The idea that reality has a hierarchical structure is familiar and intuitive. For example, it's natural to think that the psychological facts in some sense depend on and arise from the biological facts, the biological facts from the chemical facts, and the chemical from the physical. We find it natural to think of some states of affairs as more basic or fundamental than others, and to explain or account for one realm of facts in terms of others that are more fundamental. The grounding theorist—particularly, the kind of *realist* about ground I'll be concerned with here—takes this hierarchical structure in a metaphysically serious way. That is, she takes reality to genuinely have such hierarchical structure, with the distinct levels of facts in this structure being related by a worldly relation of *ground*.[1] My focus in this paper is on the question of what exactly this metaphysical commitment entails. What is involved in being a realist about a worldly, metaphysical relation of ground?

1 See Fine (2001, 2012), Schaffer (2009), and Rosen (2010) for an introduction to the general notion of ground. Some approaches (such as Audi 2012) depart from this particular conception of ground, and more generally, there is a range of views as to precisely how the notion of ground should be understood. (See Correia and Schnieder 2012 for a useful overview.) My interest here is not in the details of any particular existing theory, but rather in what I take to be a very common and intuitive understanding of the notion of ground.

I should note here that many grounding theorists take grounding to be an explanatory notion that is to be treated as an operator on sentences, rather than as a worldly relation between objects or facts.[2] And one can operate with the explanatory notion of ground without being committed to a worldly relation of ground or to any particular metaphysical structure. But my interest here is in the question of what's involved in attributing grounding structure to *reality*, or what it takes for there to be genuine grounding in the world. My focus here will thus be on the worldly notion of ground. If we think there is a worldly metaphysical relation of ground that corresponds to the explanatory notion, or that we are attributing some distinctive metaphysical structure to reality when making grounding claims, we want to understand what that worldly structure must be like: So, what must the structure of reality be like to exhibit genuine grounding?

I'll argue that there is a tension that is implicit in the notion of ground which makes it difficult to answer this question in a straightforward way. The tension is revealed via the attempt to make sense of the metaphysical status of the grounded—in particular, in the attempt to explain how it is that grounded facts can be distinct from and obtain *in addition* to their grounds on the one hand, and be "nothing over and above" their grounds at the same time. I'll argue that straightforward accounts of the metaphysics of the grounded cannot satisfy both of these requirements. Either the higher-level facts are rendered too metaphysically *separate* from their grounds, or not separate enough—there appears to be no space for the metaphysical status of the grounded to be found.[3]

I'll begin in section 1 with an initial presentation of the puzzle, which reveals the tension I take to be implicit in the notion of ground in an intuitive way. I'll then make the problem more precise in section 2 and sec-

2 Correia (2010), Schnieder (2011), and Fine (2012) formulate grounding claims in terms of an operator on sentences, while Schaffer (2009), Rosen (2010), and Audi (2012) treat ground as a relation between worldly objects or facts. It's important to note that one can prefer the operator formulation and still be open to the possibility that there is a corresponding worldly relation as well, and that, on the other hand, one can speak in terms of a worldly notion of ground and still not be committed to a robust metaphysical realism about ground of the kind I am interested in here.

3 Talk of grounding has been criticized in various ways. Some theorists have doubted the coherence of the notion of ground, while others have doubted its usefulness in metaphysical inquiry. (See, e.g., Hofweber 2009; Daly 2012; Wilson 2014; and Koslicki 2015 for critiques of the notion of ground; and Raven 2017 for a defense.) My aim here, on the other hand, is ultimately not critical. Rather, it is to illuminate the structure that grounding claims implicitly attribute to reality.

tion 3, and argue that the tension cannot be resolved as straightforwardly as it may appear. After rejecting some proposed solutions which I take to be unsatisfactory, I will present my preferred approach and explain how it can accommodate the metaphysical status of the grounded. The approach I present in section 4 appeals to the notion of a *perspective*, and utilizes a meta-metaphysical framework that I have developed in application to other domains [Solomyak (2020); unpublished book manuscript]. I'll argue that making sense of the metaphysics of ground requires that we embrace multiple *perspectives on reality*—corresponding to distinct *ways* or *senses* in which a fact can be said to obtain.

1 The Metaphysics of Ground: A Tension

The puzzle for the grounding realist can be brought out by attending to two aspects of the notion of ground, which I'll call the two *requirements of ground*. I take each of these requirements to be essential to our intuitive understanding of what it is for one fact to be grounded in some further facts. But we'll see that the two requirements pull in opposite directions, and reveal a tension in the notion of ground that causes trouble for the realist who wants to attribute genuine grounding structure to reality.

The two requirements of ground are what I'll call *Distinct Obtaining* and *Nothing Further*:

> Distinct Obtaining. For any fact [A], if [A] is grounded in Γ (where Γ is a plurality of facts), then [A] is *distinct from* and obtains in *addition to* the facts in Γ.

> Nothing Further. For any fact [A], if [A] is grounded in Γ, [A]'s obtaining is *nothing over and above* the obtaining of the facts in Γ.

Let's start with Distinct Obtaining. First, why take [A] to be distinct from its grounds? To begin with, it's important to note that on the above formulation (as on a very common conception of ground) [A] is a single fact, while its grounds is a plurality of facts.[4] So we couldn't generally take [A] to be identical to its complete grounds. And there are reasons to think that [A] must be distinct

4 See Rosen (2010), Fine (2001, 2012), and Correia (2010), though both Fine and Correia treat ground as an operator on sentences rather than as a relation between facts. I'm concerned here with what Fine calls *full* (as opposed to *partial*) ground.

doi: 10.48106/dial.v76.i1

from each individual fact in the collection of its grounds as well: First, the grounded fact might have a structure or involve objects that aren't involved in the individual facts that contribute to its grounds—a reason to think that the grounded fact is at least sometimes distinct from each of its (partial) grounds. But more generally, the facts in Γ are supposed to *explain* the obtaining of [A], or be an account of what *makes it the case* that A, and it's not clear how a fact could explain or account for its own obtaining. If [A] appeared in its own grounds, the corresponding grounding explanation would strike us as circular. For these and other reasons, grounding is commonly taken to be irreflexive, and a grounded fact is taken to be distinct from any (and all) of the facts that ground it.[5]

Ground is also typically taken to be a factive notion.[6] And on the factive notion of ground, [A]'s being grounded in Γ implies that [A] *obtains*. More generally, if some facts (which themselves obtain) *make it the case* that A, then it must *be the case* that A. It thus appears obvious and uncontroversial that grounded facts, in addition to their grounds, obtain. Distinct Obtaining, then, is implicit in a very common and intuitive understanding of what grounding involves.

Let's now consider Nothing Further. Nothing Further is an expression of the *tightness* of the metaphysical connection that is supposed to hold between a grounded fact and its grounds. Unlike weaker relations such as supervenience, where one realm of facts can supervene on another while having a kind of metaphysically independent reality of its own, grounded facts can have no such metaphysically independent reality: the grounded facts are *nothing over and above* their grounds.[7,8] Of course, this "nothing over and

5 See, e.g., Schaffer (2009), Rosen (2010), and Correia (2010). Fine (2012) distinguishes between weak and strict ground, where strict ground is the irreflexive notion I'm interested in here.

6 See, e.g., Audi (2012). Fine (2012) distinguishes between a factive and a non-factive notion of ground, and argues that the factive notion is more fundamental. This is the notion I operate with here; it is thus assumed that the facts in Γ obtain.

7 Though see Audi (2012) who explicitly rejects this requirement of ground. I'll return to discuss the costs of such a stance further on.

8 Fine (2012) explains that there can be no *explanatory* gap between the grounded and its grounds, which is not to say that the grounded cannot be real in its own right. In fact, Fine (2001) allows for grounded facts and their grounds to be equally real (though some grounded facts may also be unreal). But for a realist about ground in the sense I'm concerned with here, there must be some structure in reality that underwrites explanatory grounding claims, and thus the maximal explanatory tightness of ground is seen as reflecting an equally tight metaphysical connection between the corresponding aspects of reality. I'm interested here in making sense of such a metaphysically tight connection.

above" may not be analyzable in terms that don't ultimately appeal to the notion of ground.[9] But the intuitive thought behind NOTHING FURTHER is that the grounded doesn't carry any additional "metaphysical weight" once its grounds are in place. Given the grounds, nothing additional is required, metaphysically speaking, for the grounded facts to obtain. It is sometimes said that for a grounded fact to obtain *just is* for its grounds to obtain. For example, assuming that the existence of a table is grounded in the existence and arrangement of certain particles, we might say that for there to be a table *just is* for there to be particles arranged in this particular way.

But the "just is" in the above formulation is not meant to be the "just is" of identity—as we noted above, the intuitive notion of ground does not take the grounded to be identical to its grounds. And this is precisely where the worry for the grounding realist arises: What is the "just is" that figures in grounding claims that's metaphysically *tight* enough to satisfy the requirement of NOTHING FURTHER, while still being short of identity? In other words, where is the space between "obtaining in addition to" and "obtaining over and above" for the metaphysical status of the grounded to be found? While the two requirements we've spelled out are both aspects of a single seemingly coherent notion of ground, they appear to pull in opposite directions: How can the grounded facts be *distinct from* and obtain *in addition to* their grounds on the one hand, and yet be "nothing over and above" their grounds at the same time?

One might think that simply accepting the notion of ground as a metaphysical primitive which exhibits the features described here is enough to allay the tension I've described: Perhaps it is precisely in being *grounded* that the metaphysical status of the higher-level facts meets both of the requirements above.[10] The sense in which the higher-level facts are nothing over and above their grounds while still being distinct from them is in that the former are grounded in the latter. Ground is the primitive notion that exhibits the maximal-metaphysical-tightness-just-short-of-identity which we were after in the discussion above.

9 This phrase is sometimes used in other contexts, where it may be analyzable in terms that don't appeal to the notion of ground; but we shouldn't expect the sense in which the *grounded* is nothing over and above its grounds to be analyzable in other terms if the notion of ground is taken as primitive.

10 See, e.g., Schaffer (2009), Rosen (2010), and Fine (2012) on taking ground as a metaphysical primitive.

doi: 10.48106/dial.v76.i1

But this response avoids the real question at issue. Even for one who embraces ground as a metaphysical primitive, the question arises as to what *structural features* realist grounding claims implicitly attribute to reality. In what sense is a reality structured by a worldly relation of ground genuinely *hierarchical*? In what follows, I'll argue that the apparent tension introduced above does in fact pose a real challenge for the grounding realist—one that taking the notion ground as a metaphysical primitive does not, in itself, suffice to meet. We'll see that accommodating both requirements of ground in an account of reality's structure is far from straightforward, and this will ultimately push us to reconceptualize our understanding of the metaphysics of ground—clarifying what it *is* to attribute primitive grounding structure to reality.

2 Two Pictures of Reality

Let's begin by considering two different pictures of what the structure of reality might be like.[11] On the one hand, we have what I'll call the *Ground-Level Picture (GLP)*: On this picture, the ground-level, fundamental facts are ultimately all there is to reality—only the ground-level facts *really obtain*. Reality is thus ultimately "flat" rather than hierarchical on this picture—there is just one *real* level of facts.[12] On the other hand, we have what I'll call the *Hierarchical Picture (HP)*: On this picture, reality consists of both fundamental and non-fundamental facts. The non-fundamental facts really obtain, just as and in addition to their grounds. Reality thus has a hierarchical structure, with multiple "levels" of facts.

Of course, there is a question about how the uses of "really" and "ultimately" here should be understood. This is an issue we will return to shortly, and which will be central in the discussion that follows. For the time being, I want to appeal to an intuitive understanding of these locutions, as well as of the

11 In what follows, I will focus on a picture which assumes that there is a ground-level of fundamental facts, and that all of the higher-level facts are ultimately grounded in this fundamental level. This is not a picture that all grounding theorists will accept—one might think that not all grounding explanations bottom out at a fundamental level, or that there is no absolutely fundamental level at all. (See, e.g., Raven 2016.) One might also think that a fact can be both grounded and fundamental on a positive conception of fundamentality. For now, I'm going to set these views aside because I want to focus on the simplest and most straightforward picture of what a hierarchically structured reality might be like, and ask what it is—even in this simplest case—for reality to be genuinely hierarchical.

12 See, e.g., Bennett (2011), who discusses a variant of this view.

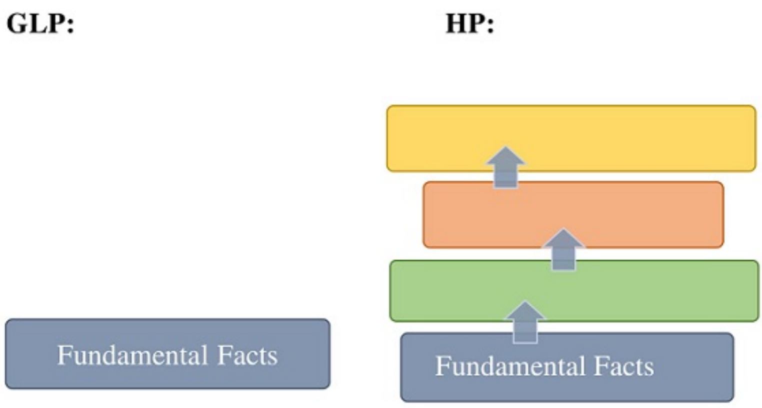


Figure 1: Two Pictures of Reality

difference between the two pictures sketched above: We have a flat, ground-level-only reality on the one hand, and a hierarchically structured, multi-level reality on the other. We have an initial, intuitive understanding of what each of these pictures amounts to, and getting a more precise understanding of the commitments they involve will be our central aim in the discussion below.

What should the realist about ground say about these two pictures of reality? Must she be committed to one of these pictures over the other? At first approach, the Ground-Level Picture appears incomplete: The fundamental facts are not all the facts; the non-fundamental facts obtain as well. Given Distinct Obtaining—the first requirement of ground introduced above—the grounded facts are distinct from and obtain in addition to the facts that ground them. Further, the grounded facts are supposed to be those we "get for free," or that "automatically arise," once the ground-level facts are in place. All this seems to point against the Ground-Level Picture, and in favor of the Hierarchical Picture instead.

More generally, the Hierarchical Picture just seems to come along with the notion of ground, or more specifically, with realism about grounding as a worldly relation. And it's a picture that is often presented as the grounding theorist's background picture of the structure of reality. Here, for example, is Jonathan Schaffer's description of the grounding theorist's background theory:

> [T]he neo-Aristotelian will begin from a *hierarchical view of reality* ordered by *priority in nature*. The primary entities form the sparse structure of being, while the grounding relations generate an abundant superstructure of posterior entities. (2009, 351)

So the realist about ground, who thinks of grounding as a worldly relation that structures reality, at least initially appears to be committed to the Hierarchical Picture.

But essential to this presentation of the two pictures is a non-trivial assumption about the difference between them. Implicit in the presentation of the two pictures as reflecting two distinct structures that reality might have, and in the conclusion that the realist about ground is committed to rejecting the GLP in favor of the HP, is what I'll call the *Assumption of Substantive Difference*. This is the assumption that the GLP and the HP *substantively differ* with respect to the structure they attribute to reality. More specifically, it is the assumption that, according to the HP, reality has a metaphysical structure that is lacking according to the GLP.

Intuitively, this seems to be a very natural assumption. Whether reality is flat or hierarchically structured, and whether there are non-fundamental facts in addition to the fundamental facts, seem to be substantive questions about the structure of reality. Is the flat, Ground-Level Picture a *complete* picture of reality, or is it missing some of the facts that genuinely obtain? Are there distinct "levels" of facts? The GLP and the HP disagree over these questions and thus appear to reflect two distinct ways in which reality might be structured. And as we saw, the realist about ground appears to be committed to the Hierarchical Picture *over* the Ground-Level Picture, in maintaining that reality genuinely has the hierarchical structure that's implicit in grounding claims.

But as we'll see, this stance leads to a difficulty for the grounding theorist. In particular, we'll see that the commitment to the HP, *with the Assumption of Substantive Difference in the background*, conflicts with Nothing Further—one of our original requirements of ground.

Recall that Nothing Further is the requirement that the obtaining of grounded facts be "nothing over and above" the obtaining of their grounds.

Of course, we don't have a precise way of cashing out how this "nothing over and above" should be understood, but we do have an intuitive sense of what it amounts to, and in what follows, I'll argue that there is a real conflict between the intuitive "nothing over and above" that we require of the grounded and the assumption of SUBSTANTIVE DIFFERENCE. It will thus turn out that favoring the Hierarchical Picture as the more accurate reflection of reality's structure, as we think the grounding realist should, conflicts with an essential aspect of our intuitive notion of ground.

In the next section, I'll consider some attempts to maintain the assumption of SUBSTANTIVE DIFFERENCE on behalf of the grounding realist and show why they are bound to be unsuccessful—arguing that NOTHING FURTHER and the assumption of SUBSTANTIVE DIFFERENCE are in fact in conflict, as I've claimed. I'll then briefly consider some further responses to the puzzle, which I take to be unsatisfactory, and finally, in section 4, present an alternative approach. There, I will argue that the assumption of SUBSTANTIVE DIFFERENCE is what must, after all, be given up. But *realism* about ground must thereby be significantly reconceptualized.

3 Maintaining the Hierarchical Picture

To see how the difficulty for the proponent of the Hierarchical Picture arises, we must ask ourselves what exactly is *missing* from the Ground-Level Picture that makes it incomplete, and which needs to be added in order to get the Hierarchical Picture. The assumption of SUBSTANTIVE DIFFERENCE says that there is something about reality that the Ground-Level Picture is missing out on, and the question is what that something could be. What could possibly be added to the GLP to get the HP, given that the obtaining of the higher-level facts is supposed to be *nothing over and above* the obtaining of the facts that ground them?

One might think that there is a trivial answer to this question: What needs to be added to the Ground-Level Picture is just *all the higher-level facts*. These facts obtain, and they are missing from the Ground-Level Picture. The problem with this response is that a proponent of the Ground-Level Picture will not deny that the higher-level facts obtain in a trivial sense. That is, someone who thinks that reality is ultimately exhausted by the fundamental facts does not deny that there is also an ordinary sense in which, e.g., there are tables. The GL-theorist simply draws a distinction between the ordinary, trivial sense in which there are tables, and the further, somehow "metaphysically

doi: 10.48106/dial.v76.i1

loaded" claim that there are *really* tables, or that this is a genuine fact of reality. For the GL-theorist, only the fundamental facts *really obtain* in this metaphysically loaded sense, but there is an ordinary, trivial sense in which the non-fundamental facts obtain as well.

This distinction relies on a metaphysically substantive notion of *reality*, or of what it is for a fact to *really obtain*. There are various ways in which such a notion might be cashed out, but to focus on one example, we might consider Kit Fine's (2001) distinction between something's being merely *the case* and something's being the case *in reality*, where the latter has a metaphysical weightiness that the former does not. The GL-theorist might thus appeal to this distinction, and maintain that the non-fundamental facts obtain, but don't obtain "in reality." The GLP is a picture of what things are like *in reality*, and thus only includes the fundamental facts—the only facts that obtain in reality according to the GL-theorist.

The real difference between the Ground-Level and Hierarchical pictures, then, must be construed in terms of some further, non-trivial, metaphysical commitment—beyond the question of whether there are tables in an ordinary sense. The pictures disagree as to whether such non-fundamental facts *really obtain* in a metaphysically weighty sense—i.e., whether they obtain *in reality*. The substantive difference between the Ground-Level and Hierarchical Picture thus concerns the metaphysical *status* of the non-fundamental facts, not their mere obtaining in the most basic, trivial sense.

But on this understanding of the Hierarchical Picture, the obtaining of the higher-level facts appears to be something "over and above" the obtaining of the fundamental facts after all: If the metaphysical weightiness of the claim that the non-fundamental facts really obtain is something that could in principle be *lacking* even after the ground-level facts are in place, then the *real obtaining* of the higher-level facts consists in something more than the obtaining of their grounds. And this conflicts with Nothing Further: the requirement that the obtaining of the grounded facts be nothing over and above the obtaining of the facts that ground them.

It appears that once the ground-level facts are in place, and we agree that those really obtain, there can be *nothing further at stake* in the question of whether the higher-level facts really obtain as well. As soon as we admit that there's a further metaphysical, or even meta-metaphysical, question to ask, an affirmative answer seems to grant the higher-level facts too much independent weight—the "real obtaining" of the non-fundamental facts in this metaphysically weighty sense is then something over and above the real obtaining of their grounds.

The grounding theorist might resist this line of thought by insisting that it doesn't *cost* anything to posit the non-fundamental facts once the ground-level facts are in place—to say the non-fundamental facts are grounded in the fundamental facts is to say that we "get them for free" once we posit the obtaining of the fundamental. But given what I've argued here, it turns out that maintaining the commitment to NOTHING FURTHER requires more than this. It is not enough to say that we get the non-fundamental facts "for free"; rather, there must be *no genuine difference*, as far as reality is concerned, between a picture that includes them and a picture that doesn't. There can be no metaphysically better answer to the question of whether they *really obtain*.

It appears, then, that the grounding theorist cannot maintain the assumption of SUBSTANTIVE DIFFERENCE by appealing to the higher-level facts, given her commitment to NOTHING FURTHER. Might the grounding theorist nevertheless maintain the assumption of SUBSTANTIVE DIFFERENCE in some other way? That is, might something other than the obtaining of the higher-level facts be what constitutes the substantive difference between the GLP and the HP? I'll briefly consider two other possibilities.

One might attempt to appeal to facts about *what grounds what* in attempting to account for the difference between the two pictures. Perhaps it is not the higher-level facts themselves, but facts about *how they are grounded* that make for the substantive difference between the GLP and the HP. There are several ways in which this response could be developed, depending on how one understands the nature and status of these facts about ground, but as we'll see, they all fall short for what is fundamentally the same reason.

On one view (e.g., as Bennett 2011; and deRosset 2013 maintain), facts about ground are themselves grounded. On such a picture, appealing to facts about ground would be of no help to the grounding theorist: if the facts about ground are among the higher-level facts, they are simply a subset of those grounded facts we considered in the discussion above, which we argued could have no metaphysical weight over and above what is already in the GLP. The facts about ground—like any other higher-level facts—could not be the source for the substantive difference between the GLP and the HP given that their obtaining (like that of all higher-level facts) can be nothing over and above the obtaining of their grounds. Alternatively, one might consider a view on

doi: 10.48106/dial.v76.i1

which the (or at least some) facts about ground are fundamental.[13] But on such a view, the facts about ground would already be included in the GLP and thus could not be missing from it. They could not thus constitute the substantive difference between the GLP and the HP.

A third possibility for treating the facts about ground is suggested by Shamik Dasgupta (2014), who argues that certain facts about how things are grounded are neither fundamental nor grounded, but rather, form a third category Dasgupta calls "not apt for grounding." At first glance, such facts might be seen as a promising candidate for identifying the substantive difference between the GLP and the HP. But closer attention to this proposal reveals it to be unsatisfactory as well. We can define the GLP+ as the Ground-Level Picture together with the facts about ground, and then ask whether there is still something missing from the resulting picture. If the grounding theorist says there is something missing, we are back to our original question of what this could possibly be, given that the higher-level facts can be nothing over and above the facts that ground them. And if the grounding theorist says there is nothing missing, she appears to have rejected the Hierarchical Picture, which includes the higher-level facts as well as those at the ground-level.[14] Irrespective of how we treat the facts about ground then, the puzzle for the grounding theorist remains: There is a tension between the grounding theorist's commitment to the Ground-Level Picture *over* the Hierarchical Picture on the one hand, and the commitment to Nothing Further on the other.

More fundamentally, the problem with this series of proposals is as follows: The facts about ground—no matter where they are to be found in the grounding theorist's metaphysical picture—concern the obtainment of the *grounding relation* between the higher-level facts and their grounds. Our central question has been how the metaphysical structure of this relation should be understood: What structure do we attribute to reality in making grounding

13 This would be to reject what Sider (2011) calls *purity* of the fundamental, but is nevertheless an option for the grounding theorist.

14 One might argue that there is a sense in which the higher-level facts wouldn't be missing from such a picture. Perhaps the inclusion of the facts about ground in the GLP simply "bring out" the fact that the higher-level facts must also be there. A picture along these lines is suggested by Glazier (2016), who argues that what he calls fundamental metaphysical laws can explain why there are any non-fundamental facts at all. But such a picture is not obviously committed to the HP *over* the GLP; on the contrary, it is a way of maintaining that the GLP may not be missing anything for the grounding realist, or alternatively, of rejecting the assumption that there is a substantive difference between the GLP and the HP. It does not then help the grounding theorist uphold the assumption of Substantive Difference as the response here was aiming for.

claims? To say that grounding structure is accounted for by *facts* about that very structure is to get things upside down. These facts *track* the structure we are after rather than bring it about.

The grounding realist might attempt another avenue of response, and claim that the Hierarchical Picture is more accurate than the GLP *holistically speaking*, simply because it describes reality as layered, or hierarchical, rather than flat: the grounding relation is real, and genuinely structures reality. But this just sweeps the central question under the rug: What is it that makes the former a more accurate description of reality? If what makes it more accurate is the addition of the higher-level facts, we are left with the problem we encountered above—how could the addition of the higher-level facts make for a substantive difference, if their obtaining is nothing over and above the obtaining of their grounds? And if the Hierarchical Picture is more accurate for some other reason, the grounding theorist owes us an explanation of what that reason could be. The question we've been concerned with all along is the question of what *is* it for reality to have a hierarchical structure organized by the relation of ground. We've seen that there is a challenge to making sense of this position, and to simply restate the position does not help resolve it. How could the HP be a *more accurate* description of reality than the GLP given the requirement of Nothing Further? Absent some further theory or explanation from the grounding theorist, it is not clear how the Hierarchical Picture can be maintained, given the requirements of ground we started out with.

One might consider rejecting Nothing Further in favor of a weaker requirement on which it is allowed that grounded facts have metaphysical weight over and above that of their grounds. This would be to embrace a distinction between what we might call *metaphysical cost* and *metaphysical weight*, and to maintain that while it doesn't *cost* anything, metaphysically speaking, for the higher-level facts to arise, their arising *is* in fact something over and above the obtaining of their grounds. The higher-level facts on such a picture would have a metaphysical *weight* of their own, in that their *obtaining* would be something over and above that of their grounds. Nevertheless, one would insist that this obtaining, or extra metaphysical weight, was of no metaphysical *cost*—the obtaining of the relevant grounds being all it *takes* for the additional layers of reality to arise.

This is perhaps closer to the Aristotelian picture of a hierarchically structured reality with multiple genuine "levels," though I think it departs from a contemporary and very intuitive conception of ground, on which the ground-

doi: 10.48106/dial.v76.i1

ing relation is supposed to be maximally "metaphysically tight." But more importantly, I am suspicious of the conceptual move of distinguishing between metaphysical cost and metaphysical weight in this way. Unlike a buy-one-get-one-free deal at the supermarket—where one can get more "weight" than the "cost" one has paid—*metaphysical* weight cannot come free of metaphysical cost. Metaphysical cost and weight cannot, in principle, come apart: What it *takes* (metaphysically speaking) for a fact to obtain and what it *is* (metaphysically speaking) for a fact to obtain are one and the same. Any additional metaphysical weight that could in principle be *lacking* given the obtaining of the ground-level facts is thereby something it *takes* for the higher-level facts to obtain. The relevant notion of cost here is not causal, but metaphysical, and so there can be no metaphysical state or status that is lacking once the assumed metaphysical costs are in place.

The thought that one can maintain the Hierarchical Picture as a grounding theorist seems to implicitly rely on the mistaken assumption that the notions of metaphysical cost and metaphysical weight are separable: One says in the same breath that the grounded facts arise "for free" but also that they really do *arise* and thus constitute a genuinely hierarchical reality. As I've argued above, this is to grant the higher-level facts their own metaphysical weight, and thus requires that their weight can be separated from their cost. But closer attention to these notions once the distinction is made explicit reveals that there is no space for them to come apart. Metaphysical cost includes *everything it takes* for a fact to obtain, and this includes any metaphysical weight this obtaining may involve.

So I think Nothing Further is essential to the notion of ground, and should not be weakened or given up. But more importantly for our purposes, it's clear that Nothing Further is essential to a very common and intuitive conception of ground, and I'm interested here in how we can make sense of realism about this particular notion.

To recap then, the puzzle for the grounding realist arises as follows: Distinct Obtaining, as well as general considerations surrounding the notion of ground, push against the Ground-Level Picture *in favor* of the Hierarchical Picture, implicitly endorsing the assumption of Substantive Difference. But this stance is incompatible with Nothing Further, as Nothing Further implies that there can be nothing of metaphysical substance at stake in this move.

As I've argued, both Distinct Obtaining and Nothing Further are essential to an intuitive and very common conception of ground. Our only real option then seems to be to reject the assumption of SUBSTANTIVE DIFFERENCE—that the Ground-Level and Hierarchical pictures reflect two genuinely distinct structures reality might have. But this seems not only counterintuitive in itself, but also counter to the intuitive notion of *ground*, which, as we've seen, naturally comes along with a Hierarchical Picture of the structure of reality.

In what follows, I'll argue that the assumption of SUBSTANTIVE DIFFERENCE is nevertheless what the grounding theorist should give up. I'll present a resolution to the puzzle on which hierarchical structure is understood to be compatible with—in fact, even dependent on—the rejection of SUBSTANTIVE DIFFERENCE.

4 The Perspectives Approach

At this point, one might find it natural to react to the line of argument I've put forward as follows: Perhaps we've simply misinterpreted both the Ground-Level and Hierarchical Pictures to begin with in thinking that they could be pitted against each other; the assumption of SUBSTANTIVE DIFFERENCE should have been rejected from the start. Just as we concluded that the Ground-Level Picture shouldn't be taken to *deny* the higher-level facts in a trivial sense, the Hierarchical Picture shouldn't be interpreted as granting them any metaphysical weight of their own. The vertical arrangement of the facts in the hierarchy, as well as the arrows going from one level to the next, are precisely meant to convey that the relation between the lower and higher-level facts is one of *grounding*, and that the higher-level facts are *nothing over and above their grounds*.

I think this line of thought is correct, but its implications must be fully appreciated. To embrace this line of thought, one must admit that each picture taken on its own is deeply misleading. The Ground-Level and Hierarchical Picture turn out to be *interdependent*: in order to interpret each picture correctly, we must have the other picture in the back of our minds. To interpret the Ground-Level Picture correctly, we must see the non-fundamental as implicitly arising from the fundamental, as depicted in the Hierarchical Picture; and to interpret the Hierarchical Picture correctly, we must see the non-fundamental facts as not really anything over and above what's already present in the Ground-Level Picture. Both pictures are thus essential to our conception of a reality that's structured by the relation of ground.

doi: 10.48106/dial.v76.i1

This is not to say that the grounding theorist cannot speak of a hierarchically structured reality, or that reality cannot truly *be* hierarchically structured. But it turns out that what it is for reality to be hierarchically structured (on this understanding) is not at all what we would have thought. To say that reality is hierarchically structured is not to say that the Hierarchical Picture (as defined above) is the picture that best reflects reality's structure. A truly hierarchical structure is, paradoxically, one that is best reflected by *both* the Hierarchical and the Ground-Level Pictures taken hand-in-hand—the hierarchical aspect of the multi-level structure is only guaranteed by the fact that the flat, ground level of the hierarchy is in some real sense all there is to the "hierarchy" at all.

In what follows, I want to suggest a way of conceptualizing this seemingly paradoxical state of affairs—one that I think can help clarify the notion of ground and further illuminate what the nature of a hierarchically structured reality must be like. The approach relies on the notion of a *perspective*, and distinguishes between two perspectives on reality implicit in the notion of ground. On the one hand, there is what I'll call the *ground-level perspective*, from which reality is exhausted by the fundamental, ground-level facts. On the other hand, there is what I'll call the *hierarchical perspective*, from which reality extends beyond the fundamental to encompass the higher-level, non-fundamental facts as well. As I've already suggested, these two perspectives are each essential to the notion of ground, as well as to a reality that is genuinely structured by a worldly grounding relation. In what follows, I'll present the background framework of perspectives in more detail and then return to explain how this approach can provide a satisfying resolution to our puzzle.

4.1 *The Perspectives Framework*

The notion of a perspective that I appeal to here will remain undefined. But an intuitive gloss and a few examples will help bring out the particular notion of a perspective that I have in mind. To get an initial sense of the relevant notion, we can consider the familiar shift from seeing reality from a first-person, or subjective, perspective to seeing or conceptualizing reality in an impersonal, or objective, way. Imagine an extreme solipsist who is not aware that there is a reality beyond her own subjective experience at all. Such a solipsist implicitly identifies her own experience with the *whole of reality*; for this solipsist, there is no distinction between something's being the case *in her experience* and something's being *the case, full stop*. Implicitly first-personal claims such as "it's painful" or "it's pleasant" will have absolute truth-values for this solipsist;

from her perspective, how things are *in reality* and how things are *in her experience* are one and the same.

This strong identification of one's own experience with the whole of reality is what I call the *first-personal perspective*, and can be contrasted with the broader *impersonal perspective*, from which one recognizes that reality extends beyond one's own experience to include other subjects and/or objective states of affairs. From the impersonal perspective, a distinction is drawn between something's being the case *in one's own experience* and something's being *the case, full stop*. First-personal claims such as "it's painful" will (from this perspective) be incomplete without reference to a subject—things can be painful for one subject but not for another, and more broadly, how things are *in one's experience* and how things are *in reality* can come apart.

Crucially, the shift from the first-personal to the impersonal perspective involves a change in one's conception of *reality*, and more specifically, one's conception of what we might call the "shape" of reality: from the first-personal perspective, reality is implicitly taken to be first-personal, while the impersonal perspective takes reality to be broader and "impersonal" in shape.[15]

We can similarly identify distinct perspectives on reality we might adopt in thinking about the metaphysics of time. On the one hand, there is the perspective of the present, or the *present-tensed perspective*. From this perspective, one identifies the present with the *whole of reality*. Crucially, this is not just an ontological stance. Rather, it is a more general conception of *reality*, which also includes a conception of what it is for something to be the case in reality, or of what it is for a fact to obtain. From the present-tensed perspective as I understand it, for something to *be the case* and for something to *be the case now* are one and the same. There is no metaphysical distinction to be drawn between something's being the case in reality and something's being the case in the present.

On the other hand, we can shift to the broader *atemporal perspective*, from which reality is seen as extending beyond the present to encompass other times and/or atemporal states of affairs. Again, this is not just a matter of ontology. From the atemporal perspective, one's conception of *reality* allows for something to be the case in reality, but not in the present—e.g., something can be the case *at another time*, or just *independently* of how things are in the present. And here, just as in the first-person case, we can see ourselves

15 Compare to Fine (2005), who describes the issues of realism about tense and the first-personal analogue of perspective as concerning the "form" of reality.

as shifting from one of these perspectives to the other, and as shifting from a narrow to a broader *conception of reality* when we make that move.

In each case, we can ask whether one perspective or the other is more fundamental or metaphysically privileged *as a perspective on reality*. That is, we can ask whether it is the narrow or broad conception of reality in each case that is getting the shape and structure of reality "right." Is reality ultimately first-personal or impersonal? Present-tensed or atemporal? The question of which, if either, perspective in each case is fundamental is a way of getting at this question about reality's structure.

More generally, then, a perspective is a way of conceptualizing *all of reality*, and comes along with a corresponding conception of what it is for something to be the case in reality, or of what it is for a fact to obtain. That is, in adopting a perspective, one identifies reality in a certain way, which allows one to answer certain questions about what it is for a fact to obtain, as well as about *which* facts obtain, just by virtue of one's conception of the metaphysical "shape" of reality. For example, in taking reality to be present-tensed, as one does from the present-tensed perspective, one thereby rules out non-present facts as well as objects from one's ontology—one's identification of reality with the present entails that what it is for a fact to obtain and what it is for a fact to obtain *now* are one and the same, and thus that obtaining at another time, or obtaining independently of time, are not ways of obtaining *in reality*. A perspective is thus associated with a certain answer to the question of what it is to be *real*, or of what it is to *obtain in reality*, from which ontological and other metaphysical commitments follow downstream.

In taking a perspective to be fundamental, one takes that perspective to be metaphysically privileged in its reflection of reality's "shape." A fundamental perspective is one that identifies the shape of reality correctly—such that what it is to *be the case from that perspective* and what it is to *be the case, full stop*, are one and the same. In other words, the "way of obtaining" that this perspective identifies with "obtaining in reality" correctly reflects what it is for a fact to *obtain*.

With this brief introduction to the perspectives framework in hand, we can return to the case of ground, and see how we might reconceptualize the surrounding metaphysical issues. As I'll argue, adopting the language of perspectives allows us to resolve the puzzle for the grounding realist in an intuitively satisfying way.

4.2 Perspectives and Ground

Turning back now to the grounding case, we can see the Ground-Level and Hierarchical Pictures as corresponding to two distinct perspectives in the sense introduced above. The *ground-level perspective* is the perspective from which the fundamental level is identified with the *whole of reality*. From this perspective, for something to *be the case* and for something to *be the case fundamentally* are one and the same; no metaphysical distinction is drawn between fundamental reality and reality. On the other hand, we can shift to the broader *hierarchical perspective*, from which reality is seen as extending beyond the fundamental to encompass the non-fundamental as well. From this perspective, for something to be the case and for something to be the case fundamentally are not one and the same; something can *be the case*, but not fundamentally.

As in the first-person and temporal cases, we can see ourselves as naturally shifting from one of these perspectives to the other: We can conceive of reality as exhausted by the fundamental and then broaden our conception of reality to include the non-fundamental as well. Importantly, this is not just a matter of "adding facts" into our picture of reality; rather, it involves shifting our *conception* of reality from a narrow sense to a broader one. We can grasp a *sense* of "reality" on which reality just *is* fundamental reality—to be real and to be fundamental are (on this conception) one and the same. On the other hand, we can also grasp a broader sense of "reality," on which reality and fundamentality can come apart. Each of these perspectives, or conceptions of reality, thus comes along with a corresponding conception of what it is for a fact to really *obtain*: From the ground-level perspective, for a fact to really obtain is for it to obtain fundamentally, while from the hierarchical perspective, a fact can *really*, but not fundamentally, obtain.

We can now understand the move from the Ground-Level to the Hierarchical Picture in a new way. Rather than holding a fixed conception of reality and positing it to include additional facts, we are shifting our *conception* of reality. The GLP understands what it *is* to really obtain in one way, while the HP understands it in another. This shift in the way reality is identified automatically gives rise to "more facts" in the Hierarchical Picture than are present in the Ground-Level Picture—but this is because the criteria for being a "real fact" have been changed, not because these facts have been granted a heftier metaphysical status. The two pictures (now understood to be perspectives)

doi: 10.48106/dial.v76.i1

thus correspond to two different ways, or *senses*, in which facts can be said to obtain.

The disagreement between the Ground-Level and Hierarchical Pictures, when seen as pitted against each other, can be seen as a disagreement about which of these two perspectives is fundamental or metaphysically privileged in its identification of the "shape" of reality: The GL-theorist holds that a conception of reality on which reality is identified with the ground-level is what best captures reality's structure, while the proponent of the Hierarchical Picture takes the broader conception of reality to be fundamental—i.e., metaphysically privileged in its reflection of what it is to be real.[16]

But as the challenge I've raised for the grounding theorist illustrates, both perspectives are essential to accommodating the requirements of ground: Distinct Obtaining, as well as the general thought that reality has a hierarchical structure, requires that we adopt the hierarchical perspective, and recognize a sense in which the non-fundamental facts really *obtain*. Nothing Further, on the other hand, requires that we adopt the ground-level perspective, and recognize that anything "beyond" the ground-level is really nothing at all—i.e., nothing over and above what is already there at the ground-level.

Recognizing this way in which we implicitly adopt and shift perspectives is the key to resolving the puzzle for the realist about ground. For the grounding realist, the GLP and the HP are not to be seen as distinct ways in which reality might be structured; rather, they are to be seen as reflecting the two distinct perspectives that are both implicit in and essential to the notion of ground. Distinct Obtaining and Nothing Further are satisfied via these two distinct, but interdependent perspectives: Distinct Obtaining is satisfied by the fact that the hierarchical perspective is a "genuine perspective on reality"—i.e., that there is a real sense in which both fundamental and higher-level facts really obtain. Nothing Further, on the other hand, is satisfied by the fact that the ground-level perspective is a "genuine perspective on reality" as well—i.e., by the fact that there is a real sense in which reality is exhausted by the fundamental.

It is important to distinguish between a number of different claims that embracing this approach could involve. First, there is the weaker, conceptual claim that both perspectives are essential to our *grasp* of the notion of ground. That is, we might say that our grasp of the notion of ground relies on our

16 I'm using "fundamental" in a new sense here: a perspective may be fundamental *qua* perspective even if it's not a perspective that "sees" only the fundamental *level* of reality. I'll return to this issue below.

ability to adopt both of these perspectives and shift back and forth between them, grasping two distinct senses of *reality* as we make that move. As I've said, I think we do in fact implicitly shift perspectives in this way, and making this explicit can help make sense of our conflicting intuitions in this area: To see the higher-level facts as genuinely *grounded* in the fundamental, we must think of them as *really obtaining* in one sense and as nothing "beyond" what *really obtains* in another.

So I think the conceptual claim goes some way towards clarifying the issues surrounding the notion of ground. But our central challenge has been the metaphysical question of what it would take for reality to genuinely *have* the kind of hierarchical structure one commits to by being a realist about ground. And this brings us to the stronger, metaphysical claim which the grounding theorist might be pushed to embrace: namely, that the two perspectives in this case are not only essential to the *concept* of ground, but also to the *metaphysical structure* of a reality that exhibits genuine grounding. More specifically, this would entail a commitment to a kind of pluralism about *reality*, on which there are genuinely two distinct *ways* in which a fact can be said to obtain. For such a pluralist, there would be no univocal answer to the question of whether the higher-level facts really obtain: only the fundamental facts really obtain in one sense, while both fundamental and non-fundamental facts really obtain in another.[17]

Importantly, this pluralist stance must be distinguished from a kind of semantic pluralism, on which the GLP and the HP simply reflect two different ways of speaking about the same reality, with no further fact of the matter as to which of the associated perspectives is fundamental. Embracing something like quantifier variance, or different—equally good—senses of the word "fact" or "obtains" would be other ways of rejecting the assumption of SUBSTANTIVE DIFFERENCE, but to embrace such a stance would be to give up on the robust metaphysical realism about ground that is of essential interest to us here:[18] a stance on which there is no metaphysically privileged way of answering the question of how reality is truly structured is thereby also an anti-realism about ground as a relation that genuinely structures reality.[19]

17 This view can be compared to ontological pluralism of the kind that is defended by McDaniel (2009) and Turner (2010), on which there are multiple ways or senses in which objects *exist*.

18 See, e.g., Hirsch (2002) for such an approach to ontology.

19 See, e.g., Sider (2009) for a way of conceptualizing the kind of metaphysical realism I take to be in the background here.

doi: 10.48106/dial.v76.i1

The kind of pluralism I present here is thus more radical than it appears, and faces a number of conceptual and metaphysical difficulties, but I think it is in some ways best suited to reflect the commitments of the realist about ground. In what follows, I'll sketch the proposal in a bit more detail, and explain how it can accommodate genuinely hierarchical structure.

4.3 *Maintaining Hierarchical Structure*

I've argued that there are two perspectives that are implicit in our thinking about the metaphysics of ground: the ground-level perspective, from which reality is identified with the fundamental, and the hierarchical perspective, from which reality is seen as extending beyond the fundamental to encompass the non-fundamental as well. A pluralist about reality takes each of these perspectives to correspond to a real *way* of being the case, or a sense in which facts can *obtain*. Only the fundamental facts *really obtain* in one sense, while both fundamental and non-fundamental facts *really obtain* in another. For the pluralist, there is thus no univocal answer to the question of whether the higher-level facts really obtain in addition to their grounds: from the ground-level perspective on reality, they do not, while from the hierarchical perspective, they do.

It is important to note that on this approach, no distinction is drawn between something's being merely *the case* and its being the case *in reality*, as it is on Fine's (2001) view. The pluralist has no need for this additional distinction, given that she accepts multiple senses in which facts can (really) obtain in the first place. The "really" in the pluralist's claim that the non-fundamental facts "really obtain" in one sense but not in another is thus metaphysically redundant—it only serves to make clear that, from the relevant perspective, what we are concerned with is as "hefty" a metaphysical status as there is. For the pluralist, there are simply two such metaphysically privileged statuses—i.e., two senses in which a fact can (*really*) obtain.[20]

It is also important to make clear that the pluralist needn't maintain that the two perspectives at issue here are metaphysically on a par, or that they are both *maximally* fundamental. It may be that one of the two perspectives is *more* fundamental than the other, but that both are still metaphysically

20 My thinking about these issues draws heavily on Sider (2011), particularly on Sider's (2011, chap. 11) discussion of metaphysical saturation and redundancy.

privileged in that they each truly reflect something about reality's structure. There are several options one might pursue here.

One might find it natural to think that the ground-level perspective is more fundamental than the hierarchical perspective, or even that it alone is maximally fundamental, while the hierarchical perspective is not. The latter stance has the significant benefit of bringing together the two distinct senses of fundamentality I've been employing here: the perspective that is fundamental *qua* perspective is taken to be the perspective that "sees" only the fundamental level, i.e., from which only fundamental facts really obtain. The hierarchical perspective would then be seen as non-fundamental, though essential to making sense of the metaphysics of ground. Though such a stance would be natural for the grounding theorist to adopt, it carries the odd and somewhat counterintuitive consequence that, in the fundamental sense, hierarchical structure wouldn't be *real* after all. Realism about ground is in some sense rendered impossible on this view—reality cannot *really* be hierarchically structured or organized via the relation of ground, in that grounding structure is only "visible" from a non-fundamental perspective on reality.

Importantly, this is not just the trivial claim that grounding structure cannot be found *at the ground-level*. One might find it natural to think that despite grounding structure's not being fundamental in this sense, it is still *real* in a fundamental sense of the term. To maintain that grounding *genuinely* structures reality (i.e., that it's real *in a fundamental sense*, even if not itself fundamental), one must take both the hierarchical and ground-level perspectives to be fundamental *as perspectives*—i.e., to correspond to genuine, "joint-carving" conceptions of reality.

But as I said above, this still leaves several options open, in that one might take the two perspectives to be equally but not maximally fundamental, or take one perspective to be more fundamental than the other—depending, of course, on whether one allows for a comparative notion of perspective-fundamentality. Developing a more thorough conception of perspective-fundamentality, as well as of the various routes one might take here, is beyond the scope of this paper. My aim here is just to provide a rough sketch of the picture, which I think most faithfully reflects the grounding theorist's commitments, as well as the resolution such a picture could offer for the realist's puzzle. The intuitive appeal of the solution it offers can motivate the further development of variants of the pluralist view.

Nevertheless, it's important to flag that the questions surrounding perspective-fundamentality and the precise formulation of the pluralist's

doi: 10.48106/dial.v76.i1

view raise some serious conceptual and metaphysical challenges. For one thing, it is not clear how the view can be formulated without reliance on a third sense of "reality"—in saying that there **are** two distinct, joint-carving conceptions of reality, we seem to employ a third sense in which this can **be**. There are various routes one might take in response to this worry. One is to claim that the third sense of "reality" is simply non-fundamental, and that the two that the pluralist has identified are simply the two that best "carve at the joints." Another is to say that the third sense is *more* or even maximally fundamental, but not one we ordinarily employ. I explore each of these options elsewhere [in my Solomyak (2013); unpublished book manuscript], and wish to remain neutral on this issue here. For our purposes here, we can take the pluralist's claim to simply be that there are two metaphysically privileged conceptions of reality, corresponding to two distinct *ways* in which facts can be said to obtain.

With this claim in hand, the grounding theorist can make sense of hierarchical structure in a way she was previously unable to do. What it is for reality to have hierarchical structure, on the pluralist's understanding, is the following: *In one real sense*, there are both fundamental and higher-level facts, and *in another real sense*, reality is exhausted by the fundamental. The fundamental facts are thus real no matter which perspective one adopts, while the higher-level facts are real in one sense and unreal in another. This secures both the special status of the fundamental facts and the distinctive status of the grounded. The special status of the ground-level facts is secured by the fact that in addition to obtaining as the higher-level facts do, there is a real sense in which they exhaust reality; while the status of the grounded facts is secured by the fact that they really obtain in one sense but not in another—Distinct Obtaining secured by the fact that there is a real sense in which they obtain, and Nothing Further secured by the fact that there is a real sense in which they do not. For the pluralist, obtaining in one real sense but not in another is (part of) what it *is* for something to be a grounded, higher-level fact.[21]

21 One might worry that in taking this approach, the grounding theorist will need to admit many more than two senses in which facts can really obtain, assuming that a hierarchically structured reality has more than two "levels." That is, one might worry that there must be a distinct "perspective on reality" corresponding to each "level" of the hierarchy. But the pluralist can accept that facts on two distinct non-fundamental levels (that stand in a relation of ground) obtain in different ways without accepting that they obtain in *fundamentally* different ways. The pluralist can take all non-fundamental facts to obtain in a single fundamental *sense*, even if there are non-fundamental ways of distinguishing between them as well.

One might worry that there is a sense in which the reality of grounding structure is not secured after all, in that one of the essential requirements of ground is not satisfied from each of the two fundamental perspectives on reality. The two perspectives are incompatible, and each is missing out on an essential element of the metaphysics of ground. How is it that the pluralist gets to satisfy *both* requirements by embracing the two perspectives rather than, ultimately, neither (or at least, always not-both)?[22] To provide a fully satisfying response to this worry would require answering the questions posed above about the precise formulation of the pluralist view and the complications this raises. In particular, what is the more general sense of "reality" we are to use in stating that there *are* two fundamental perspectives? It is in this sense that the pluralist can say that grounding structure is *real*—each of the two requirements of ground is satisfied from one of the two fundamental perspectives. The objector might press on: Can we not equally say that each of the two requirements is, from some fundamental perspective, *unsatisfied*? But the intuitions behind the requirements of ground do not demand that each requirement be met *in every sense*, or even in every fundamental sense. What is essential is that there is *some real sense* in which the grounded obtains in addition to its grounds, and *some real sense* in which it is nothing over and above these grounds. It is thus enough that each requirement is satisfied from some fundamental perspective. This allows for hierarchical structure to be "*real*" in the more general, and yet to be fully explicated, sense of the term.[23]

So this, then, is how the metaphysical status of the genuinely *grounded* is to be found. To be grounded is not to belong to a mysterious in-between state, wedged between "obtaining in addition to" and "being nothing over and above." Rather, the grounded is what *in one sense* obtains in addition to, and *in another sense*, is nothing over and above. What makes the structure of reality genuinely *hierarchical* is that the "higher-levels" are, in one sense, there and, in another sense, literally *nothing* over and above the fundamental.[24]

22 Thanks to two anonymous reviewers for raising this worry.

23 Depending on how one understands this more general sense of "reality," it may thus turn out that the sense in which grounding is real is not itself fundamental, a possibility raised briefly above. Whether one takes this to be problematic for the grounding realist is an issue that requires further explication, and depends both on precisely how we understand realism in this context, as well as on how we understand the relationship between the various senses of "reality" embraced by the pluralist.

24 Interestingly, such a view is central to the metaphysics of Hasidic mysticism in the Jewish tradition. As Rabbi Shneur Zalman Borukhovich of Liadi (1745–1812) explains in what is known as the *Tanya* (1973, pts. 1, chaps.20–21), the Hasidic view is one on which only the fundamental—on

While what I have presented here is far from a developed metaphysical picture, it can be taken as a rough sketch of the direction in which I think the puzzle for the grounding realist pushes. The intuitive appeal of the pluralist's solution can bring out the ways in which we implicitly adopt and shift perspectives in thinking about the metaphysics of ground, and bring to light what simpler solutions to the realist's puzzle are thereby missing: A univocal answer to the question of whether the grounded facts *really obtain* cannot fully accommodate the competing requirements of ground. Recognizing that there are two genuine perspectives on reality here is the key to making sense of genuinely hierarchical structure, and developing a more thorough framework within which this claim can be understood is my aim elsewhere.

5 Conclusion

We've seen that the realist about ground faces a difficulty in accommodating genuinely hierarchical structure. The notion of ground requires that the grounded facts *obtain*, distinct from and in addition to their grounds, and that they be "nothing over and above" their grounds at the same time. These requirements push the realist about ground in seemingly competing directions: The realist is pushed to accept the Hierarchical Picture *over* the Ground-Level Picture on the one hand, and to reject that there is any substantive difference between these two pictures on the other. The higher-level, grounded facts are thus mysteriously elusive in their metaphysical status: we find them either too "metaphysically weighty" to be grounded or too "metaphysically light" to support the Hierarchical Picture, which seems essential for the grounding realist to maintain.

I've argued that the best route for resolving this puzzle involves rejecting what I've called the Assumption of Substantive Difference; the Ground-Level and Hierarchical Picture are not to be seen as reflecting distinct ways in which reality might be structured after all. But appreciating the implications of this stance and making it intuitive require some further reconceptualization of the issues. To make sense of the way in which the two pictures do not substantively differ, we can adopt the framework of *perspectives*: From one perspective, reality is exhausted by the fundamental, and from another perspective, reality encompasses the grounded, non-fundamental facts as well.

this picture, God—exists in one sense, and "everything else" exists in another. The unique sense in which God—seen as the ground or basis of the rest of reality—exists, is one in which *nothing else* really exists in addition.

The deep pluralism of this stance can provide a unique approach to understanding the metaphysical status of the grounded: the grounded is *in one sense* real, and in another sense literally *nothing* over and above the fundamental.*

Olla Solomyak
0000-0001-7177-0860
Shalem College, Jerusalem
os17@nyu.edu

References

Audi, Paul. 2012. “Grounding: Toward a Theory of the *In-Virtue-Of* Relation.” *The Journal of Philosophy* 109(12): 685–711, doi:10.5840/jphil20121091232.

Baruchovitch, Shneur Zalman. 1796. *Likutei Amarim (Tanya)*. Slavita: Baal HaTanya.

—. 1973. *Likutei Amarim (Tanya) [English Standard Revised Edition]*. New York: Kehot Publication Society. Translated by Nissan Mindel, Nissan Mangel, Zalman Posner, Jacob Immanuel Schochet; original publication: Baruchovitch (1796).

Bennett, Karen. 2011. “By Our Bootstraps.” in *Philosophical Perspectives 25: Metaphysics*, edited by John Hawthorne, pp. 27–41. Hoboken, New Jersey: John Wiley and Sons, Inc., doi:10.1111/j.1520-8583.2011.00207.x.

Correia, Fabrice. 2010. “Grounding and Truth-Functions.” *Logique et Analyse* 53(211): 251–279, https://logiqueetanalyse.be/archive/issues87-220/LA211/02correia.pdf.

Correia, Fabrice and Schnieder, Benjamin Sebastian. 2012. “Grounding: An Opinionated Introduction.” in *Metaphysical Grounding: Understanding the Structure of Reality*, edited by Fabrice Correia and Benjamin Sebastian Schnieder, pp. 1–36. Cambridge: Cambridge University Press, doi:10.1017/cbo9781139149136.001.

Daly, Christopher John. 2012. “Skepticism about Grounding.” in *Metaphysical Grounding: Understanding the Structure of Reality*, edited by Fabrice Correia and Benjamin Sebastian Schnieder, pp. 81–100. Cambridge: Cambridge University Press, doi:10.1017/cbo9781139149136.003.

Dasgupta, Shamik. 2014. “The Possibility of Physicalism.” *The Journal of Philosophy* 111(9/10): 557–592, doi:10.5840/jphil20141119/1037.

deRosset, Louis. 2013. “Grounding Explanations.” *Philosophers’ Imprint* 13(7), http://hdl.handle.net/2027/spo.3521354.0013.007.

Fine, Kit. 2001. “The Question of Realism.” *Philosophers’ Imprint* 1(2), http://hdl.handle.net/2027/spo.3521354.0001.002.

* Many thanks to Claire Benn, Sharon Berry, David Chalmers, Kit Fine, Martin Glazier, Casper Storm Hansen, Silvia Jonas, Leora Dahan Katz, Erica Shumener, Ted Sider, Georgie Statham, and audiences at the University of Helsinki as well as at the *eidos* “Truth and Grounds” Conference for invaluable comments and discussion.

—. 2005. “Tense and Reality.” in *Modality and Tense: Philosophical Papers*, pp. 261–320. Oxford: Oxford University Press, doi:10.1093/0199278709.003.0009.

—. 2012. “Guide to Ground.” in *Metaphysical Grounding: Understanding the Structure of Reality*, edited by Fabrice Correia and Benjamin Sebastian Schnieder, pp. 37–80. Cambridge: Cambridge University Press, doi:10.1017/cbo9781139149136.002.

Glazier, Martin. 2016. “Laws and the Completeness of the Fundamental.” in *Reality Making*, edited by Mark Jago, pp. 11–37. Mind Association Occasional Series. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780198755722.003.0002.

Hirsch, Eli. 2002. “Quantifier Variance and Realism.” in *Philosophical Issues 12: Realism and Relativism*, edited by Ernest Sosa and Enrique Villanueva, pp. 51–73. Oxford: Blackwell Publishers, doi:10.1111/j.1758-2237.2002.tb00061.x.

Hofweber, Thomas. 2009. “Ambitious, Yet Modest, Metaphysics.” in *Metametaphysics: New Essays on the Foundations of Ontology*, edited by David J. Chalmers, David Manley, and Ryan Wasserman, pp. 260–289. Oxford: Oxford University Press, doi:10.1093/oso/9780199546046.003.0009.

Koslicki, Kathrin. 2015. “The Coarse-Grainedness of Grounding.” in *Oxford Studies in Metaphysics*, volume IX, edited by Karen Bennett and Dean W. Zimmerman, pp. 306–344. New York: Oxford University Press, doi:10.1093/acprof:oso/9780198729242.003.0011.

McDaniel, Kris. 2009. “Ways of Being.” in *Metametaphysics: New Essays on the Foundations of Ontology*, edited by David J. Chalmers, David Manley, and Ryan Wasserman, pp. 290–319. Oxford: Oxford University Press, doi:10.1093/oso/9780199546046.003.00010.

Raven, Michael J. 2016. “Fundamentality without Foundations.” *Philosophy and Phenomenological Research* 93(3): 607–626, doi:10.1111/phpr.12200.

—. 2017. “New Work for a Theory of Ground.” *Inquiry* 60(6): 625–655, doi:10.1080/0020174x.2016.1251333.

Rosen, Gideon. 2010. “Metaphysical Dependence: Grounding and Reduction.” in *Modality: Metaphysics, Logic, and Epistemology*, edited by Bob Hale and Aviv Hoffmann, pp. 109–136. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199565818.003.0007.

Schaffer, Jonathan. 2009. “On What Grounds What.” in *Metametaphysics: New Essays on the Foundations of Ontology*, edited by David J. Chalmers, David Manley, and Ryan Wasserman, pp. 347–383. Oxford: Oxford University Press, doi:10.1093/oso/9780199546046.003.0012.

Schnieder, Benjamin Sebastian. 2011. “A Logic for ‘Because’.” *The Review of Symbolic Logic* 4(3): 445–465, doi:10.1017/s1755020311000104.

Sider, Theodore. 2009. “Ontological Realism.” in *Metametaphysics: New Essays on the Foundations of Ontology*, edited by David J. Chalmers, David Manley, and Ryan Wasserman, pp. 384–423. Oxford: Oxford University Press, doi:10.1093/oso/9780199546046.003.0013.

—. 2011. *Writing the Book of the World*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199697908.001.0001.

Solomyak, Olla. 2013. "Actuality and the Amodal Perspective." *Philosophical Studies* 164(1): 15–40, doi:10.1007/s11098-013-0096-8.

—. 2020. "Temporal Ontology and the Metaphysics of Perspectives." *Erkenntnis* 85(2): 431–453, doi:10.1007/s10670-018-0034-4.

Turner, Jason. 2010. "Ontological Pluralism." *The Journal of Philosophy* 107(1): 5–34, doi:10.5840/jphil201010716.

Wilson, Jessica M. 2014. "No Work for a Theory of Grounding." *Inquiry* 57(5/6): 535–579, doi:10.1080/0020174x.2014.907542.

doi: 10.48106/dial.v76.i1

Structural Realism and the Interpretation of Mathematical Structure

Noah Stemeroff

Structural realists typically appeal to the explanatory and predictive success of science to suggest that the mathematical structure of scientific theory, which is continuous across theory change, provides an accurate description of some aspect of the structure of the world. In this paper, I present a challenge to this claim that concerns how the relevant structure in nature is identified and represented in the context of a physical theory. I argue that the structures, on which many structural realists base the historical support for their position, can only be taken to represent "physical structures" in the context of a broader theoretical framework and that this framework is not necessarily preserved through theory change.

Structural realism holds that science comes closest to comprehending nature, not in its account of its constituents but in its account of its structure (e.g., see Stein 1989, 58). In its epistemic variant, structural realism suggests that scientific knowledge is limited to a structural description of reality. In its metaphysical variant, it defends a radical structural ontology of science. However, in both cases, the structural realist maintains that the significance of successful scientific theories consists in their ability to provide an accurate description of the structure of the world (e.g., see Ladyman and Ross 2007, 92).[1] This structure can be held to be metaphysically basic, or defined over a set of fundamental objects, but in either case, the scientific account of reality is taken to be essentially structural in nature.

1 Here, I take scientific realism, more generally, to be characterized by the belief that our best scientific theories provide a true, or approximately true, description of some aspect of the natural world (i.e., in both its observable and unobservable features).

There are three major motivations for the structural realist position (e.g., see Votsis 2017).[2] The first, historical, motivation is drawn from a particular response to the problem of scientific theory change (Worrall 1989; Ladyman 1998, 2011; Votsis and Schurz 2012). The history of science has shown that science is fallible. Many, if not all, scientific theories of the past are now considered to be false by the standards of modern science, and our current scientific theories will likely suffer the same fate. This pessimistic induction from the history of science is considered to be one of the strongest arguments against scientific realism (for more, see Laudan 1981). In response, many scientific realists have sought to defend selective forms of realism grounded on the portions of scientific theory that are preserved through theory change. The historical motivation for structural realism is based on the apparent continuity in the formal structure of scientific theory through the progress of science.

The second, epistemological, motivation for structural realism is derived from formal studies of the highly abstract nature of modern physics (e.g., Cao 1997; Morganti 2004, 2011; French 2014).[3] Here, it is argued that physics has become, in part, a study of the abstract mathematical structures that are taken to characterize the fundamental features of the natural world. In particular, the essential role that group-theory now plays in modern physics seems to entail that our knowledge of reality can only be determined up to an isomorphism—i.e., a given class of structure (e.g., see Lyre 2004). Thus, scientific knowledge itself may be formally limited to a description of the general structure of reality.

The third, metaphysical, motivation for structural realism takes the epistemological argument a step further (e.g., French and Ladyman 2003, 2011; Lyre 2004; Ladyman and Ross 2007; Esfeld 2013, 2017; Esfeld and Lam 2009, 2011; French 2014). Ontological structural realists argue that modern physics is not only in tension with, but can be taken to present a challenge to, the traditional object-based ontology of classical physics. For instance, the permutation invariance of quantum theory has been taken to directly undermine the individuality of quantum objects. This, along with a host of other examples of underdetermination drawn from both quantum mechanics and general

2 This is not to mention the additional motivation stemming from recent work on the structuralist methodology of science (e.g., Brading and Landry 2006).

3 In this context, it is also important to note the additional epistemic motivation that has come from a renewed interest in Russell's structuralist epistemology (e.g., Votsis 2005).

relativity, suggests that modern physics should be taken to support a theory of structural metaphysics (e.g., see Ladyman and Ross 2007; French 2014).

The epistemological and metaphysical motivations for structural realism have garnered the lion's share of attention in recent debates (e.g., see Bokulich and Bokulich 2011). But this does not mean that scholars have lost sight of the significance of the historical motivation for the position (e.g., Zahar 2007; Ladyman 2011; Vickers 2013, 2019; Saatsi 2019). Indeed, the historical motivation for structuralism continues to support much of the broader interest in the position as a form of scientific realism (i.e., as opposed to an account of the methodology of modern science). Of course, the historical arguments supporting structural realism have not gone unquestioned (e.g., Saatsi 2005; Wright 2017), but many structural realists continue to feel that the historical motivation for structuralism is sound—the cases of structural continuity in the history of science are clear, and the only remaining question concerns how best to understand this continuity (e.g., Ladyman 2011, 2018; French 2014). However, the structural realists' portrayal of the history of science as a progressive series of structural descriptions of reality, and the broader framework that defines this sense of progress, is often taken for granted. This raises the question of what exactly constitutes structural continuity through the progress of science in the first place.

At the outset, it is important to note that a mere formal continuity of structure through theory change, although necessary, is not sufficient to support a viable realist position. The structural realist must demonstrate that the continuous structure of scientific theory represents some aspect of the world—as opposed to simply providing a convenient language to express observable facts (e.g., see Duhem 1991, 151). The retention of structure must mark a sense in which different theories can be said to accurately represent the same reality, at least in some sense (e.g., see Psillos 1995, 1999; Chakravartty 2007). Otherwise, one could easily argue that the retention of structure is simply "a pragmatic feature of scientific practice" (Psillos 1999, 152).

Putnam (1975, 73) famously wrote that scientific realism "is the only philosophy that doesn't make the success of science a miracle." Successful scientific theories explain and predict the outcome of experiments. It is this ability to explain and predict an empirical phenomenon that leads us to conclude that a successful theory provides a true, or approximately true, description of the world. Worrall (1989, 121) suggests that structural realism offers "the best of both worlds" by charting a middle path between Putnam's "no miracles" argument and the pessimistic induction. However, to apply the "no miracles"

doi: 10.48106/dial.v76.i1

argument, the structural realist must demonstrate that the structure of scientific theory is, at least in part, responsible for its explanatory and predictive success. The structural realist position would collapse into instrumentalism if the structure of physical theory cannot be said to have some "grip on reality" (Ladyman 1998, 418).[4]

To defend the historical motivation for structural realism, structural realists must show that the structure, which is continuous across theory change, can be taken to represent the same structure in nature, and that this structure can account for the relevant physical phenomena. The concern here is to not mistake a continuity of symbolism for a continuity of representation (Cao 2003, 14). The structural realist must not only demonstrate that there exists a continuity in the formalism of physical theory, but also that this continuity entails a continuity of representation. In this context, it is important to note that in order for a structure to account for a physical phenomenon, it must be accompanied by a suitable interpretation. It is the interpretation that correlates a given structure to the natural world.[5] This may be a trivial point, but it represents a non-trivial problem for the structural realist. If a continuity of structure is not sufficient to establish a continuity of representation, then the structural realist must demonstrate that the relevant structure, along with a suitable interpretation, is maintained across theory change.

In what follows, I will argue that this concern presents a challenge to the historical motivation for the structural realist position. This challenge concerns the way in which the mathematical structure of a physical theory is interpreted as a description of the structure of the natural world. In particular, I will argue that the structural realist faces a problem in specifying how a mathematical structure is correlated to nature across the progress of science. To support this criticism, I will present two case studies concerning two of the most prominent articulations of the historical motivation for the structural realist position—i.e., Worrall (1989) and Ladyman and Ross (2007).

Against Worrall's structural realism, I will argue that the mathematical structure, on which he bases his realism, cannot provide a description of the relevant physical structure in nature—at least in the context of an actual experiment—as it requires a theoretical interpretation. To defend this claim,

4 Here, instrumentalism is characterized by the belief that our best scientific theories provide an accurate description of the observable features of the natural world, but nothing more.

5 In structuralist circles, this notion of correlation is typically cashed out in terms of a structural isomorphism or a similar representation relation (e.g., see Pincock 2005, 2012; Brown 2012; Bueno and French 2018).

I will present a detailed re-examination of Worrall's seminal historical case study concerning the transition from Fresnel's optical theory to Maxwell's. Through this case study, I will show that the holistic nature of the interpretation of the mathematical structure of a physical theory threatens to undermine the continuity of structure supporting Worrall's structural realism. The problem is that to correlate a mathematical structure to a physical structure in the world, Worrall's structural realist needs to specify the formal and theoretical framework required to characterize the physical structure and the system of which it is a part. The question is, then, to what extent is this framework maintained, or suitably translated, in the transition to a new theory, and can we still claim that the two theories describe the same physical structure in the world?

In contrast, Ladyman and Ross (2007) do not seem to fall prey to this concern, as their structural realism is based on a more general appeal to the modal structure of reality. On this account, the modal structure of reality is identified with "real patterns" in observational data rather than physical structures in nature. However, against Ladyman and Ross, I will argue that the holistic nature of the interpretation of the mathematical structure of a physical theory may still present a challenge to their modal structural realism in the context of the history of modern particle physics, which is one of the key case studies they take to support their position. Once again, the concern relates to how the abstract mathematical structure of modern physical theory is interpreted as a representation of the modal structure of nature, which they argue is identified in an experiment.

1 Interpreting Mathematical Structures

Ladyman and Ross (2007, 67) define structural realism as "the view that our best scientific theories describe the structure of reality, where this is more than saving the phenomena, but less than providing a true description of the natures of the unobservable entities that cause the phenomena." But what is this "structure" that is "described" by a scientific theory? In the context of the historical argument for structural realism, the structure of reality is often taken to be described by certain aspects of the mathematical formalism of a scientific theory. However, it is not always clear in what sense we should interpret a given mathematical structure as a description of a given structure in nature, or when we should interpret a given mathematical structure as continuous across theory change. On the one hand, it is clear that past and present scientific

doi: 10.48106/dial.v76.i1

theory adhered to entirely different theoretical and experimental practices, not to mention methodologies. On the other hand, the interpretation of the structure of past science must be explained in light of the success of current scientific theory, which is taken to provide an accurate description of the structure of the world—at least approximately. The question is then: how much of the broader framework of modern scientific theory do we need to project back onto past scientific theory to interpret parts of its mathematical structure as a continuous across theory change?

An initial problem relates to the general interpretation of the mathematical formalism of past science. Here, the concern is that we need to specify a formal framework to even determine the meaning of a mathematical structure. This is a fairly general concern—to be applicable, mathematical structures must be definable. We cannot speak coherently of a mathematical structure divorced from the formalism that gives it meaning. Take, for example, a dynamical equation that is taken to provide a description of the evolution of some system, say a ball thrown in the air. It would be meaningless to say that the dynamics of the system can be represented by a solution to a given equation without specifying the underlying formal framework in which the equation is defined. This framework delimits the manifold and metrical structure required to ensure the differential structure of a dynamical equation is well-defined and the constraints on its domain of application. It is only within the context of this formal framework that the equation can be taken to characterize a path in a geometrical space. It is this path that is actually taken to describe the evolution of the system.[6] However, the successful theories of the past often lacked what we would now consider to be a proper formal framework, and it is not always clear how we should interpret their mathematical structure.

A subsequent problem relates to the manner in which we interpret a given mathematical structure as a description, or representation, of nature. To interpret a mathematical structure in an empirical setting, we need to specify how the structure is to be situated within the context of a physical system or experimental result. The problem is that it is the theoretical framework of a physical theory that is responsible for delimiting its domain of application. Returning to the case of a dynamical equation, it is clear that in order to say that the evolution of the system is characterized by a solution to a given dynamical equation, we need to specify how the equation is to be understood

6 This may be slightly pedantic, but it is important to note that a given dynamical equation may define entirely different paths depending on the mathematical framework in which it is formulated.

in the context of a given empirical setting. It is the theoretical framework of a physical theory, e.g., classical mechanics, that provides an account of the physical space through which a given object moves, the vantage point from which the motion is defined, and the constraints that may be present on the system, as they constitute essential features of the physical context in which the structure is taken to apply. The structure that a given dynamical equation is taken to describe cannot be properly situated or understood outside the theoretical framework of a physical theory. However, the scientific theories of the past were formulated within vastly different theoretical frameworks, and it is, once more, not entirely clear how we interpret their mathematical structure as a representation of a physical system from the perspective of modern science.

These two concerns are closely related. They both pertain to the fact that to ground a realist account of mathematical representation, we must first ensure that the representation is well-defined—i.e., that the relevant mathematics is well-defined and applicable (i.e., interpretable) in a physical setting. To delimit its definition, we must provide a formal framework in which the mathematical structure is defined. To delimit its applicability, we need to specify the physical system in the world that it is taken to represent. This must be done prior to any question of the correlation between a given mathematical structure and nature.

The worry here is that if the formal and theoretical framework of past scientific theory is inconsistent with that of today or not entirely well-defined (from the perspective of modern science), then the structural realist may be forced to project too much of the formalism of modern scientific theory onto past science. Otherwise, it might be impossible for the realist to define the sense in which two seemingly identical equations can be taken to represent the same structure in nature. But the structural realist must be able to identify the sense in which past science, on its own, can be taken to describe the same structure in nature. Otherwise, they run the risk of simply imposing continuity rather than identifying it. In what follows, I argue that these concerns pose a distinct challenge to the historical motivation for structural realism.

2 Worrall, and the Problem of Physical Structure

Worrall suggests that structures are preserved through theory change because they play an essential role in accounting for physical phenomena. For example, he (1989, 116) suggests that there "was an important element of continuity in

doi: 10.48106/dial.v76.i1

the shift from Fresnel to Maxwell—and this was much more than a simple question of carrying over the successful *empirical* content into the new theory [...] the continuity is one of *form* or *structure*." The continuity, in this case, is in description of the phenomena of diffraction and the reflection and refraction of polarized light. Worrall continues, "it is no miracle that [Fresnel's] theory enjoyed the empirical predictive success that it did; it is no miracle because Fresnel's theory, as science later saw it, attributed to light the right *structure*" (1989, 117). However, it is not entirely clear how Fresnel's equations actually characterize the structure of light. Worrall assumes that a continuity of mathematical structure entails a continuity of representation—that the two are coextensive (1989, 117–119). However, an argument is needed.

Worrall bases his defence of structural realism on a detailed historical case study concerning the transition from Fresnel's ether-based theory of light to Maxwell's theory of the electromagnetic field. This case study is meant to demonstrate that "if we restrict ourselves to the level of mathematical equations—*not* notice the phenomenal level—there is in fact complete continuity between Fresnel's and Maxwell's theories" (Worrall 1989, 119). Worrall suggests that this continuity in the mathematical structure of scientific theory represents a continuity in the description of the structure of the world (1989, 116). In Worrall's view, this structure was responsible for the empirical success of Fresnel's theory and was retained in the transition to Maxwell's theory of light. However, I will argue that it is not entirely clear that this case study actually supports Worrall's conclusion.

2.1 *Fresnel's Theory of Light*

Fresnel championed the wave theory of light over the corpuscular, or emissionist, ray theory that was dominant at the time. His work on diffraction and the reflection/refraction of polarized light is often credited with bringing about the widespread acceptance of the wave theory of light (Buchwald 1989, 291–310). However, the development and success of Fresnel's mathematical description of light can only be understood within its historical context, as this context determined its interpretation. Worrall's case study focuses on the mathematical structure of Fresnel's theory, but in order to understand how these equations were empirically interpreted, we must first address the specific experiments that Fresnel's equations were taken to describe.

Fresnel's defence of the wave theory of light began with a wave-theoretic account of the phenomenon of diffraction—that is, the bending of light around an obstructing object. The phenomenon of diffraction was first observed by Grimaldi in the seventeenth century. In Fresnel's time, diffraction was easily explained within the context of the ray theory of light, which developed out of Newton's corpuscular theory of light. Newton held that white light is composed of a collection of particles of different size, shape, and velocity. In Newton's view, the size and velocity of the particles accounted for their colour. The primary advantage of the corpuscular theory over the wave theory was the ease through which it accounted for the rectilinear propagation of light. Newton held that the fundamental flaw in the wave theory of light was its failure to account for this fact (Darrigol 2012, 104). Within the corpuscular ray theory, the phenomenon of diffraction was explained by the existence of a localized force in the neighbourhood of the boundary of a diffracting object. This force accounted for the observed inflection of the corpuscles of light at the boundary. In light of the success of the corpuscular ray theory, Fresnel had to show that the wave theory could account for the inflection of light and its rectilinear propagation. Famously, Fresnel was able to account for both diffraction and the rectilinear prorogation of light through an application of Huygens' principle and the principle of interference (Buchwald 1989, 160–162).

Huygens' principle states that each element of a wavefront of light serves as the source of a new outgoing wave. The waveform at any given point in space and time can be determined through the principle of interference. Fresnel simply needed to sum the contributions from each outgoing wave that reaches a given point at a given time. Fresnel's mathematical treatment of diffraction identifies the source of diffraction in the wavefront that passes unimpeded around the diffracting object. He accounted for the interference pattern that is observed in the shadow of a diffracting object by applying Huygens' principle and the principle of interference to sum the outgoing waves from each element of the unobstructed wavefront.

Fresnel was able to integrate over the unobstructed wavefront, and found that the resulting oscillation at any point P beyond the diffracting body is proportional to:

$$\int \cos(\omega t + kz^2)dz, \tag{1}$$

where z is the distance from the source point on the unimpeded wavefront to the point P, ω is the angular frequency, k is the wave number, and t is the time (Darrigol 2012, 204). Fresnel defined the amplitude of the wave at the

doi: 10.48106/dial.v76.i1

point P to be proportional to $\sqrt{U_c^2 + U_s^2}$, with:

$$U_c = \int \cos(kz^2)dz, \tag{2}$$

$$U_s = \int \sin(kz^2)dz. \tag{3}$$

These equations are collectively known as Fresnel's integrals and constitute the essence of Fresnel's prize-winning paper to the French Academy of Sciences in 1818.[7] If you apply Fresnel's integrals to account for the diffraction of light through a slit in a screen and take the limit as the width of the slit tends to infinity, then you observe that the light's propagation beyond the slit is rectilinear. This is simply due to the effects of destructive interference. This result established the first formal proof of the rectilinear propagation of light within the burgeoning wave theory. In conjunction with Fresnel's mathematical description of diffraction, this result established the formal viability of the wave theory of light.

However, to establish the veridicality of the wave theory, Fresnel had to show that it could successfully account for the observable phenomenon of diffraction. To observe diffraction, we need a source of light, an object or surface, and a screen upon which to cast a shadow. As a simple example, we can consider a variant of the famous diffraction experiment that Poisson devised to test the predictions of Fresnel's prize-winning paper. In this experiment, light is cast on a circular disk, and the diffraction pattern is observed on a screen. Poisson recognized that Fresnel's wave theory of light predicted that a bright spot would appear behind the diffracting disk at the centre of the

7 In its modern form, the three-dimensional equation states that:

$$U(x, y, z) = \frac{k}{2\pi i}\int_{x_0'}^{x_1'}\int_{y_0'}^{y_1'}\int_{z_0'}^{z_1'} U_1(x', y', z')\frac{e^{ikr}}{r^2}(\hat{n}\cdot\mathbf{r})dx'dy'dz'. \tag{4}$$

Where $U(x, y, z)$ denotes the amplitude of the displacement of the wave at the location (x, y, z) (neglecting polarization for the sake of simplicity), $x_0' \rightarrow x_1'$ are the x components, $y_0' \rightarrow y_1'$ are the y components, and $z_0' \rightarrow z_1'$ are the z components of the wavefront that passes unimpeded around the obstructing object, $U_1(x', y', z')$ is the amplitude of each surface element of the wavefront, and e^{ikr}/r^2 is the amplitude of each propagating wavefront, k is the wavenumber, $\hat{n}$ and $\mathbf{r}$ are the vectors that define the normal of the incoming wavefront and distance to the point under consideration, and $r = \|\mathbf{r}\|$. It is important to note that Kirchoff was the first to provide the formal basis for this mathematical description of diffraction. Before Kirchoff, the mathematics of Huygen's principle was not well-formulated or even well-defined (Buchwald 1989, 188).

screen. This central spot was indeed observed, and this experiment served as an important early verification of the wave theory of light.

At this point in the story, everything seems to be going according to plan for Worrall's structural realist. Fresnel's diffraction integrals appear to correctly predict the outcome of this novel experiment, and the reason why could very well be that the mathematical structure of Fresnel's equations accurately describes the structure of light. However, the problem remains to show how Fresnel's diffraction integrals can be interpreted to provide a well-defined representation of the structure of light. Fresnel's integrals are not well-defined when separated from the mathematical framework of Fresnel's optical theory. This framework is required to define the underlying fixed spatial and temporal structure through which wave propagation is defined. This framework is required to not only effectively solve Fresnel's integrals, but to ensure that they actually define the structure of wave propagation—i.e., to interpret the mathematical structure. If the structural realist wants to argue that Fresnel's diffraction equation can account for the structure of light, then the mathematical formalism that is required to interpret Fresnel's integrals must be written out explicitly and included in the set of equations that define Fresnel's account of light. This formalism must then be maintained, or at least suitably translated, in the transition to Maxwell's theory.

In addition, it is not clear whether Fresnel's integrals, on their own, can be interpreted to provide a prediction of the outcome of a diffraction experiment, or to situate the structure of light within it. The integrals only describe light in free propagation, but we never observe light in free propagation; we can only observe light when it interacts with matter. In fact, there is nothing in Fresnel's integrals that makes reference to matter or the condition of observation. Although the integrals are thought to describe the propagation of light and the interference pattern that results from the propagation of the unobstructed wavefront, they cannot take into consideration the actual physical setup of the experiment. What is lacking is an account that serves to correlate the observable outcome of the experiment to the structure of Fresnel's equations—i.e., to show that the observed result is a consequence of this structure. But this would require an account that locates this structure within the experimental setup, which can only be defined by certain aspects of Fresnel's optical theory that account for the initial emission, reflection, and observation of light. To define a continuity of interpretation across theory change, this framework must also be maintained, or at least suitably translated, in the transition to Maxwell's theory.

doi: 10.48106/dial.v76.i1

Unfortunately, the structural realist seems to run into similar problems in the case of Fresnel's equations for the reflection and refraction of light—that is, the reflection of light off of a surface boundary and the bending of light as it passes through the surface boundary into a medium with a higher or lower refractive index. After his initial success with diffraction, Fresnel turned his attention to the newly discovered phenomenon of polarization. Although the phenomena of reflection and refraction had been studied since antiquity, the phenomenon of polarization was first observed by Malus in the early nineteenth century. Initially, the phenomenon of polarization was easily accounted for by the corpuscular theory under the umbrella of the ray theory of light.[8] Ray theorists held that light consists of a bundle of rays and that each light ray possessed an inherent asymmetry—Buchwald (1989, 50–51) uses the analogy of a broom handle with a nail hammered in it (the broom handle represents the direction of the ray and the nail its asymmetry). Ray theory offered a theoretical means to treat polarization as a property of a bundle of rays. The theory suggested that, under normal circumstances, the distribution of the asymmetries in a given bundle of rays is entirely random and unobservable. However, under rare circumstances, they held that the rays with a particular asymmetry could be preferentially selected, thus resulting in a skewed distribution in a given bundle. To the ray theorist, polarization was nothing but a prevalence of a certain asymmetry within a given bundle. The ray bundle theory of polarization successfully explained a number of early polarization experiments. Unfortunately, this all changed with Arago's discovery of chromatic polarization and Fresnel's discovery of rectilinear, circular, and elliptical polarization (Buchwald 1989, 67–85, 222–231).

In contrast to the static ray theory of polarization, Fresnel formulated a dynamical transverse wave theory of light. He suggested that light waves oscillate in time perpendicular to the normal of the wavefront. This dynamical conception of polarization marked a profound reconceptualization of the structure of light. Fresnel's theory took polarization to be a local feature of every element of a wavefront. This meant that Fresnel had to trace the dynamical propagation of every single element of the wavefront in order to explain the observed behaviour of light. Despite this challenge, Fresnel was able to devise a successful account of the reflection and refraction of light.

8 Within the ray theory, light was taken to be constituted out of luminous corpuscles that formed rays. The rays were assumed to be countable and were taken to correspond to the ray tracks in geometric optics.

Fresnel's equations define the polarization-dependent angle of reflection and refraction at the interface between two transparent substances. To derive these equations, Fresnel made two assumptions. First, he assumed conservation of energy across the surface that defines the boundary between the two substances. Second, he assumed that the amplitude of the transverse polarization is continuous across the surface. Given these conditions, the law of reflection, and Snell's law, Fresnel was able to derive his reflection/refraction equations.[9] Fresnel's equations state:

$$\frac{U_y^{reflected}}{U_y^{Incident}} = \frac{\sin(i-r)}{\sin(i+r)}, \tag{5}$$

$$\frac{U_x^{Reflected}}{U_x^{Incident}} = \frac{\tan(i-r)}{\tan(i+r)}, \tag{6}$$

$$\frac{U_x^{Refracted}}{U_x^{Incident}} = \frac{2\sin(r)\cos(i)}{\sin(i+r)\cos(i-r)}, \tag{7}$$

$$\frac{U_y^{Refracted}}{U_y^{Incident}} = \frac{2\sin(r)\cos(i)}{\sin(i+r)}. \tag{8}$$

Where U denotes the transverse amplitude of the displacement of the light wave at the interface, the subscripts x and y refer to the orthogonal components in the plane of polarization, i refers to the angle of the incident and reflected waves, and r the angle of the refracted wave (both measured relative to the normal to the surface).

To establish the veridicality of the wave theory of polarization, Fresnel had to show that it accounts for the observable phenomena of reflection and refraction. To observe reflection and refraction, we need a source of light, a block of a homogeneous transparent substance (e.g., glass), and two screens—one to observe the reflected light and one to observe the refracted light. Fresnel's equations are able to accurately describe the observed location of the reflected and refracted light in a diffraction experiment.

9 The law of reflection states that the angle of incidence equals the angle of reflection $\theta_{incident} = \theta_{reflected}$, both measured relative to the normal of the surface. Snell's law, or the sine law for refraction, states that $n_1 \sin\theta_1 = n_2 \sin\theta_2$, where n_1 is the index of refraction of the first substance (e.g., air), n_2 is the index of refraction of the second substance (e.g., glass), θ_1 is the angle of the incident light, and θ_2 the angle of the refracted light (both measured relative to the normal to the surface).

However, once again, it is not clear whether Fresnel's equations for the reflection and refraction of polarized light, on their own, can be interpreted to provide a well-defined representation of the structure of light. To reiterate, Fresnel's equations for the reflection and refraction of light are not well-defined when separated from the mathematical formalism of Fresnel's optical theory. This framework serves to define the fixed spatial and temporal structure through which the waves propagate, and the very distinction between transverse and longitudinal wave propagation. Just as in the case of diffraction, if the structural realist wants to argue that Fresnel's equations can account for the physical structure of reflection and refraction, then the mathematical formalism that is required to interpret Fresnel's equations must be written out explicitly and shown to be suitably maintained in the transition to Maxwell's theory.

Similarly, it is not clear whether Fresnel's equations, on their own, can be interpreted to provide a prediction of the outcome of a reflection and refraction experiment or situate the structure of light within it. Fresnel's equations only describe the structure of light at the interface between the air and the refractive substance, but that is not what we observe. The structural realist needs to clarify the sense in which this result is due to the structure of the phenomenon that these equations are taken to describe. The problem is that these equations, on their own, are not able to locate this structure in the world. There is nothing in Fresnel's equations that makes reference to the condition of observation, and they cannot take into consideration the actual physical setup of the experiment. What is lacking, once again, is an account that relates the observable outcome of this experiment to the structure of Fresnel's equations—i.e., to show that the observed result is a consequence of this structure. Again, this account must be maintained, or appropriately translated, in the transition to Maxwell's theory.

In response to these concerns, the structural realist might simply acknowledge that the relevant structures require an interpretation to be formally defined and correlated to the appropriate structure in nature. Of course, this will be done differently in each theory, and some features of these interpretations may be abandoned through theory change, but the underlying structure remains and can still be effectively correlated to the relevant phenomena. The problem is that it is not at all clear that Fresnel's equations will describe the same structure in this case. We need to be very careful to not mistake a continuity of symbolism for a continuity of structure. Worrall's structural realist needs to show that despite the change in the theory, the relevant mathematical structure still depicts the same structure in the world.

It is important to note that there is no question here of the instrumental value of Fresnel's equations. The question is whether they can be taken to depict the structure of light, and whether this structure is responsible for the explanatory and predictive success of the theory. Worrall's structural realist needs to show that the explanatory and predictive success of the theory is a consequence, at least partially, of the accurate description of the structure of light. To do this, they need to clarify the sense in which this structure is responsible for the outcome of the relevant experiments—i.e., they need to directly correlate this structure to the relevant observable phenomena. In the case of Fresnel equations, they need to present an account of the structure of the wave propagation throughout the physical system and correlate this structure to experimental observation. This requires that they situate Fresnel's equations within a framework to account for how light interacts with the experimental setup. Once this is done, the question is to what extent this account is maintained, or suitably translated, in the transition to Maxwell and whether we can still claim that the two theories describe the same structure in the world.

2.2 *The Transition to Maxwell's Theory*

It is clear that within Maxwell's theory of the electromagnetic field, one finds equations in the symbolism of Maxwell's theory that appear to be formally similar to Fresnel's equations for the diffraction and reflection and refraction of light. This continuity is not in question. The challenge is to determine whether this continuity is merely a symbolic continuity, or whether it represents a continuity of description. This is a question of the interpretation of the shared mathematical structure of the theories. Worrall's structural realist needs to show that it is the shared structure that is responsible for the shared success of the theories. However, both theories describe the structure of light in terms of a transverse wave equation, and both theories seem to refer to this structure to explain the phenomena of diffraction, reflection, and refraction of light, so there may not be much of a problem. The only worry is that they correlate this structure to observable phenomena in slightly different ways. This concern primarily involves the way in which light is taken to interact with matter.

doi: 10.48106/dial.v76.i1

Fresnel initially attempted to address the interaction between light and matter through an account of the phenomenon of dispersion—that is, the wavelength-dependent refraction of light. He knew that the effects of dispersion had to be taken into account and that his neglect of dispersion in the case of diffraction and reflection/refraction meant that his results were only approximate in nature (Buchwald 1989, 306–310). In fact, Fresnel put forward an intriguing idea for the development of a dynamical theory of dispersion (e.g., see Buchwald 2012). He suggested that dispersion might be the result of the coarse-grained nature of matter. He assumed that matter is composed of many "atoms" with a certain characteristic spacing. Fresnel thought that each "atom" would place a stress on the ether, which he considered to be an elastic medium. Fresnel suggested that we could use this periodic loading of the ether to account for the phenomena of dispersion. In Fresnel's theory, dispersion was taken to be dependent upon the ratio of the wavelength of light to the characteristic spacing of the "atoms" in a substance.

Sadly, Fresnel passed away at the age of thirty-nine, before he was able to complete an account of the interaction between light and matter (Buchwald 1989, 307–308). However, three years after Fresnel's death, Cauchy took up Fresnel's suggestion for a theory of dispersion. By applying Navier's theory of elastic solids and point-centres of force, he was able to derive a modified wave equation for the propagation of light within a dispersive substance. Cauchy's modified wave equation states:

$$\frac{\partial^2\eta}{\partial t^2} = \alpha\left(\frac{\partial^2\eta}{\partial x^2}\right) + \beta\left(\frac{\partial^4\eta}{\partial x^4}\right) + \gamma\left(\frac{\partial^6\eta}{\partial x^6}\right) + \dots, \tag{9}$$

where α, β, and γ are constants, η is the displacement of the ether, and x is the direction of propagation (Whittaker 1951, 165). Substituting in the solution $\eta = e^{2\pi i(x-c_1 t)/\lambda}$, Cauchy then solved for the velocity of light in a dispersive medium:

$$c_1^2 = \alpha - \beta(\frac{2\pi}{\lambda})^2 + \gamma(\frac{2\pi}{\lambda})^4 + \dots, \tag{10}$$

where c_1 is the phase velocity of light, and λ is the wavelength of light. This expression shows that in a dispersive medium, the velocity of light is wavelength-dependent, as we would expect. The index of refraction for a dispersive substance is then given as:

$$\mu^2 = \frac{\alpha}{c^2} - \frac{\beta}{c^2}(\frac{2\pi}{\lambda})^2 + \frac{\gamma}{c^2}(\frac{2\pi}{\lambda})^4 + \dots. \tag{11}$$

The essential feature of Cauchy's account is the use of a modified wave equation to represent the effects of dispersion. Since Fresnel placed the locus of polarization on the wavefront itself, a structural account of an optical experiment requires that we trace the propagation of a wavefront throughout the experimental setup. Cauchy's modified wave equation would have allowed Fresnel to do exactly that.

The Fresnel-Cauchy theory of dispersion was eventually shown to be fundamentally flawed by the discovery of anomalous dispersion by Leroux in 1862. Leroux observed that a prism filled with iodine gas refracted red light more than blue light. This contradicted the Fresnel-Cauchy theory of dispersion, which predicted that the refractive index increases with the frequency of light. Stokes pointed out that effects of anomalous dispersion could be explained if we simply posit that every substance possesses certain natural frequencies of vibration. He suggested that matter itself is a dynamical system that possesses natural vibratory frequencies that interact with the incident vibrations of light. Stokes also noted that the effects of anomalous dispersion could account for the surface colour of objects.

Maxwell devised a theory of dispersion that took account of the crucial discovery that every substance possesses a set of natural dispersive frequencies (Whittaker 1951, 262–265). Maxwell suggested that material bodies are formed out of an immense number of "atoms," which occupy holes in the ether. He thought that each "atom" consists of a number of shells, where the outermost shell is in contact with the ether. In Maxwell's view, dispersion was a result of the natural vibrational character of the shells within each "atom." The idea is that as light propagates through a material substance, it can set the atoms in motion. Since each "atom" has certain allowable oscillatory frequencies, each frequency represents a natural basis for dispersion.

Maxwell derived a modified wave equation for the propagation of light in a dispersive substance by specifying the kinetic and potential energy of the ether between the "atoms" of a substance. Assuming that the system conserves energy, he was able to derive the equation of motion for light propagation in a dispersive medium. He found that the propagation of light in a dispersive substance with a single natural vibrational frequency is given by the following equation:

$$\left(1+\frac{\sigma}{\rho}\right)\frac{\partial^2\eta}{\partial t^2}-c^2\left(\frac{\partial^2\eta}{\partial x^2}\right)+\frac{1}{p^2}\left(\frac{\partial^4\eta}{\partial t^4}\right)-\frac{c^2}{p^2}\left(\frac{\partial^4\eta}{\partial x^2\partial t^2}\right)=0, \tag{12}$$

doi: 10.48106/dial.v76.i1

where η is the displacement of the ether, σ is the mass of the atomic particles per unit volume, ρ is the ethereal density, and p is the vibrational frequency of the “atom” (Whittaker 1951, 263).[10] Assuming that the substance through which the light propagates has a natural frequency of vibration, n, Maxwell found that the index of refraction, μ, in a dispersive substance—within the limits of the visible spectrum—is given as:

$$\mu^2 = 1 + \frac{\sigma}{\rho}\left(1 + \frac{n^2}{p^2} + \frac{n^4}{p^4} + \ldots\right). \tag{13}$$

Maxwell then expanded his dispersion relation to allow for a possibly infinite series of natural oscillatory frequencies and determined the refractive index for a substance through the following relation:

$$\mu^2 = 1 + \frac{c_1}{p_1^2 - n^2} + \frac{c_2}{p_2^2 - n^2} + \ldots, \tag{14}$$

where c_1 refers to the velocity of a light wave of frequency p_1, c_2 to the velocity of a light wave of frequency p_2, and so on.[11] Through this relation, Maxwell was able to determine the dispersion of light in any substance once the natural oscillatory frequencies of the atomic constituents had been found. Maxwell's theory of dispersion was confirmed at the end of the nineteenth century by Rubens (Whittaker 1951, 265).

Comparing Cauchy's and Maxwell's modified wave equations, we can see that Cauchy's structural description of dispersion is not maintained across the transition to Maxwell's theory. This is not a surprise, since Cauchy and Maxwell had a different understanding of the structure of both matter and the ether. To Cauchy, dispersion was a result of the coarse-grained nature of matter, whereas to Maxwell, it was the result of the interaction between the light and the natural oscillatory frequencies of matter. The question is whether this apparent discontinuity poses any real challenge to the structural realist. It is clear that both Cauchy and Maxwell agree that light will satisfy a modified wave equation with additional derivative terms accounting for the dispersive nature of the substance. Despite the fact that Cauchy and Maxwell disagree about the nature of the dispersive terms, there is a sense in which they agree

10 Note that either one of the last two terms in the above relation is sufficient to produce dispersion within the substance.

11 This result is based on an account given by Whittaker (1951). A slightly different account concerning dispersion in prisms is given in Maxwell (1867–1868).

as to the nature of dispersion. That is, that dispersion should be represented by a modified wave equation.[12] It seems that this subtle discontinuity should not pose a real challenge to the structural realist.

However, it is important to recall that Worrall argues that the success of Fresnel's theory is due to the fact that he "attributed to light the right *structure*" (1989, 117). Given that both theories defend a wave theory of light in which the key structural features propagate on the wavefront itself, this entails that they must be able to locate this wave structure throughout an experiment to correlate it to the observable outcome. Otherwise, one could easily argue that the equations are of mere instrumental value. The problem is that the observable phenomena, in the case of the diffraction, reflection, and refraction of light, are not correlated directly to the shared wave structure. Due to their differing accounts of the interaction between light and matter, the specific mathematical structures picked out by Fresnel and Maxwell as responsible for the observable phenomena are actually subtly different. The structures picked out by Fresnel and Maxwell differ because of their disparate accounts of the emission and dispersion of the propagating wave. It is not the actual description of the physical structure of light that is continuous, but rather the more general wave-like nature of this structure. In both cases, we are detecting something that has the mark of a transverse wave, but not the same physical structure. However, this may not pose a significant problem for the structural realist, as one could argue that the shared type of structure is responsible for a key part of the explanatory and predictive success of the theory.[13]

In the end, this challenge may be manageable. All that Worrall needs to show is that, from the perspective of modern science, we can continue to reinterpret both Fresnel's and Maxwell's theories to account for the right general structure, instead of a specific physical structure. In this case, it would seem that we can easily mitigate the challenge posed by the holistic nature of the interpretation of a mathematical structure in the context of an experiment. The real problem is that this challenge has only become worse with time. The past two centuries have witnessed a dramatic shift in the structural account

12 In addition, it can also be shown that Cauchy's formula converges to Maxwell's when n, the natural vibratory frequency of matter, is taken to fall within a specific range (Whittaker 1951, 264).

13 In addition, in the transition from Fresnel to Maxwell, the structural realist may be able to mitigate some of these concerns if they can identify an appropriate notion of approximate structural representation or approximate continuity. However, Saatsi's (2005) appeal to explanatory approximate truth may not be of much help in this case, as it also seems to run afoul of the holistic nature of mathematical representation in the physical sciences.

doi: 10.48106/dial.v76.i1

of matter and the dynamics of light. In particular, the transition to quantum theory and quantum electrodynamics has redefined our fundamental understanding of the constitution of matter and the structure of light. In the transition to quantum theory, there can be no question that there has been a large-scale change in the account of the nature of light, not to mention the interaction between light and matter and the nature of observation. The structure of light depicted by these theories can still be defined in terms of a transverse wave equation, and one can find analogues of Fresnel's equations in many cases. However, the actual structures picked out by these equations are so different, given their theoretical setting, that they simply no longer constitute a description of the same physical structure in the world.

In addition, it is hard to see how the general structure described by these equations could account for any physical phenomena or support a robust realism in this case—at least in anything but a vacuous sense. It is for this reason that Ladyman and Ross (2007) frame their structural realism in terms of an account of modal structure, rather than the structure of a specific phenomenon. But even in this case, as I will argue in section 3, lingering concerns remain about whether they can account for the holistic nature of mathematical representation.

3 Ladyman and Ross, and the Problem of Situating Structures

Ladyman and Ross (2007) provide a compelling structuralist interpretation of the epistemology and metaphysics of modern science. Their structural realism is based on a model-theoretic account of scientific representation. On this view, scientific theories are taken to present a family of formal, i.e., mathematical, models, and these models are assumed to relate to nature through a structural similarity. Specifically, Ladyman and Ross defend a modal structural realism, through which parts of the mathematical structure of successful scientific theories are held to map onto the modal structure of reality.[14] In response to the pessimistic meta-induction, they argue (2007, 123) that the "idea that science describes the objective modal structure of the world is not undermined by theory change in the history of science, since all the

14 Here, I should highlight the radical ontological nature of their view. As Ladyman and Ross (2007, 130) characterize it: "Ontic Structural Realism (OSR) is the view that the world has an objective modal structure that is ontologically fundamental, in the sense of not supervening on the intrinsic properties of a set of individuals."

well-confirmed modal relations expressed by old theories are approximately recovered in their successors." In addition, to account for the no-miracles argument, they note (2007, 153) that if "science tells us about objective modal relations among the phenomena (both possible and actual), then occasional novel predictive success is not miraculous but to be expected."

In response to the concerns presented in section 2, Ladyman and Ross can simply acknowledge that the mathematical structure that Worrall highlights fails to map onto the physical structure of light. The problem was Worrall's narrow focus on the physical structure, rather than the modal structure, of nature. This modal structure is defined in terms of their account of "real patterns" in nature (e.g., see Dennett 1991; Ladyman and Ross 2007, 190–257). Following Dennett (1991), a pattern is termed "real" when we can make successful predictions concerning it. These "real patterns" are often identified through data models, which are taken to represent the underlying phenomena. The patterns within these models are real, in this sense, when they can be taken as a basis for predictions.

In the transition from Fresnel to Maxwell, we are no longer concerned with identifying the relevant physical structure that is responsible for the observed phenomena. Rather, we need to explicate the manner in which the patterns in the observable phenomena—i.e., the location of the diffracted, reflected, and refracted light—are accounted for in terms of the relevant modal structure, where for "modal" one could read "nomological" (Ladyman and Ross 2007, 130). The laws governing Fresnel and Maxwell's accounts of diffraction, reflection, and refraction, are, indeed, formally similar. They are expressed through the same mathematical structures, and these structures are both derived through an appeal to a similar set of principles (e.g., the conservation of energy, Huygen's principle, etc.). The remaining question is whether there is sufficient continuity in the broader formal and theoretical framework to actually show that the same modal structure in the world is identified and represented in the transition from Fresnel to Maxwell.

In response to the challenge posed by the holistic nature of the interpretation of mathematical structure in section 2, Ladyman and Ross can simply accept that, to a certain extent, we need to be more careful to articulate the structures that we presuppose in characterizing a mathematical representation. Of course, these presuppositions will certainly be weaker when we are only concerned with representing the general modal structure of reality. Here, we no longer face the challenge of situating a structure in a physical setting, but rather, situating a structure in a set of experimental data. Mathematical

doi: 10.48106/dial.v76.i1

modality simply needs to represent the physical modality in a given domain. All that is required, then, is that both the data model and the mathematical structure can be defined within the same basic theoretical formalism. Given that one can formulate both Fresnel's theory and Maxwell's theory within the context of classical mechanics, we can ensure that their mathematical structures are well-defined and can be formally related to one another, and, on the assumption that the data model is well-understood, the structural realist can simply define a mapping between the shared structure and the "real pattern" in nature.[15]

However, this solution does not entirely alleviate the challenge for the structural realist. There remains a concern with how we interpret the data produced in an experiment in the context of the mathematical structure of a physical theory. Recall that a data pattern is termed "real" when it can serve as the basis for successful predictions, but to make a prediction in a novel situation, we need to know the sense in which a given pattern is to be both located in nature and interpreted.

In response to the question of how "real patterns" are located in nature, prior to their representation, Ladyman and Ross (2007, 121) suggest that "one picks out a real pattern independently of its structural description by an ostensive operation—that is, by 'pointing at it.' " But here we should "think of 'pointing' as meaning 'directing a measurement apparatus.' " In this context, they (2007, 122) are quick to point out that they "are not suggesting that one begins by locating real patterns and then discovers their structural descriptions." Rather,

> Location is a recursive practice, and generally goes on against the background of some already developed structure. In practice, then, a locator will be a partial interpretation of a structure in the context of another, presupposed, structure. (2007, 122)

Here, Ladyman and Ross suggest that as theory progresses, it specifies the location of a "real pattern" with greater precision within the context of a "presupposed" structure that is developed through the refinement of empirical theory.

This is an important point. In the context of, say, modern particle physics, it is not sufficient to simply state the energy range in which one might encounter some novel structure—i.e., of where a real pattern may be located. To even

15 The hope would be that one could do something similar in the transition to quantum mechanics, as well.

understand the sense in which a given experiment provides a probe of a certain energy range, one must presuppose a vast theoretical framework to account for the operation of the measurement apparatus. Thus, in modern particle physics, one needs to be careful to clearly specify the relevant theoretical structure that is presupposed and its role in locating and interpreting the "real patterns" in nature. The challenge, in this case, is to delineate the sense in which the broader theoretical and formal framework of particle physics determines which patterns are real, i.e., detectable, and the manner in which they can serve as the basis for prediction. This is essential to both locate a real pattern and correlate it to the nomological structure of the standard model of particle physics.

In defence of their view, Ladyman and Ross (2007, 130–189) specifically appeal to modern particle physics, which they take to not only undermine the individuality of objects required by traditional scientific realism, but to also motivate their formal account of scientific understanding. It is clear, even to a casual observer, that modern particle physics is now largely based on a study of the abstract mathematical structures that characterize the natural world. Indeed, no pursuit better encapsulates the profound structural nature of modern theoretical physics better than the historical development and conceptual foundation of the standard model. Here, elementary "particles" are defined through the group-theoretic structures that characterize their properties. The standard model is a gauge theory—i.e., a theory through which one appeals to local symmetry structures to derive the relevant fields and their interactions. Thus, the nature of reality is described, at its fundamental level, through the structural relations it obeys.

The potential problem is that the standard model of particle physics has been tested in some of the most elaborate experiments ever devised. To even understand the output, i.e., data pattern, of one of the ATLAS detectors at the high-energy particle accelerator at CERN (the European Centre for Nuclear Research), we need to interpret the results within a broad theoretical framework that includes the standard model itself. The data produced from one of these detectors is so vast that it cannot possibly be processed. We must discard the overwhelming majority of it by an initial filtering, which is based upon theoretical expectations. This filtering process is guided by the standard model itself. But, more generally, the data itself cannot even be processed until it is "understood" through a theory that defines the detector function. This theoretical framework includes quantum field theory, non-relativistic quantum mechanics, solid state theory, electromagnetic theory, classical mechanics,

doi: 10.48106/dial.v76.i1

chemistry, and computational theory—just to name a few. In the context of this disparate and inconsistent theoretical framework, it is not always clear how exactly we should interpret the structure identified by the detector and the recursive theoretical process through which "real patterns" are precisely located and related to the modal structure of the standard model.

In the move from "physical" to "modal" structure (within the "real patterns" account), the hope was that the problem of situating structures "in nature" would be resolved. But situating a structure in a set of data may be no less problematic, and for the very same reasons. Once again, we can only interpret a data pattern within a theoretical and formal framework, and in the transition to a new theory, one will again face the same concerns relating to how we interpret these "real patterns" across inconsistent frameworks. It is not all clear the sense in which a "real pattern" from classical physics, or even non-relativistic particle physics, is approximately maintained in modern high-energy particle physics, given the vast theoretical change and the deeply theory-laden nature of the measurement procedure.

However, the hope may be that the theoretical assumptions that go into the location and interpretation of the data patterns produced by experiments in modern particle physics are so well-grounded, or general enough, that they will likely survive any future theory change—at least as an approximation. Indeed, there is a tradition in the philosophy of physics that has argued for the necessity of theory-laden experimentation, as an essential feature of scientific enquiry (e.g., Stein 1994; Smith 2014; Curiel 2019). In particular, Koberinski and Smeenk (2020) and Koberinski (2022) have brought attention to the fundamental role that the framework of modern quantum field theory plays in the precision tests of the standard model, and the search for new physics. They highlight the importance that this framework plays in securing theoretical continuity in the search for novel phenomena. But these merits presuppose that quantum field theory is itself on the right track—i.e., in the sense that it will be maintained as a low-energy approximation to whatever future theory replaces it. Thus, the continuity required by the modal "real patterns" account of structural realism may be easily secured, but only within the framework of quantum field theory. The concern is that this would pin structural realism (in the context of modern particle physics) to a particular "assumed structure." The modal structural realist would be required to presuppose the framework of quantum field theory to locate real patterns in nature. But this would sit uneasily with the structural realist response to the pessimistic induction on the history of scientific theory change. At the very least, these problems seem to pose a potential challenge to the structural realism of Ladyman and Ross (2007), and its subsequent defence (e.g., Ladyman 2011, 2017, 2018).[16]

In addition, these issues may reach beyond the historical motivation for structural realism, as they bring into question the manner in which the abstract formal structures of modern physics are related to reality, more generally. Although this paper has focused on only two articulations of the historical motivation for structural realism, the assumptions underwriting these positions are shared by a number of other variants of structural realism (e.g., see Frigg and Votsis 2011). The common assumption is that modern physical theory presents us with a family of models, or formal structures, and that the problem of realism can be solved if we can simply specify how these structures map onto nature. This "mapping" or "model-theoretic" account of structural realism (e.g., French 2014) has led to a profound new understanding of the nature of mathematical representation in physics (e.g., Bueno and French 2018), but it has yet to sufficiently articulate how the "structures" in nature are themselves individuated and identified (Stemeroff 2021). Thus, the concerns addressed in this paper may pose a general challenge to the modern structural realist, as they may need to pay closer attention to the practice of how the abstract structure of modern physics is related to the reality that it is taken to describe.

In this context, there lies a further problem concerning the consistency of modern science. Here, the issue is that the broader mathematical framework of high-energy particle physics is, itself, not even consistent.[17] The theory currently lacks a well-defined formulation. Given that the definition of a mathematical structure essentially depends on the formalism of a theory, it is unclear whether a mathematical structure within a poorly defined or inconsistent formalism can be said to represent a structure in nature. It is interesting to note that this is not a problem in quantum field theory alone. A pertinent example from the case study in section 2 is Fresnel's use of a flawed dynamical formalism (e.g., see Kelvin 1904, 424). In this context, it is also important to note that Kirchoff was actually the first to provide the viable formal foundation for Fresnel's diffraction integrals in the late nineteenth century. Before Kirchoff, the mathematics of Huygen's principle was not

16 This later work has served to further refine the metaphysics and justification of the position, but has largely maintained the "real patterns" account under consideration (e.g., see Ladyman 2018, 105).

17 This is not even to mention the manner in which this framework will be maintained in any subsequent theory.

doi: 10.48106/dial.v76.i1

even well-defined (e.g., see Buchwald 1989, 188). The structural realist can reformulate Fresnel's theory in the context of modern mathematical physics and relate it to a modern reformulation of Maxwell's theory. But this sort of formal inconsistency has been quite common in the history of science—e.g., one could argue that the entire field of mechanics dealt with poorly-defined structures before the calculus was reformulated and finally placed on a rigorous foundation. The concern is that our current physical theories may suffer the same fate, and we may have to concede that our theories will generally fail to specify well-defined structures from the perspective of future science.

4 Conclusion

The structural realist seems to face a challenge in accounting for the holistic nature of the interpretation of the mathematical structure of physical theory. To provide an interpretation of a mathematical structure, we need to specify the theoretical and formal framework required to give it meaning. The problem is that even when structures are maintained, their broader interpretations are often not. The case studies presented in the paper illustrate the need for a more refined structural realism, one that is able to present a viable account of how we interpret and situate the structures of a physical theory.*

Noah Stemeroff
0000-0003-1667-0677
University of Bonn
nstemero@uni-bonn.de

References

BOKULICH, Alisa and BOKULICH, Peter, eds. 2011. *Scientific Structuralism*. Boston Studies in the Philosophy of Science n. 281. Dordrecht: Springer, doi:10.1007/978-90-481-9597-8.

BRADING, Katherine and LANDRY, Elaine M. 2006. "Scientific Structuralism: Presentation and Representation." *Philosophy of Science* 73(5): 571–581, doi:10.1086/518327.

* I would like to thank Joseph Berkovitz, James Robert Brown, Michael Stuart, Stuart Presnell, Michael Miller, and the anonymous reviewers for their helpful comments on earlier drafts of this paper. I also want to thank the Azrieli Foundation and the Dan David Foundation for generously supporting this research.

Brown, James Robert. 2012. *Platonism, Naturalism and Mathematical Knowledge*. Routledge Studies in the Philosophy of Science n. 10. London: Routledge, doi:10.4324/9780203155356.

Buchwald, Jed Z. 1989. *The Rise of the Wave Theory of Light: Optical Theory and Experiment in the Early Nineteenth Century*. Chicago, Illinois: University of Chicago Press.

—. 2012. “Cauchy’s Theory of Dispersion Anticipated by Fresnel.” in *A Master of Science History: Essays in Honor of Charles Coulston Gillispie*, edited by Jed Z. Buchwald, pp. 399–416. Archimedes: New Studies in the History and Philosophy of Science and Technology n. 30. Dordrecht: Springer, doi:10.1007/978-94-007-2627-7_22.

Bueno, Otávio and French, Steven. 2018. *Applying Mathematics: Immersion, Inference, Interpretation*. Oxford: Oxford University Press, doi:10.1093/oso/9780198815044.001.0001.

Cao, Tian Yu. 1997. *Conceptual Development of 20th Century Field Theories*. Cambridge: Cambridge University Press, doi:10.1017/cbo9780511563997.

—. 2003. “Structural Realism and the Interpretation of Quantum Field Theory.” *Synthese* 136(1): 3–24, doi:10.1023/a:1024139713910.

Chakravartty, Anjan. 2007. *A Metaphysics for Scientific Realism: Knowing the Unobservable*. Cambridge: Cambridge University Press, doi:10.1017/cbo9780511487354.

Curiel, Erik. 2019. “Schematizing the Observer and the Epistemic Content of Theories.” ArXiv preprint, doi:10.48550/arXiv.1903.02182.

Darrigol, Olivier. 2012. *A History of Optics: From Greek Antiquity to the Nineteenth Century*. Oxford: Oxford University Press.

Dennett, Daniel C. 1991. “Real Patterns.” *The Journal of Philosophy* 88(1): 27–51, doi:10.2307/2027085.

Duhem, Pierre. 1914. *La théorie physique, son objet, sa structure*. 2nd ed. Paris: Chevalier et Rivière.

—. 1991. *The Aim and Structure of Physical Theory*. Princeton, New Jersey: Princeton University Press. With a new introduction by Jules Vuillemin; foreword by Louis de Broglie; translation of Duhem (1914) by Philip P. Wiener, doi:10.2307/j.ctv1nj34vm.

Esfeld, Michael. 2013. “Ontic Structural Realism and the Interpretation of Quantum Mechanics.” *European Journal for Philosophy of Science* 3(1): 19–32, doi:10.1007/s13194-012-0054-x.

—. 2017. “How to Account for Quantum Non-Locality: Ontic Structural Realism and the Primitive Ontology of Quantum Physics.” *Synthese* 194(7): 2329–2344, doi:10.1007/s11229-014-0549-4.

Esfeld, Michael and Lam, Vincent. 2009. “Holism and Structural Realism.” in *Worldviews, Science and Us: Studies of Analytical Metaphysics. A Selection of Topics from a Methodological Perspective – Proceedings of the 5th Metaphysics of Science Workshop, Ghent, Belgium, 2-3 June 2005*, edited by Robrecht Vanderbeeken and Bart

D'HOOGE, pp. 10–31. Singapore: World Scientific Publishing Co., doi:10.1142/9789814299053_0003.

—. 2011. "Ontic Structural Realism as a Metaphysics of Objects." in *Scientific Structuralism*, edited by Alisa BOKULICH and Peter BOKULICH, pp. 143–160. Boston Studies in the Philosophy of Science n. 281. Dordrecht: Springer, doi:10.1007/978-90-481-9597-8_8.

FRENCH, Steven. 2014. *The Structure of the World: Metaphysics and Representation*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199684847.001.0001.

FRENCH, Steven and LADYMAN, James. 2003. "Remodelling Structural Realism: Quantum Physics and the Metaphysics of Structure." *Synthese* 136(1): 31–56, doi:10.1023/a:1024156116636.

—. 2011. "In Defence of Ontic Structural Realism." in *Scientific Structuralism*, edited by Alisa BOKULICH and Peter BOKULICH, pp. 25–42. Boston Studies in the Philosophy of Science n. 281. Dordrecht: Springer, doi:10.1007/978-90-481-9597-8_2.

FRIGG, Roman and VOTSIS, Ioannis. 2011. "Everything You Always Wanted to Know About Structural Realism but Were Afraid to Ask." *European Journal for Philosophy of Science* 1(2): 227–276, doi:10.1007/s13194-011-0025-7.

KELVIN, William Thomson. 1904. *Baltimore Lectures on Molecular Dynamics and the Wave Theory of Light*. London: C.J. Clay and Sons, ark:/13960/t04x5581c.

KOBERINSKI, Adam. 2022. "Fundamental 'Constants' and Precision Tests of the Standard Model." *Philosophy of Science* 89(5): 1255–1264, doi:10.1017/psa.2022.41.

KOBERINSKI, Adam and SMEENK, Christopher. 2020. "Q.E.D., QED." *Studies in History and Philosophy of Science. Part B: Studies in History and Philosophy of Modern Physics* 71: 1–13, doi:10.1016/j.shpsb.2020.03.003.

LADYMAN, James. 1998. "What is Structural Realism?" *Studies in History and Philosophy of Science* 29(3): 409–424, doi:10.1016/s0039-3681(98)80129-5.

—. 2011. "Structural Realism versus Standard Scientific Realism: the Case of Phlogiston and Dephlogisticated Air." *Synthese* 180(2): 87–101, doi:10.1007/s11229-009-9607-8.

—. 2017. "An Apology for Naturalized Metaphysics." in *Metaphysics and the Philosophy of Science: New Essays*, edited by Matthew H. SLATER and Zanja YUDELL, pp. 141–162. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199363209.003.0008.

—. 2018. "Scientific Realism Again." *Spontaneous Generations: A Journal for the History and Philosophy of Science* 9(1): 99–107, doi:10.4245/sponge.v9i1.29356.

LADYMAN, James and ROSS, Don. 2007. *Every Thing Must Go: Metaphysics Naturalized*. Oxford: Oxford University Press. With David Spurrett and John Collier, doi:10.1093/acprof:oso/9780199276196.001.0001.

LAUDAN, Laurens [Larry]. 1981. "A Confutation of Convergent Realism." *Philosophy of Science* 48(1): 19–49, doi:10.1086/288975.

Lyre, Holger. 2004. "Holism and Structuralism in $U(1)$ Gauge Theory." *Studies in History and Philosophy of Science. Part B: Studies in History and Philosophy of Modern Physics* 35(4): 643–670, doi:10.1016/j.shpsb.2004.07.004.

Maxwell, James Clerk. 1867–1868. "On the Best Arrangement for Producing a Pure Spectrum on a Screen." *Proceedings of the Royal Society of Edinburgh* 6: 238–242. Reprinted in Maxwell (1890, 96–100), doi:10.1017/S0370164600045879.

—. 1890. *The Scientific Papers of James Clerk Maxwell. Volume II*. Newcastle upon Tye: Cambridge Scholars Publishing. Edited by W.D. Niven.

Morganti, Matteo. 2004. "On the Preferability of Epistemic Structural Realism." *Synthese* 142(1): 81–107, doi:10.1023/b:synt.0000047712.39407.c3.

—. 2011. "Is there a Compelling Argument for Ontic Structural Realism?" *Philosophy of Science* 78(5): 1165–1176, doi:10.1086/662258.

Pincock, Christopher. 2005. "A Role for Mathematics in the Physical Sciences." *Noûs* 41(2): 253–275, doi:10.1111/j.1468-0068.2007.00646.x.

—. 2012. *Mathematics and Scientific Representation*. New York: Oxford University Press, doi:10.1093/acprof:oso/9780199757107.001.0001.

Psillos, Stathis. 1995. "Is Structural Realism the Best of Both Worlds?" *Dialectica* 49(1): 15–46, doi:10.1111/j.1746-8361.1995.tb00113.x.

—. 1999. *Scientific Realism: How Science Tracks the Truth*. London: Routledge, doi:10.4324/9780203979648.

Putnam, Hilary. 1975. *Mathematics, Matter and Method. Philosophical Papers, Volume 1*. Cambridge: Cambridge University Press. Second edition: Putnam (1979).

—. 1979. *Mathematics, Matter and Method. Philosophical Papers, Volume 1*. 2nd ed. Cambridge: Cambridge University Press. First edition: Putnam (1975), doi:10.1017/cbo9780511625268.

Saatsi, Juha. 2005. "Reconsidering the Fresnel-Maxwell Theory Shift: How the Realist can Have Her Cake and Eat It Too." *Studies in History and Philosophy of Science* 36(3): 509–538, doi:10.1016/j.shpsa.2005.07.007.

—. 2019. "What is Theoretical Progress of Science?" *Synthese* 196(2): 611–631, doi:10.1007/s11229-016-1118-9.

Smith, George E. 2014. "Closing the Loop: Testing Newtonian Gravity, Then and Now." in *Newton and Empiricism*, edited by Zvi Biener and Eric Schliesser, pp. 262–351. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199337095.003.0011.

Stein, Howard. 1989. "Yes, but ...– Some Skeptical Remarks on Realism and Anti-Realism." *Dialectica* 43(1/2): 47–66, doi:10.1111/j.1746-8361.1989.tb00930.x.

—. 1994. "Some Reflections on the Structure of our Knowledge in Physics." in *Logic, Methodology and Philosophy of Science IX: Proceedings of the Ninth International Congress of Logic, Methodology, and Philosophy of Science, Uppsala, Sweden, August 7-14, 1991*, edited by Dag Prawitz, Brian Skyrms, and Dag Westerståhl, pp. 633–655. Studies in Logic and the Foundations of Mathematics n. 134. Amsterdam: Elsevier Science Publishers B.V., doi:10.1016/s0049-237x(06)80067-4.

Stemeroff, Noah. 2021. "Structuralism and the Conformity of Mathematics and Nature." *Studies in History and Philosophy of Science* 86: 84–92, doi:10.1016/j.shpsa.2021.01.004.

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

—. 2019. "Towards a Realistic Success-to-Truth Inference for Scientific Realism." *Synthese* 196(2): 571–585, doi:10.1007/s11229-016-1150-9.

Votsis, Ioannis. 2005. "The Upward Path to Structural Realism." *Philosophy of Science* 72(5): 1361–1372, doi:10.1086/508974.

—. 2017. "Structural Realism and its Variants." in *The Routledge Handbook of Scientific Realism*, edited by Juha Saatsi, pp. 108–119. Routledge Handbooks in Philosophy. London: Routledge, doi:10.4324/9780203712498-10.

Votsis, Ioannis and Schurz, Gerhard. 2012. "A Frame-Theoretic Analysis of Two Rival Conceptions of Heat." *Studies in History and Philosophy of Science* 43(1): 105–114, doi:10.1016/j.shpsa.2011.10.010.

Whittaker, Edmund Tayler. 1951. *A History of the Theories of Aether and Electricity. Volume 1: The Classical Theories*. 2nd ed. London: Thomas Nelson and Sons Ltd.

Worrall, John. 1989. "Structural Realism: The Best of Both Worlds?" *Dialectica* 43(1/2): 99–124, doi:10.1111/j.1746-8361.1989.tb00933.x.

Wright, Aaron Sidney. 2017. "Fresnel's Laws, *ceteris paribus*." *Studies in History and Philosophy of Science* 64: 38–52, doi:10.1016/j.shpsa.2017.07.008.

Zahar, Elie G. 2007. *Why Science Needs Metaphysics: A Plea for Structural Realism*. LaSalle, Illinois: Open Court Publishing Co.

Focus Effects in Number Sentences Revisited

Katharina Felka

There are easy arguments for numbers: Arguments that derive the existence of numbers in a few, simple steps from uncontroversial premises like the premise that I have ten fingers. In recent literature some authors have argued that easy arguments rely on a mistaken linguistic analysis of number sentences like ‘The number of my fingers is ten’: While such sentences are traditionally considered as identity sentences, they are rather specificational sentences. However, Barlew (2017) has disputed this line of argument: He argues that *in easy argument contexts* the pertinent number sentences function as identity sentences even though they function as specificational sentences *in their standard use*. Hence, Barlew concludes, the rebuttal of easy arguments fails. The aim of the present paper is to defend the linguistic objection to easy arguments against Barlew’s criticism.

When philosophers discuss whether numbers exist, they usually assume that they discuss a hard question that does not have an easy answer. However, surprisingly, there seem to be very easy arguments for the existence of numbers. Just look! I have ten fingers. If I have ten fingers, then the number of my fingers is ten. Hence, there is a number! Or look at my legs! I have two legs. If I have two legs, then the number of my legs is two. Hence, there is a number! In such arguments the existence of numbers is derived from completely uncontroversial premises, like the premise that I have ten fingers or that I have two legs. That makes the arguments very puzzling: How can it be that philosophers have discussed for thousands of years whether numbers exist if the existence of numbers can be derived from completely uncontroversial premises in a few, simple steps?

In recent literature some authors have argued that easy arguments fail to establish the existence of numbers on linguistic grounds. They argue that easy arguments rely on a mistaken linguistic analysis of number sentences like

'The number of my fingers is ten' or 'The number of my legs is two': While such sentences are traditionally considered as identity sentences in which the number words 'ten' and 'two' appear in singular term position, they are rather specificational sentences in which the number words appear in determiner position.[1] However, in a recent paper Barlew (2017) has disputed this line of argument: He argues that *in easy argument contexts* the pertinent number sentences do function as identity sentences even though they function as specificational sentences *in their standard use*. Hence, Barlew concludes, the rebuttal of easy arguments fails. The aim of the present paper is to defend the linguistic objection to easy arguments against Barlew's criticism.

The structure of the paper is as follows. Section 1 sketches the linguistic objection against easy arguments. Section 2 presents Barlew's (2017) argument to the effect that number sentences function as identity sentences rather than as specificational sentences in easy argument contexts, in contrast to what opponents of easy arguments have claimed. Section 3 argues that Barlew's argument fails and, thus, that it is warranted to object to easy arguments on linguistic grounds.

1 A Rebuttal of Easy Arguments

Paradigmatic easy arguments start from a fairly uncontroversial assumption that does not say anything about numbers. For instance, it is commonly assumed that Mars has two moons and, thus, that sentence (1) is true:

(1) Mars has two moons.

If sentence (1) is true, then sentence (2) is true as well:

(2) The number of moons of Mars is two.

But, so the argument goes, sentence (2) is true only if numbers exist. Hence, numbers exist!

Apart from the assumption that sentence (1) is true, the argument relies on the following two assumptions:

(P1) If sentence (1) is true, then sentence (2) is true.

1 See Felka (2014, 2016) and Moltmann (2013). The first elaborated criticism of the traditional analysis of the pertinent number sentences, however, is due to Hofweber (2005). But, in contrast to Felka and Moltmann, he does not defend a specificational analysis of those sentences.

(P2) The truth of sentence (2) requires the existence of numbers.

(P2) is based on a certain linguistic analysis of the pertinent number sentence that was most famously proposed by Gottlob Frege. In his *Foundations of Arithmetics* Frege writes:

> [T]he proposition ‘Jupiter has four moons’ can be converted into ‘the number of moons of Jupiter is four’. Here the word ‘is’ should not be taken as a mere copula, as in the proposition ‘the sky is blue’ [...] Here ‘is’ has the sense of ‘is identical with’ or ‘is the same as’. (Frege 1980, para. 57)

Frege, thus, assumes the following:

> ID. Sentences of the form ‘The number of *F*s is *n*’, where ‘*n*’ is a placeholder for a number word, are identity sentences in which ‘*n*’ functions as a singular term.

(ID) Sentences of the form ‘The number of *F*s is *n*’, where ‘*n*’ is a placeholder for a number word, are identity sentences in which ‘*n*’ functions as a singular term.

If (ID) is correct, then the number word ‘two’ contained in sentence (2) functions as a singular term. Since sentences containing singular terms can be true only if the singular terms refer, (2) can be true only if numbers exist.

However, in recent literature some authors have rejected (ID) (Felka 2014, 2016; Hofweber 2007, 2016; Moltmann 2013). Some of them have argued that sentence (2) is a so-called specificational sentence while specificational sentences are the elliptical remainders of question-answer pairs (Felka 2014, 2016; Moltmann 2013). According to this analysis, sentence (2) is analysed as follows:[2]

(2*) [~~What~~ the number of moons of Mars ~~is~~] is [~~Mars has~~ two ~~moons~~.]

If this analysis is correct, then the number word ‘two’ is the elliptical remainder of sentence (1). Since the number word functions in sentence (1) as a determiner, it functions in sentence (2) as a determiner as well. Hence, it does

2 Following Barlew, I focus here on the question-answer analysis proposed in Felka (2014, 2016). See Moltmann (2013) for a different variant. For the present discussion it does not matter what specificational analysis we rely on.

doi: 10.48106/dial.v76.i1

not function as a singular term and, thus, does not bring it about that the truth of sentence (2) requires the existence of numbers, as proponents of easy arguments assume.[3]

2 Barlew's Defence of Easy Arguments

Barlew (2017) concedes that number sentences of the form 'The number of *F*s is *n*' function as specificational sentences *in their standard use*. However, he argues that *in easy argument contexts* the number sentences function as identity sentences and, thus, that easy arguments go through. In the following I will first explain a distinction between narrow and broad focus on which Barlew relies in his argument and then explain how he uses this distinction to establish his claim.

2.1 *Narrow and Broad Focus*

Intuitively, the focus of an utterance of a sentence is that part of information conveyed with the utterance that is most important in the utterance context.[4] Take, for instance, the sentence

(3) Paul shattered the china.

When the question under discussion is 'Who shattered the china?', the focus is on the information provided by 'Paul'. When the question under discussion is 'What did Paul shatter?', the focus is on the information provided by 'the china'.

There are different ways to mark the focus of an utterance. Firstly, we can mark it by putting intonational stress on some part of the utterance (here marked with bold letters):

(4) PAUL shattered the china.
(5) Paul shattered THE CHINA.

3 One might argue that the definite description still induces that (2) is true only if numbers exist. However, it has been argued that it only induces a *pragmatic* presupposition and, thus, that 'Mars has two moons' can be a true answer to the question even if numbers do not exist (see Felka 2016; and Brogaard 2007 for further discussion).

4 See Hofweber (2016) for a more detailed explanation as well as the pertinent references from the linguistic literature.

(4) marks the information provided by 'Paul' as the focus of the utterances; (5) the information provided by 'the china'. Secondly, we can mark the focus of an utterance by choosing a specific syntactic structure. Consider:

(6) It was the china that Paul shattered.

(6) marks the information provided by 'the china' as the focus of the utterance due to its syntactic structure. Sentences that exhibit such an intonation-independent structural focus are called *focus constructions*.

A striking feature of focus constructions is that they give rise to a specific question-answer behaviour which allows us to check (i) whether some sentence is a focus construction and (ii) what part exactly carries the information marked as the focus. In relation to (i), consider the following exchanges:

(7) Who shattered the china? # It was the china that Paul shattered.
(8) What did Paul shatter? It was the china that Paul shattered.

The question-answer behaviour of (6) makes obvious that the sentence marks the information provided by 'the china' as the focus. For since this information is marked as the focus and, thus, as particularly important, the sentence cannot felicitously be uttered to answer the first question that does not ask about it. In contrast, it can felicitously be uttered to answer the second question. In relation to (ii), notice that the expression that carries the information marked as the focus constitutes an appropriate short answer to question (9):

(9) What did Paul shatter?
(6) It was the china that Paul shattered.
(10) The china.

Thus, we can check what short answers are appropriate in order to determine what expression exactly carries the information marked as the focus.

The example sentence considered above is a case of narrow focus in which the focus is on a *single* constituent ('the china'). Barlew points out that there are also cases of broad focus in which the focus is on the complete utterance. For illustration, consider a context in which (11) ist the question under discussion:

(11) What happened?
(3) Paul shattered the china.
(10) # The china.

doi: 10.48106/dial.v76.i1

In this utterance context the focus of an utterance of sentence (3) is not on a single constituent like ‘the china’. Rather, it is on the complete utterance. Accordingly, no single constituent will be an appropriate short answer to the question under discussion; we have to utter the complete sentence to answer the question appropriately. This is a case of broad focus.

2.2 *The Number Sentences in Easy Argument Contexts*

Both opponents of easy arguments and their critic Barlew assume that specificational sentences are copular sentences that are distinguished by exhibiting a structural focus on the post-copular expression.[5] They also agree that at least in their standard use number sentences of the form ‘The number of *F*s is *n*’ exhibit a structural focus on the post-copular expression.[6] The latter claim is based on the question-answer behaviour of the number sentences. Consider:

(12) Who has ten fingers? # The number of my fingers is ten.

(13) What is the number of your fingers? The number of my fingers is ten. // Ten.

An utterance of the number sentence (or simply the number word ‘ten’) is an appropriate answer to a question that asks about the information provided by the number word while it is not an appropriate answer to a question that does not ask about that information. Since this is to be expected if the number sentence exhibits a structural focus on the post-copular expression, both opponents of easy arguments and Barlew assume that the sentence exhibits such a focus and, thus, functions as a specificational sentence in its standard use.

However, following Higgins (1973) and others, Barlew points out that many copular sentences allow for different uses. Therefore, Barlew says, it is “essential to determine which reading of [the number sentence] arises” in easy argument contexts (Barlew 2017, 421). According to Barlew, easy argument contexts are not “contexts where the interlocutors are wondering about numbers of moons or planets” since “a philosopher making the easy argument doesn’t actually care how many moons [Mars] has” (Barlew 2017, 421). Rather,

5 See, e.g., Higgins (1973), Heycock (1995), Mikkelsen (2005) for this view. In the philosophical literature, a detailed defence can be found in Felka (2014, 2016).

6 This observation is due to Hofweber (2005) and is employed in Felka (2014, 2016) to argue for the claim that the pertinent number sentences function as specificational sentences.

they are contexts in which philosophers discuss the entailments of ontologically innocent sentences like ‘Mars has two moons’.[7] Thus, Barlew says, we have to determine how number sentences of the pertinent kind are used in contexts in which philosophers discuss the entailments of ontologically innocent sentences.

In order to do so, Barlew presents the following example of such a context:

(C) Al and Betty are philosophers. Al is also an amateur astronomer with a decent telescope but not much background knowledge. After a night of star gazing Al tells Betty: “Guess what, Mars has two moons.” Betty replies: “Hmm, I wonder what we can infer from this, or what other sentences we might say that are true in virtue of this.”

According to Barlew, this is an easy argument context since the question under discussion is (14):

(14) What are the entailments of ‘Mars has two moons’?

However, Barlew observes, an utterance of the number word ‘two’ is not an appropriate answer to the question under discussion, while an utterance of the complete sentence (2) is:

(2) The number of moons of Mars is two.
(15) # Two.

Thus, Barlew concludes, in the present context the focus is not on the number word ‘two’ (or any other single constituent); rather, the focus is on the complete utterance. We thus have a case of broad focus, rather than a case of narrow focus on the number word (or any other constituent of the sentence).

If Barlew’s consideration were correct, it would present a major difficulty for the objection to easy arguments presented above. As we have seen, the objection crucially relies on the claim that number sentences of the form ‘The number of *F*s is *n*’ are specificational sentences. But if in easy argument contexts the number sentences do not exhibit narrow focus on the post-copular term, they do not function as specificational sentences in such contexts. Rather, they function as identity sentences, just like proponents of easy arguments assume.

7 The distinction between “ontologically innocent” and “ontologically loaded” sentences is due to Hofweber (2007).

doi: 10.48106/dial.v76.i1

3 Rebuttal of Barlew's Defence

Barlew's defence of easy arguments is successful only if he manages to establish (i) that the (allegedly special) philosophical use of the sentence 'The number of moons of Mars is two' he considers is the one pertinent for easy arguments and (ii) that the sentence functions as an identity sentence in that use. In the following I will argue that Barlew fails on both counts.

3.1 *What Are the Pertinent Uses of the Number Sentences?*

As presented above, Barlew concedes that the number sentence 'The number of moons of Mars is two' functions as a specificational sentence in its standard use. But, he argues, in the uses pertinent for easy arguments the sentence functions as an identity sentence and, thus, the arguments go through. According to Barlew, the pertinent uses are uses of the sentence in contexts in which metaphysicians are concerned with the entailments of ontologically innocent sentences rather than with astronomical facts concerning Mars and its moons. That is, they are uses in *philosophical* rather than in *ordinary* contexts.

However, Barlew's assumption that easy arguments target uses of the number sentence in philosophical contexts is mistaken. There certainly are contexts in which metaphysicians discuss entailments of ontologically innocent sentences rather than astronomical facts concerning Mars and its moons. And in these contexts metaphysicians are concerned with uses of number sentences. But this does *not* imply that the uses of number sentences they discuss are uses in philosophical contexts: Surely, in a given context C_1, one can discuss features of sentences (including their apparent entailments) *as they are used in a different context* C_2. And this is exactly what is going on in easy argument contexts: In such contexts, metaphysicians discuss features of number sentences as they are used by ordinary speakers in non-philosophical contexts. Proponents of easy arguments take every opportunity to emphasise this. Here is a representative quotation from Thomasson:

> [...] the relevant conditions of existence are determined by the application [...] conditions for the terms speakers use [...] the truths [...] uncovered by metaphysicians are just ways of making explicit the ontological implications of the rules we master in learning to use expressions. (2009, 450)

As Thomasson emphasises in this quotation, in easy argument contexts metaphysicians take expressions in their standard use by ordinary speakers and investigate their existence entailments in that very use.

Barlew might want to try the following defence strategy:

> It is correct that proponents of easy arguments like Thomasson focus on standard uses of number sentences by ordinary speakers. However, a more successful strategy to argue for the existence of numbers in an easy way is to focus on philosophical uses of such sentences since philosophical uses of number sentences are identity rather than specificational uses.

The next subsection shows that this defence strategy fails as well, since Barlew is unable to establish that the philosophical use of the number sentence he considers is a non-standard identity rather than a standard specificational use.

3.2 *A Case of Broad Focus?*

Let us consider whether Barlew has established that the philosophical use of the number sentence he considers is a non-standard identity rather than a standard specificational use. Recall that in the context he presents the question under discussion is supposed to be (14):

(14) What are the entailments of 'Mars has two moons'?

To this question, Barlew claims, sentence (2) is an appropriate answer:

(2) The number of moons of Mars is two.

This could not be the case if the sentence were exhibiting a structural focus on the number word 'two' since then an utterance of the sentence could only be an appropriate answer to a question that asks about the information provided by the number word. Thus, Barlew says, the sentence does not exhibit such a focus and, hence, does not function as a specificational sentence since such sentences are distinguished by exhibiting a structural focus on the post-copular term.

However, Barlew's claim that sentence (2) is an appropriate answer to the question under discussion is mistaken. For the question under discussion requires *examples of sentences*. In particular, it requires examples of sentences that are entailed by the sentence ‘Mars has two moons’. But an utterance of sentence (2) does not give an example of such a sentence: An utterance of sentence (2) does not say anything about sentences or other linguistic expressions; it only says something about Mars and its moons. Therefore, it does not constitute an answer to the question under discussion. In contrast, an utterance of sentence (2_{Q}) does constitute an answer to the question under discussion:

(2_{Q}) ‘The number of moons of Mars is two.’

An utterance of sentence (2_{Q}) is the short version of the following complete answer to the question under discussion, which, indeed, is also an appropriate answer to (14):

(2_{L}) ‘Mars has two moons’ entails ‘The number of moons of Mars is two’.

But from the observation that (2_{L}) constitutes an appropriate answer to the question under discussion we cannot derive anything about the information structure of some other sentence. In particular, we cannot derive anything about the information structure of sentence (2), with which opponents of easy arguments are concerned.

To drive my point home, consider the following argument that is analogous to the one that Barlew presents. In the previous section we considered the sentence ‘It was the china that Paul shattered’ as an example of a focus construction that marks the information provided by ‘the china’ as the focus. One might now try to establish that in some contexts the sentence does not mark the information provided by ‘the china’ as the focus. Take, for instance, a context in which the question under discussion is (16):

(16) What is an example of a focus construction?

To this question, one might argue, an utterance of sentence (6) is an appropriate answer while an utterance of (10) is not:

(6) It was the china that Paul shattered.
(10) # The china.

Thus, so the argument would go, the sentence ‘It was the china that Paul shattered’ does not mark the information provided by ‘the china’ as the focus in the present context since then an utterance of the sentence could only be an appropriate answer to a question that asks about that information. But, again, the argument fails since it relies on the mistaken assumption that an utterance of sentence (6) is an appropriate answer to the question under discussion while in fact only an utterance of sentence (17) or of its short version (18) is—and it is exactly *since* sentence (6) marks the information provided by 'the china' as the focus:

(17) An example of a focus construction is 'It was the china that Paul shattered'.
(18) 'It was the china that Paul shattered.'

For the very same reason Barlew's argument fails to establish that the sentence 'The number of moons of Mars is two' does not mark the information provided by 'two' as the focus in the specified context. Therefore, it also fails to establish that sentence (2) functions as an identity sentence in that context.

Let me finally point out that Barlew might try to rescue his point by modifying the question under discussion such that it does not ask for examples of sentences anymore. For instance, the question could also be:

(19) What follows from the fact that Mars has two moons?

However, an utterance of sentence (2) is not an appropriate answer to this question either; eventually only '(From the fact that Mars has two moons it follows that) the number of moons of Mars is two' is. But, again, from the observation that the latter sentence is an appropriate answer to question (19) we cannot derive anything about the information structure of sentence (2). The same holds for every other question one might want to try to bring to Barlew's rescue I can think of. I thus conclude that Barlew's argument fails.

4 A More General Reply

Finally, let me give a more general reply to Barlew's criticism that is independent of the subtleties of the previous discussion. As pointed out at the outset, our easy argument relies on the following two premises:

(P1) If sentence (1) is true, then sentence (2) is true.
(P2) The truth of sentence (2) requires the existence of numbers.

doi: 10.48106/dial.v76.i1

Proponents of easy arguments usually rely on standard uses of sentence (2) in the premises (P1) and (P2). But if they rely on standard uses, then premise (P2) fails. For the justification of (P2) is based on the assumption that number sentences like (2) are identity sentences. But in their standard use such sentences function as specificational rather than as identity sentences.

One may now follow Barlew and try to argue that there are also non-standard—perhaps special philosophical—uses of the number sentence in which it does function as an identity sentence. Indeed, one may simply stipulate that one takes the sentence in the sense of 'The number of moons of Mars = the number two'. But if proponents of easy arguments rely on such a special non-standard use of the sentence, then premise (P1) of the easy argument becomes highly unobvious. For the justification of premise (P1) is based on the observation that *ordinary speakers* take the two sentences to be truth-conditionally equivalent. Since the pertinent speakers' intuition relies on standard uses of the number sentences, premise (P1) loses its justification if one does not rely on such uses.

Thus, if one agrees that number sentences like (2) function as specificational sentences in their standard use (like Barlew does), then it does not matter whether there are any further non-standard uses of the sentences in which they function as identity sentences. For if one relies on such non-standard uses, then premise (P1) of the easy argument loses its justification and the argument fails nevertheless.

5 Conclusion

Barlew recently argued that in easy argument contexts number sentences like 'The number of moons of Mars is two' are used in a non-standard way: They are used as identity rather than as specificational sentences. Thus, Barlew claims, a rebuttal of easy arguments on linguistic grounds is unconvincing. The present paper defended the linguistic objection to easy arguments against Barlew's criticism. In particular, it has been argued that (i) the uses that are pertinent for easy arguments are standard uses and (ii) Barlew's considerations do not even show that there are non-standard uses of the number sentences in which they function as identity sentences. Since Barlew's defence of easy arguments thus fails, the linguistic objection against easy arguments stands. Arguing from 'Mars has two moons' to 'The number of moons of Mars is two' is no quick and easy way to establish the existence of numbers, since such an argument has to rely on a mistaken linguistic analysis of the pertinent number sentence.*

Katharina Felka
0000-0002-4921-8815
University of Graz
katharina.felka@uni-graz.at

References

Barlew, Jefferson. 2017. “Focus on Numbers.” *Linguistics and Philosophy* 40(4): 401–426, doi:10.1007/s10988-017-9217-8.

Brogaard, Berit. 2007. “Number Words and Ontological Commitment.” *The Philosophical Quarterly* 57(226): 1–20, doi:10.1111/j.1467-9213.2007.465.x.

Felka, Katharina. 2014. “Number Words and Reference to Numbers.” *Philosophical Studies* 168(1): 283–294, doi:10.1007/sl1098-013-0129-3.

—. 2016. *Talking about Numbers: Easy Arguments for Mathematical Realism*. Studies in Theoretical Philosophy n. 3. Frankfurt a.M.: Vittorio Klostermann, doi:10.5771/9783465138792.

Frege, Gottlob. 1884. *Die Grundlagen der Arithmetik: Eine logisch-mathematische Untersuchung über den Begriff der Zahl*. Breslau: Wilhelm Koebner. Reissued as Frege (1961).

—. 1950. *The Foundations of Arithmetic*. Oxford: Basil Blackwell Publishers. Translation of Frege (1884) by J.L. Austin, doi:10.4324/9781003072874.

—. 1961. *Die Grundlagen der Arithmetik: Eine logisch-mathematische Untersuchung über den Begriff der Zahl*. Hildesheim: Georg Olms.

—. 1980. *The Foundations of Arithmetic*. 2nd ed. Evanston, Illinois: Northwestern University Press. Translation of Frege (1884) by J.L. Austin; first edition: Frege (1950).

Heycock, Caroline. 1995. “The Internal Structure of Small Clauses: New Evidence from Inversion.” in *Proceedings of the North East Linguistic Society 25, University of Pennsylvania. Volume One: Papers from the Main Sessions*, edited by Jill N. Beckman, pp. 223–238. Amherst, Massachusetts: University of Massachusetts, Department of Linguistics.

Higgins, Francis Roger. 1973. “The Pseudo-Cleft Construction in English.” PhD dissertation, Cambridge, Massachusetts: Massachusetts Institute of Technology. Republished as Higgins (1976) and Higgins (1979).

* Many thanks to the audience of the CLLAM meeting in Stockholm, Alex Steinberg as well as two anonymous referees for helpful discussions and comments.

—. 1976. *The Pseudo-Cleft Construction in English*. 310 Lindley Hall, Indiana University, Bloomington, Indiana 46405: Indiana University Linguistics Club. Republished as Higgins (1979).

—. 1979. *The Pseudo-Cleft Construction in English*. New York: Garland Publishing Co., doi:10.4324/9781315693545.

Hofweber, Thomas. 2005. "Number Determiners, Numbers, and Arithmetic." *The Philosophical Review* 114(2): 179–225, doi:10.1215/00318108-114-2-179.

—. 2007. "Innocent Statements and their Metaphysically Loaded Counterparts." *Philosophers' Imprint* 7(1), http://hdl.handle.net/2027/spo.3521354.0007.001.

—. 2016. *Ontology and the Ambitions of Metaphysics*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780198769835.001.0001.

Mikkelsen, Line. 2005. *Copular Clauses: Specification, Predication and Equation*. Amsterdam: John Benjamins Publishing Co., doi:10.1075/la.85.

Moltmann, Friederike. 2013. "Reference to Numbers in Natural Language." *Philosophical Studies* 162(3): 499–536, doi:10.1007/s11098-011-9779-1.

Thomasson, Amie L. 2009. "Answerable and Unanswerable Questions." in *Metametaphysics: New Essays on the Foundations of Ontology*, edited by David J. Chalmers, David Manley, and Ryan Wasserman, pp. 444–471. Oxford: Oxford University Press, doi:10.1093/oso/9780199546046.003.0015.

What Is the Point of Persistent Disputes?

The Meta-Analytic Answer

Alexandre Billon & Philippe Vellozzo

Many philosophers regard the persistence of philosophical disputes as symptomatic of overly ambitious, ill-founded intellectual projects. There are indeed strong reasons to believe that persistent disputes in philosophy (and more generally in the discourse at large) are pointless. We call this the pessimistic view of the nature of philosophical disputes. In order to respond to the pessimistic view, we articulate the supporting reasons and provide a precise formulation in terms of the idea that the best explanation of persistent disputes entails that they are pointless. We then show how to answer the pessimistic argument. Taking a well-known mathematical controversy as our paradigm example, we argue that some persistent disputes reflect substantive disagreements at the "meta-analytic" level, i.e., disagreements about the best way, among quite different candidates, to understand the topic at issue, and the best associated cluster of analytic truths one should accept concerning it. Moreover, our concrete example shows that such meta-analytic disagreements can, in principle, be settled and yield a genuine theoretical (as opposed to merely pragmatic) breakthroughs. We conclude optimistically that persistent disputes can be an important means of fostering epistemic progress.

It is commonplace to observe that people tend to disagree and argue about a multitude of issues, from the most trivial to the most sophisticated. Some disputes last but briefly, others endure for more significant lengths of time (sometimes for decades or even, arguably, for centuries). The history of philosophy is replete with disputes of the latter, long-lasting kind; or, as we call them, 'persistent' disputes. To take a few illustrative examples, consider the nominalism-realism debate, the "free will" debate, or the mind-body problem.

The persistence of philosophical disputes has often been taken as symptomatic of over-ambitious and wrongheaded intellectual projects; the very ubiquity of such disputes has been used as an argument for the need for an extensive overhaul of the field.[1] For instance, Descartes dismissed both the philosophy and the science of his predecessors as dubious and ultimately ill-grounded, "seeing that it has been cultivated for many centuries by the best minds that have ever lived, and that nevertheless no single thing is to be found in it which is not subject of dispute." Ironically, Locke subsequently accused Cartesians of breeding "disputes [...] never coming to any clear Resolution [...] proper to only continue and increase their Doubts, and to confirm them at last in a perfect Skepticism" (Locke 1975, 31).[2] In his autobiography, Hume explained that he was struck very early by the fact that "Philosophy [...] contain[s] little more than endless Disputes, even in the most fundamental Articles."[3] Kant began his first critique with a gloom-ridden reflection on the fact that metaphysics is nothing but "a battle-field of endless controversies." In the 20th century, Wittgenstein and Schlick, among others, expressed a similar verdict. "Two thousand years of experience, argues Schlick, seem to teach that efforts to put an end to the chaos of systems and to change the fate of philosophy can no longer be taken seriously" (Schlick 1930, 53–54). Wittgenstein famously construed this chaos as a series of "endless misunderstandings."[4]

Yet these philosophers disagreed both on the exact diagnosis and on the best treatment of persistent disputes. While Descartes thought that philosophy needed a constructive reestablishment that would put an end to its persistent disputes by answering the questions that had given rise to them,

---

1 This theme is developed by Rescher (1985, chap. 1).

2 In a fragment on medicine, he also says that rationalists "lay a foundation for endless disputes" (Locke 1975, xxiv).

3 In the *Enquiry Concerning Human Understanding*, Hume distinguishes two kinds of pointless, persistent disputes that we will later review:

> It is true; if men attempt the discussion of questions, which lie entirely beyond the reach of human capacity, such as those concerning the origin of worlds, or the œconomy of the intellectual system or region of spirits, they may long beat the air in their fruitless contests, and never arrive at any determinate conclusion. But if the question regard any subject of common life and experience; nothing, one would think, could preserve the dispute so long undecided, but some ambiguous expressions, which keep the antagonists still at a distance, and hinder them from grappling with each other. (Hume 1758, sec. 8, pt.1)

4 "Ordinary language leads to endless misunderstandings" (Wittgenstein 1929). For all these references, we have drawn on Rescher's (1985, chap. 1) useful survey of philosophical diversity.

other philosophers thought that the revisions needed would turn out to be destructive rather than constructive, appearing to defuse persistent disputes yet without answering the questions that had given rise to them. While Rationalists such as Descartes took it that persistent philosophical disputes could be solved in principle, if only the proper rational steps were taken, Empiricists and Kantians believed that they could only be "dissolved." For these latter philosophers, the very fact that the enduring disputes had lasted for such a long time meant that they could not be solved at all (thus the endless characterization), and that it was simply pointless for the contending parties to continue to argue over the disputed matters. In what follows, we shall call 'pessimistic' the claim that persistent disputes are always pointless and 'optimistic' the claim that they are not always so. We shall come back to the question of why Rationalists, but not Empiricists, tend to be optimistic about persistent disputes.Despite its impressive philosophical pedigree and the admittedly strong intuition it embodies, the pessimistic stance on persistent disputes has seldom been adequately defended. Contemporary researchers do often appeal explicitly to pessimistic intuitions, usually in order to dissolve some perennial disputes (in metaphysics, think of Hirsch 2016, 138; 2009, 241; in epistemology, of Alston 2005, 21–23; in metaphilosophy, of van Inwagen 2017, 129–131; or Stoljar 2017).[5] However, they hardly try to justify or deepen that intuition. To our knowledge, there is no direct argument in the literature purporting to show that, in philosophy or elsewhere, persistent disputes must be pointless in virtue of their very persistence. Moreover, no one has explicitly pointed out what is supposed to be wrong with the fact that a dispute persists for a long time. This paper aims at filling this lacuna while delineating the optimistic defense of persistent disputes.

After having defined disputes in section 1 and their persistence in section 2, we survey in section 3 the different ways in which a dispute may be said to be pointless. We then put forward in section 4 what we take to be the strongest pessimistic challenge to the optimistic claim that persistent disputes may in fact have a point. The challenge relies, as we shall see, on the fact that when a dispute persists for a long time, the best explanation for its persistence seems to render it pointless. In section 5, we consider a real-world example of a persistent dispute that has turned out demonstrably to have a point: the

5 Some make it clear that persistent disputes may be "interesting" even though they are pointless (more on that soon). Stoljar (2017) suggests that if philosophical disputes were all persistent, philosophy would be pointless. He argues, however, that they are much less persistent than they seem.

doi: 10.48106/dial.v76.i1

"Functions Controversy." Drawing on this example, we argue that some persistent disputes do have a point, and that their point is meta-analytic, implicitly concerning the best way, among quite different candidates, of understanding the terms and objects at issue. We show that such meta-analytic disputes can be settled and yield genuine theoretical (as opposed to merely pragmatic) progress.

The topic of disagreement has recently come to the fore of the philosophical agenda, yielding a multiplicity of debates about faultless disagreements, peer disagreements, deep disagreements, philosophical disagreements, and the a priori, conceptual engineering and metalinguistic negotiations. The question of persistent disputes, as we shall see, cuts across a variety of debates. It is therefore difficult (if not impossible) to do full justice to the precise ways in which these varied approaches interact. In the penultimate section 7, however, we connect our optimistic defense of persistent disputes to some of these recent debates and argue that it can prove fruitful for our understanding of the importance of metalinguistic negotiations and related phenomena in science and philosophy.

1 Disputes

At first approximation, a dispute over a sentence q is a situation in which different parties

- seem to disagree about q: while Pro asserts q's truth, Con denies it,[6]
- argue against each other in order to find out which one is correct, and which one is incorrect.

Note that there are countless ways in which one might object to this first approximation, going on consequently to build in complex and precise detail by way of refinement and exactitude. For our present purposes, however, a brief characterization should suffice.

6 We assume throughout the paper the following equivalence schema: the proposition expressed by a given sentence use is equivalent to the proposition asserting the latter's truth. We also assume that assertion and denial are incompatible speech-acts (one cannot coherently assert and deny the same proposition) and exhaustive speech-acts (someone who has settled his mind about a proposition should be disposed to assert it or to deny it).

For simplicity, we suppose (against, e.g., dialetheists such as Priest 2006) that asserting that q is not true amounts to denying that q is true and thus to engaging in a dispute with someone who asserts that q is true.

Two parties disagree when one asserts and the other denies q('s truth).

Parties The parties involved in a dispute might be single individuals, or collectively, they might form groups. Moreover, the weight or preponderance of the argument on each side might well be asymmetric. Consider, for example, the dispute over whether the earth is flat, opposing (in the present day) a negligibly small number of flat-earthers to virtually everyone else. Or, to take a limit case, think about the disputes opposing some delusional patients to their doctors and families (see Hohwy and Rosenberg 2005; Bayne and Pacherie 2004 for a couple of relevant case reports).

Aims We assume in this paper that the *primary aim* of a dispute is to find out who is right or wrong, that is, whether Pro's assertion is true, and Con's is false or the opposite. Some might object that the aim of a dispute should be construed in terms of knowledge, or of some other norm of assertion, rather than truth (see e.g., Williamson 2000). This point is well taken. Because it will make things simpler, however, and because we believe it does not affect the main thrust of our arguments, we will neglect alternative, knowledge-based, views of the primary aim of disputes.

Be that as it may, a dispute might be quite useful even when it does not fulfil its primary aim. Pursuing it might, for example, allow the disputants to attain other valuable cognitive goals, such as finding out that it is impossible to reach the primary aim of the dispute, or that they need further evidence, or again that this dispute is connected in surprising ways to other classical disputes, and so on. In the case of collective or group-based disagreements, the dispute might allow a select few individuals to realize whether or not they are correct, even in the absence of a collective forming of opinion. In such cases, we might say that the dispute has fulfilled some of its *secondary aims*, and that it is accordingly *interesting*, even if it has no point. Finally, when neither its primary aim nor its secondary aims can be reached, a dispute might still serve what we might call *adventitious aims*: aims that are not directly related to epistemic values. Participating in a philosophical dispute to which one has skillfully and adeptly devoted time and effort, for example, might help one keep one's job as a philosopher and pay one's rent on time.

doi: 10.48106/dial.v76.i1

2 Persistence

What about *persistent* disputes, one might ask? ‘Persistent’ is a rather vague and context-sensitive word. In order to make explicit what we mean by it, we need to be clear as to the role we assign to the notion of persistent disputes. This notion is epistemologically useful and significant, we believe, because the persistence of a dispute casts a doubt on its having a point. For the doubt to arise, two things are necessary. First, a persistent dispute must have existed long enough to allow all participating parties to share their evidence, extensively argue, and thoroughly assess the arguments put forward. Although this might depend on the pace of exchanges and on the number of people involved, it might be surmised that several decades should suffice for the process to be completed. However, this condition is neither precise, nor sufficient, for constraining the analysis. To see why, suppose that new and important experimental results for and against *q* reliably emerge over a short period of time (say, every year), and that as a result, a couple of antagonistic scientists contend over *q* for decades. The very long time they have spent arguing would not be epistemically challenging, or not quite as much as it would have been, had the relevant empirical evidence remained constant all along. It would indeed be easily explained by the continuous discovery of new empirical data, contributing to each new iteration of their argument. Accordingly, if we do not want to deprive the category of persistent disputes of much of its epistemological usefulness and significance, we should say that a dispute over *q* is persistent only if, *while the relevant available empirical evidence did not significantly change, it has lasted long enough to allow all parties to share their evidence, extensively argue, and thoroughly assess the arguments put forward*. Conversely, we count as persistent a dispute satisfying this condition. The examples from the history of philosophy given above do not all qualify as persistent disputes in this sense, as for some of them (most notably, the “free will” debate), the relevant empirical evidence has in fact significantly changed over the centuries. But one thing the debates we adverted to should have in common is that they all involve a *series* of persistent disputes in our given sense. Thus, one might say that both the discovery of classical mechanics and the discovery of quantum mechanics ended a form of persistent dispute over free will and, at the same time, gave rise to a new variant. Similarly, when Thomas Young made his two-slit experiment, one arguably persistent dispute over the nature of light (wave or particle) ended, and another one took its place.

3 Varieties of Pointless Disputes

3.1 *What Is a Pointless Dispute?*

Throwing a rock at the sky is pointless if it is aimed at knocking the moon off orbit or at causing rainfall in the Sahara. It is not pointless if it is part of a game or play. More generally, an action has a point if and only if, given one's capacities and the laws of nature, it allows one to reach the aim we assign to it. A dispute is a kind of action too, albeit a collective action. And just like throwing a rock, it will have a point if and only if it permits the disputants to achieve the aim of the dispute. As we have seen, disputes can be assigned many aims. Previously, we distinguished the primary aim of a dispute (finding out who is right and who is wrong) from its secondary aims (such as finding out whether the primary aim can be attained) and adventitious aims (such as keeping one's job as a philosopher). No one would be tempted to say that a dispute has a point only because it allows one to reach some of its adventitious aims. The matter is less straightforward when it comes to secondary aims. There is, in any case, an interesting category of disputes that are pointless in the sense that, *given their epistemological profile, taking part in these disputes cannot allow the disputants to reach the primary goal of these disputes, that is, cannot allow them to find out who is right and who is wrong about q*. By the "epistemological profile" of a dispute, we mean not only the rationality of the parties, broadly understood (that is, their epistemic virtues and capacities, and the various epistemic vices, motivational influences, and cognitive biases that might hinder the exercise of the former), but also the way rationality itself (in terms of capacities, virtues, biases, and influences) evolves over time. We should also include in the epistemological profile of a dispute the distribution of the relevant available evidence and its relation to both parties (i.e., how easily accessible it is to both) and other relevant epistemological factors. In what follows, we focus on disputes that are pointless in this primary sense. Importantly, if a dispute is pointless (in that sense), the fact that the parties want to find out who is right and who is wrong gives them no practical reason to keep arguing against each other. If that is the only thing they are hoping to achieve, then the debate is indeed terminally devoid of point, and the disputants would be better off engaged in other pursuits.

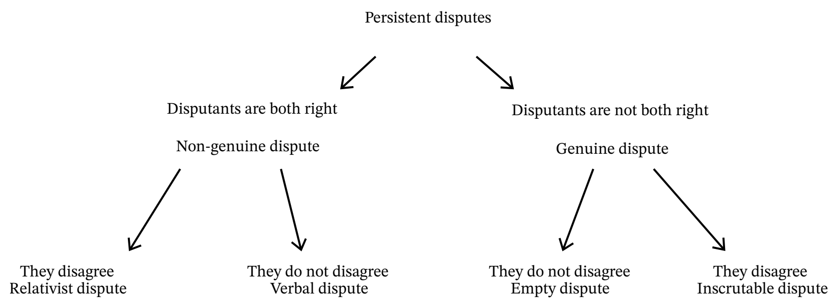


Figure 1: Four ways for a (persistent) dispute to be pointless.

3.2 *A Typology of Pointless Disputes*

It is possible to distinguish four types of pointless disputes. Notice that appearances notwithstanding, opposing parties engaged in a dispute might in fact both be right. In such a case, we should say that the dispute is *not genuine*. If a dispute is not genuine, then neither of the disputants is wrong; it is accordingly impossible to find out which of the two is wrong and *a fortiori* to settle the issue by arguing antagonistically. Non-genuine disputes are, therefore, manifestly pointless. There are, however, two different ways for a dispute to be non-genuine, as we shall now explain.

Verbal Disputes Typically, a non-genuine dispute is one in which both parties *do not genuinely disagree*. Yet, one might ask, how can two speakers fail to disagree if one asserts that q is true, while the other denies it? Such an eventuality might easily obtain if the speakers misunderstand each other, for example, if q contains ambiguous terms, and the disputants are linguistically at odds over the various intended contents. In such a case, it turns out that if there is disagreement at all, it is about how to use words and their possible meanings, and not matters of deeper substance. Thus, the apparent dispute is, contrary to first impressions, merely *verbal*.[7] For instance, it is sometimes claimed that in matters of taste, disputes are verbal because what 'tasty'

7 Some disputes might be verbal and substantive, rather than *merely* verbal, if q is itself about language. For simplicity, we will suppose that q is not about language and that verbal disputes are all merely verbal.

means is tantamount to 'tasty for the speaker who utters it', and will, on this analysis, mean different things as uttered by different speakers. The claim that metaphysical disputes are verbal corresponds to a form of metaphysical pluralism. Carnap (1950) seems to have held such a view about ontology.[8] Hirsch (2009) has recently revived that view, arguing that many (but not all) metaphysical disputes are verbal.

Relativist Disputes There are moreover some non-genuine disputes in which the disputing parties nevertheless genuinely do disagree. That is to say, there is no linguistic misunderstanding of the type above, and yet, intuitively at least, both parties really do put forward conflicting proposal. A similar conundrum arises: How can two people, who are said to genuinely disagree with each other, nevertheless both be correct? The answer is that such a predicament might occur if the truth of the disputed sentence is *relative* to certain parameters, be they moral standards or standards of taste, theoretical frameworks or paradigms, and similar. Goodman (1978) argues that even when they are genuinely conflicting and not ambiguous, a sentence *q* and its negation can be both correct because they are not correct in or relative to the same "world." Goodman calls his view radical relativism, and his relativism is indeed radical in the sense that it is universal. More recently, some philosophers have advocated circumscribed forms of relativism (see MacFarlane 2014). Some have argued that disputes about taste are not usually verbal because adversaries in matters of taste do not talk past each other; when I say that spinach is tasty and you deny it, our speech acts bear on the same proposition, and our disagreement is tangibly real. Such a disagreement, it has been claimed, might nevertheless be faultless (in the relevant sense that permits both of us to be right) if truth about matters of taste is made relative to latent standards of assessment. Goodman (1978) has held that metaphysical disputes are relativist and endorsed metaphysical relativism. As we understand him, Gallie (1956) argues that many disputes in the political and social domain are "endless," because they are relativist. Williams's (2011, chaps. 7–10) seems to hold a similar view (which he calls 'non-objectivism') regarding many moral disputes.

8 More precisely, Carnap's (1950) view was that there are two possible readings of ontological questions: on one "internal" reading, they are verbal; on another "external" reading, they are empty. It should be reminded, however, that Carnap granted a useful, pragmatic role to certain external ontological questions, namely that of helping us choose and coordinate on a given ontological framework (see Flocke 2020).

doi: 10.48106/dial.v76.i1

Empty Disputes Genuine disputes are disputes in which at least one party is *not* right about *q*. Yet these types of debates might be pointless too. Starkly, this will obtain when neither party happens to be correct about the matter at issue. In such a case, the primary aim of the dispute—finding out which one of the two parties is right and which is wrong—will, as before, be impossible to achieve. One might say by way of a stipulative definition that when both opponents are not right, their dispute is *empty*. Trivially, if the disputed sentence *q* is meaningless, the dispute over *q* is empty. In such a case, it is a moot point whether the parties do in fact disagree.[9] Expressivists about taste might thus argue that "spinach is tasty" or " 'spinach is tasty' is true" are merely expressions of feelings which are neither true nor false and that disputes about such matters are always empty. In metaphysics, the claim that disputes are pointless because they are empty has been maintained by the Logical Empiricists. It expresses a form of metaphysical anti-realism (Carnap 1931). Thomasson (2009) and Yablo (2009) argue that some metaphysical disputes might indeed be empty.

Empty disputes constitute a central case of the category of pointless genuine disputes. We now come to a third.

Inscrutable Disputes In order to have a point, a dispute must be genuine and non-empty. Let us call 'substantive' a dispute in which one party is right while the other party is not. Not all substantive disputes have a point. A substantive dispute will indeed be pointless if it is impossible for the parties to come to an agreement through rational exchange, that is, if the epistemic reasons justifying the assertion or the denial of *q* are inaccessible to one of the parties. Note that the impossibility and inaccessibility at stake in this context are epistemological. They depend on what we have called the epistemological profile of the dispute, and in particular on the rationality of the disputants. We call those disputes whose epistemological profile makes it impossible to convince by dint of reasons the error-committing party, inscrutable disputes. The claim that traditional metaphysical disputes are pointless because they are inscrutable expresses a form of metaphysical skepticism. This Humean or Kantian view has contemporary advocates. Kriegel (2013) puts forward an argument to the effect that they are always inscrutable. Bennett (2009) claims that some of them are.

9 We have assumed that the norm of assertion is truth and truth only, and we will suppose that a meaningless sentence cannot be true and should not accordingly be asserted.

Verbal, relativist, empty, and inscrutable disputes are subcategories of pointless debate. Conversely, if a dispute is neither verbal nor relativist, empty, or inscrutable, it is a substantive dispute in which, given their cognitive capacities, the disputants can, in principle, come to an agreement over q by means of argument and rational persuasion. It will accordingly be a dispute that has a point. Verbal, relativist, empty, and inscrutable disputes thus nicely partition the field of pointless disputes (see figure 1).

Note that there is an interesting contrast between verbal, relativist, and empty disputes, on the one hand, and inscrutable disputes, on the other. Whereas the first three types are pointless for a semantic or an ontological reason, the last type is pointless for an epistemic reason. Importantly, as we have emphasized, a dispute might be pointless but still interesting and, accordingly, worth having. Bennett (2009) claims that this has been the case for some ontological disputes that are inscrutable, and Sosa (2010, 281) argues that this is the case for many philosophical disputes that are verbal.

4 The Pessimistic Challenge to Persistent Disputes

Our general discussion of the futility of disputes is directly relevant to *persistent* disputes, which may turn out to be pointless in precisely four different ways, on the present analysis: they may be verbal, relativist, empty, or inscrutable. Our question is now the following: Is there something in the persistence of a dispute that makes it likely to fall into one of these categories? It may be assumed that *some* persistent disputes are pointless, but why should the very persistence of a dispute always make it pointless? Since we are envisaging an internal connection between persistence and pointlessness, we need to examine *general* arguments for the pointlessness of persistent disputes. We shall see, however, that these general arguments can also be applied on a case-by-case basis, yielding more cogent conclusions for some persistent disputes, as opposed to others.

There is an obvious inductive argument which infers endlessness from persistence: if a dispute has existed for a very long time without having been settled successfully, it will never be. When Empiricists or Kantians say that metaphysical disputes are *endless*, they seem to appeal implicitly to an argument of this kind. The weakness of the inductive argument is easy to see once the latter is made explicit. For instance, a similar argument would have

doi: 10.48106/dial.v76.i1

concluded twenty years ago that the perennial search for a demonstration of Fermat's last theorem was pointless, which we know is obviously false.[10]

More significantly, the inductive argument fails to distinguish between persistent *disputes* and persistent *questions*. When a dispute exists for a very long time, the intuitive worry is not so much that a complicated question fails of an answer (persistent questions are legion in mathematics and natural sciences, and few would claim that their persistence means that they are pointless). The intuitive worry is rather that, *despite their common knowledge of the unsettled issues, the parties do not relinquish their dispute and continue to hold and argue for (apparently) dissenting views*. In a genuine persistent dispute, one of the parties does not know that she is not right, and that she does not know the answer to the question. But this is not so, in general, with a persistent question (think again of the many conjectures and open problems in mathematics and physics which do not yield persistent disputes). Unlike persistent questions, persistent disputes involve a form of *reflective opacity*. Accordingly, they seem much more worrying from an epistemological point of view than mere persistent questions.[11]

This intuitive worry forms the basis of a serious philosophical challenge, a challenge that is abductive rather than inductive. The challenge is to explain the persistence of a given dispute without assuming that it is pointless. What might account for the fact that parties persist in disputing a sentence's truth if their dispute is not, in one way or another, pointless? Below, we will introduce two important and connected problems that the theorist we have

10 The inductive argument is probably stronger in the special case of *philosophical* persistent disputes, as one could contend that none of the philosophical disputes that have lasted a long time have been solved. This last claim, however, is less obvious than it seems. Stoljar (2017) has for example argued, rather convincingly, that if we individuate philosophical disputes properly (and distinguish, for example, the various questions that we loosely put under the heading mind-body problem but that have been raised at very different historical periods and are indeed quite different), the track record of philosophy is similar to that of other fields, and that once philosophical questions are properly individuated, philosophical disputes last much less than is usually assumed (many remain persistent in our sense). Moreover, there are other fields in which disputes tend to last. In a classical paper, Gallie (1956) argued that some concepts are "essentially contested," i.e., are bound to lead to endless disputes, in part because of their evaluative character. These include the central concepts of political science and legal theory. In any case, our question at this point is not limited to philosophy: we are wondering whether the persistence of a dispute per se *generally* gives us reason to deem it pointless.

11 There is an additional difference between persistent questions and persistent disputes that we shall not consider here. In a persistent dispute, the parties typically know that someone who is likely to be a peer disagrees with them. This knowledge gives rise to the problem of "peer disagreement" (see Feldman 2003; Elga 2007; Christensen 2007; as well as Kelly 2005, 2010).

characterized as *optimist* must face in order to answer this challenge. The first one is, roughly, that if a dispute *which has a point* persists, both parties should become competent enough to settle it after a reasonable time. This dispute should not, accordingly, be persistent. This is the *competence problem*. The second one, which we call the *problem of apt a priori disagreement*, can be stated thus: when a dispute persists and involves sufficiently rational subjects who can share the relevant empirical evidence, it reflects a persisting *a priori* disagreement among rational subjects whose judgments are both apt. But it is hard to see how such a thing could be possible. Taken together, these two problems suggest that the challenge cannot be met and that persistent disputes are pointless.

4.1 *The Competence Problem*

How can a dispute persist if it is not pointless? A successful explanation should first grant that the dispute is substantive: one party must be wrong and the other right; otherwise, the dispute would be merely verbal, relativist, or empty, and hence pointless. It should accordingly explain the persistence of the dispute in epistemological terms, invoking a bad epistemological profile of the dispute. The epistemological profile must not be too bad, however; that is to say, it must not be incorrigibly bad, for otherwise the dispute would be inscrutable and pointless. In other words, the parties should be competent enough to settle the dispute, but their performance should be impeded by some epistemological obstacles liable to be overcome, albeit extremely slowly.

Let us see how this might happen by singling out the epistemological obstacles that might plausibly explain persistent disagreements—call these persistent disagreement factors[12]—and see whether they can explain a persistent dispute. Persistent disagreement factors all hinge on an asymmetry in the distribution of certain epistemic features that need to be overcome.

Asymmetric Access to Empirical Evidence Rational agents do not, as a rule, have equal access to all available empirical evidence relevant to a given question. This fact explains many of our persistent disagreements. For instance, I believe that the male rather than the female of the seahorse species carries eggs because I recall coming across this information in a book on marine life roughly thirty years ago. My partner believes the opposite because

12 We borrow the term 'disagreement factor' from Frances (2014).

it seems to him less implausible as a scientific hypothesis. We have disagreed all this time (to be frank, we never much talked about it).

Similarly, I can disagree with my neighbor about the claim that vaccines are, on the whole, more dangerous than the disease against which they offer immunity, at least in part because I happen to have access to far more reliable scientific sources than he does and because my sources, but not his, inform my opinion correctly in view of the relevant facts. Consequently, the disagreement can rage on unabated for a considerable period.

Some theists likewise explain their disagreement with atheists, as well as with advocates of rival religions, by claiming that *they* have experienced the presence of (their version of) God (among philosophers, see, among others, the influential accounts of Plantinga 2000; Alston 1991).

Differences in Rationality Psychologists have shown that we are almost without exception affected by cognitive biases and that consequently, different thinkers display different cognitive "styles." They have also shown that our motivations can significantly affect our beliefs and their entrenchment. It is safe to suppose that cognitive and motivational biases can account for a range of persistent disagreements.

Take the following puzzle, a paradigm case for attracting disagreement. Suppose Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. Which is more probable? That Linda is a bank teller, or that Linda is a bank teller and is active in the feminist movement? Kahneman, Slovic and Tversky (1982) have argued convincingly that many people wrongly believe that the second option is the more probable because they use a misleading representativeness heuristic to assess probabilities.

Moreover, it has been observed that psychological factors can affect real-life persistent philosophical disagreements. Enoch (2011, 192–195) has argued that many moral disagreements are partly grounded on the distorting effects of self-interest. As an illustration, he mentions the view advocated by Peter Singer and Peter Unger to the effect that unless we give almost all our money to famine relief, we are nearly as morally condemnable as murderers. As he says, "refusing to see the (purported) truth of Singer's and Unger's claims thus has tremendous psychological payoffs" (Enoch 2011, 193).

Feltz and Cokely (2013) have likewise argued that some "persistent philosophical disagreements" are predicted by individual differences, in particular by personality traits, which determine certain cognitive biases. They show, for instance, that extroverts tend to endorse the compatibility of free will and determinism.

Different Epistemic Paths and Starting Points Finally, some persistent disagreements can be explained by factors that do not directly depend on differences in rationality or access to the relevant empirical evidence, but only on what we might call the *topography* of the disagreement. That is, on the different starting points, and on the different paths taken in the course of a disagreement. The idea is to compare the evolution of someone's opinion on a given topic to climbing a mountain. Even if two people are aiming at the same terminus (by analogy, truth), and even if they are in a perfect physical condition (by analogy, even if they are perfectly rational and have common access to the relevant empirical evidence), they might end up in different places simply because they had different starting points, took different paths thereafter, and because the landscape itself is full of pitfalls.[13]

Arguably, the most notable pitfalls are what philosophers call *vicious epistemic circles*. Often, such circles successfully entrap ordinary subjects, altering the form of their beliefs and creating the conditions for long-standing divergence of opinion. Thus, the prevalence of conspiracy theories in some social contexts has been explained in terms of the fact that some people do not trust the accredited experts because they do not trust the institutions bestowing credentials upon them. But they do not trust the institutions accrediting the latter because they believe in conspiracy theories. Conspiracy theorists are trapped in a vicious epistemic circle. Basham (2001, 270ff) has argued that, as a result, we are not, in general, in a position to find out whether a claim of conspiracy is correct. We cannot but assume a prior answer to the core question of how conspiracy-prone our society is, in order to derive a well-justified position on the issue (Basham 2001, 274). If Basham's view is correct, those who start by trusting institutions end up rejecting conspiracy theories, and those who distrust them are bound to adopt conspiracy theories. Vicious epistemic circles have also been invoked to explain the fixity of delusional beliefs (Hohwy and Rosenberg 2005; Bayne and Pacherie 2004) and the persistent disagreement between for-vaccine and anti-vaccine factions, and flat-earthers

13 We borrow the term 'starting point' from Feldman (2006).

doi: 10.48106/dial.v76.i1

and their opponents (see, e.g., Levy's 2019 account of scientific denialism; and Nguyen's 2020 account of echo chambers).

Note that these disagreement factors can explain persistent *disagreements*. Can they explain persistent *disputes*, and persistent disputes that have a point, and are hence substantive and "scrutable"? In a dispute (as opposed to in a mere disagreement) the parties argue to correct and convince each other.[14] In a scrutable dispute, moreover, the epistemological profile must be good enough to allow the parties ultimately to settle the dispute by means of rational argument. The parties must be sufficiently rational (the cognitive biases and the motivational influences on beliefs affecting them must be benign and corrigible), vicious circles must be eschewed, and the relevant available evidence must be equally accessible to both. In such a scrutable dispute, time will accordingly have a beneficial effect. It will progressively cancel not only performance errors, but also the impact of differences in rationality (due, for example, in cognitive biases and motivational influences) as well as the asymmetries in the access to the empirical evidence. For a scrutable dispute to persist, this beneficial effect of time must be real, but extremely slow. The gist of the competence argument is that in most cases, such a very slow effect is simply implausible: either the disputants are competent enough to settle the dispute, and it should be settled in a reasonable amount of time, or they are not competent, and the dispute is pointless.

Let us see how this works on the above examples. It is reasonable to suppose that the disagreement about seahorses' eggs and the disagreement about probabilities in the Linda example will not yield anything like a persistent dispute, or at least not one that is scrutable. If the dispute is scrutable, both parties have the capacity to acknowledge without further ado the decisive evidence to the effect that male seahorses carry eggs or that it is more likely that Linda is a bank teller rather than a bank teller and something else. It is hard to see what could prevent them, then, to quickly come to an agreement.

In the theistic example, on the other hand, it seems that the disagreement could indeed yield a persistent dispute, but it is dubious that the asymmetries in the access to the relevant empirical evidence can be redressed by means of simple debate. As James (1902, 371) emphasized, religious experiences are usually very difficult to communicate. They seem to provide what is sometimes called subjective, private, or first-person evidence (Schellenberg 2007, chap.

14 Connection with active sense of disagreement.

8). Accordingly, if the religious disagreement case yields a persistent dispute, this is likely to be merely of the pointless kind.

The moral disagreement case, the free will case, and the anti-vaccine case are less straightforward to analyze. Historically, disagreements of their type have given rise to genuine disputes, both at the factional (group) level and at the level of individual thinkers. There are reasons, however, to believe that such group-level disputes are pointless. Take the moral disagreement, for example, and suppose, for the sake of the argument, that Singer and Unger are right and that their opponents are simply self-deceived. For the dispute to have a point, it must be possible, through rational exchange, to correct the distorting influence of their self-interest on their beliefs and have them change their mind. But even if we could do that, it would not suffice to settle the debate, as there would always be new, self-deceived comers joining the ranks of Singer's and Unger's opponents who have not yet benefited from the virtues of rational redress. The ensuing dispute would arguably be pointless. A similar analysis might deal with the anti-vaccine and the free will cases. The problem in such cases is that new members of one group seem to be selected by their epistemic limitations (more precisely, by how they fare on some disagreement factor), which prevents the dispute from being settled.

The competence challenge is not a knockdown objection against persistent disputes that have a point. Nothing prevents, theoretically at least, the possibility that a dispute exists that is shaped by cognitive biases, asymmetries in the access to evidence, or differences in starting points and epistemic circles that can be overcome, albeit extremely slowly. The competence challenge can, however, yield a schema of abductive arguments that should be applied on a case-by-case basis, as we have illustrated above. For a given persistent dispute, depending on the precise details of the case, the strategy of appealing to the argument schema implies that the best explanation of why such a dispute persists makes it pointless. It is in fact arguable that many pessimistic views about the debates in metaphysics and elsewhere stem implicitly from the idea that in these cases of interest, disagreement factors are set at rest once by one, making persistence mysterious.

There is moreover a broad category of cases to which the competence challenge can be applied directly, as opposed to on a case-by-case basis, characteristic of our argument schema strategy. It is the category of disputes in which differences in rationality are sufficiently easy to overcome, the epistemic circles sufficiently easy to escape, the starting points sufficiently close, the relevant available evidence sufficiently easy to access or share, and the

doi: 10.48106/dial.v76.i1

epistemic profile of the dispute, more broadly, sufficiently auspicious. Call such disputes *virtuous disputes*. In such cases, the disagreement factors we discussed, which might explain the dispute without making it pointless, will most likely be cancelled after a short period of rational exchanges (perhaps in the space of a couple of years). Virtuous disputes, it should be emphasized, need not have a point. Virtuous disputes are such, however, that their epistemic profile seems incapable of explaining their persistence without making them pointless. But it is hard to see what else could explain their persistence; accordingly, the argument concludes, they will only persist because they are pointless, i.e., because they are verbal, relative or empty.

The point raised above is interesting, since many disputes seem at least prima facie virtuous, and some of these seem persistent too. Think of disputes among researchers on such topics as mereological composition in ontology, or fundamental axioms in mathematics (for example, disputes about the truth of the continuum hypothesis). Or consider, in biology, the disputes over the choice of a classification system based on phylogeny vs. interbreeding (LaPorte 2004, 70–76); or, in cognitive neuroscience, the dispute over the neural correlates of visual consciousness; or, in cosmology, disputes over the status of multiverses. Many people engage in these disputes with the hope of settling them in a reasonable time, and they seem to believe that these disputes are virtuous (the case of ontological debates is perhaps moot). There is no question that self-interest sometimes plays a role in them, some researchers being motivated, for example, by the perspective of promotions and social recognition. It is, however, at least prima facie plausible that such motivational influence and other aspects of the epistemic profile cannot explain the persistence of these disputes. At least this is what many researchers engaged in these disputes seem to believe.

In short, the competence challenge enjoins us to find an explanation why some disputes stubbornly persist, which does not entail pointlessness. In many cases, it is difficult to understand how the dispute may persist for protracted periods of time without being pointless, since, as we have outlined, if the dispute has a point, the participating parties must be sufficiently competent to settle it, and the passage of time must bring with it adequate and timely redress. This then is the *Competence Problem*. It might be possible to meet this challenge for some forms of persistent debates. It is difficult, however, to see how this might proceed, especially in the case of virtuous disputes.

4.2 *The Problem of Apt A Priori Disagreements*

The *Competence Problem* is related to a second, cognate difficulty, namely, the *Problem of Apt A Priori Disagreements*. Roughly sketched, this says that (i) when a virtuous dispute is persistent, it becomes a priori; (ii) however, given that the disputants involved in a virtuous dispute are equally competent to assess a priori claims, it is very hard to see what could explain the persistence of their dispute. We tackle these two premises in turn.

Since the disputants engaged in a persistent virtuous dispute are said to gain quick and easy access to a shared empirical body of *relevant* evidence, one might suppose that their disagreement would at some early point become independent of relevant empirical evidence. Since other pieces of empirical evidence are, ex hypothesis, not relevant to this dispute, the disagreement is also independent of these latter. Overall, the dispute thus becomes independent of *all* empirical evidence, relevant as well as irrelevant, and, accordingly, a priori.

If the virtuous dispute over the sentence q is not pointless, the persistent disagreement will in fact be grounded on a (more or less explicit) disagreement over a more fundamental sentence q^*, to the effect that the available empirical evidence provides decisive reasons for q. The sentence q^* will be a priori not only because the difference in attitudes toward it (namely, one party believes that q^* is true, the other one that it is false) is not grounded on a difference in empirical evidence, but also because, if the parties were to settle the dispute, their correct attitude toward q^* would not be similarly grounded either.

There are classical Platonic and Kantian arguments to the effect that fundamental disagreements in metaphysics and ethics hinge on a priori claims.[15] Our argument is much simpler and much more modest than these. First, our argument relies on a dialectical and quasi-operational conception of the a priori (expressed by the necessary condition that, to the effect that a disagreement that does not depend on problems of rationality or on empirical evidence, must be a priori) that remains neutral on the cognitive mechanisms implied.[16]

15 These arguments hinge roughly on the idea that fundamental claims in ethics and metaphysics are necessary, and that necessary claims are a priori. See Wedgwood (2019) for an updated defense of the Kantian argument concerning ethics.

16 Interestingly, this conception would classify as a priori a dispute that hinges on the weighing of different theoretical "super-empirical" virtues. We side with Hirsch (2009, 233, n.3) who takes such disputes to be straightforwardly a priori, and against Hawthorne (2009, 217) here.

This conception also sidesteps an influential objection raised by Williamson (2007) against the significance of the a priori / a posteriori distinction. On the one hand, his objection relies

doi: 10.48106/dial.v76.i1

Moreover, our argument only targets disputes (not just disagreements) with a point, and only those, moreover, that are both persistent and virtuous. To reiterate, for a dispute to have a point, the relevant empirical evidence must be equally accessible . If the dispute is, moreover, virtuous and persistent, this equally accessible evidence must quickly become equally accessed in actual fact. Hence the dispute must quickly become a priori, depending only on a priori claims.[17]

Let us illustrate this point with an example. For the last two decades, neuropsychologists have disagreed about the neural correlates of visual consciousness; all the while, the accessible relevant empirical evidence did not change significantly. Roughly, while some (call them Pro) believe that the neural correlate necessarily involves frontoparietal networks, others (call them Con) believe that an activation of primary visual areas in the occiput is sufficient for visual consciousness.[18] Strikingly, they all agree on the data collected by both camps and on their prima facie relevance to the debate. While some have characterized this debate (in this and ancillary areas) as merely verbal (see, for example, Bayne 2007, 100; Rosenthal 2002, 660; and even more specifically, Gottlieb 2018), it is arguable that nevertheless the dispute is substantive, granted that they disagree on the way the universally accepted common data should be weighed and interpreted, and that their disagreement is grounded on a priori claims about scientific methodology and scientific concepts. Pro scientists explicitly suggest, for example, that consciousness is a priori tied to reportability and that the only scientifically tractable concept of consciousness is that of "cognitive access"; while Con scientists argue that consciousness is not tied a priori to reportability but is still scientifically tractable (see, for example, Block's 2007b insightful analysis of this debate).[19]

---

heavily on what he considers the mechanisms of a priori knowledge should be, an issue on which we remain neutral. On the other hand, our dialectical conception and the pervasiveness of persistent virtuous dispute do suggest that our notion of the a priori is indeed quite natural and philosophically important.

17 Our thesis here should not be confused with the claim made by Chalmers (2011) to the effect that a sentence is such that any "dispute over it involving a competent disputant is verbal" iff it is, in a sense, analytic. Our claim, we shall see, allows for persistent virtuous disputes that have a point (and hence are not verbal) and that are a priori but arguably synthetic rather than analytic.

18 Advocates of the first "Pro" view include Dehaene and Naccache (2001); Sergent and Dehaene (2005); Naccache (2005); Kouider et al. (2007); Kouider, Sackur and Gardelle (2012). Advocates of the second, "Con" view include Zeki and Ffytche (1998); Lamme (2004); Block (2005, 2007b).

19 One might concede that a virtuous persistent dispute that has a point quickly becomes independent of the empirical evidence that is *directly* relevant to the dispute, and hinges on background disagreements concerning, say, methodological principles or wide-ranging philosophical or

It may already seem mysterious that thinkers disagree on an a priori truth, but when being rational, they are competent enough to find out that it is indeed true. It gets all the more mysterious when their disagreement persists despite lively rational exchanges, since we can safely assume that they correct each other's performance errors and that their disagreement does not stem from such errors—it is an apt a priori disagreement. The problem here is not so much that one of the parties persistently fails to assent to a truth (q^* or its negation) that is a priori even though he is rational enough to do so and does not commit performance errors. After all, many competent subjects have persistently failed to see that some complex mathematical claims, such as Fermat's theorem or Poincaré's conjecture, follow from the relevant axioms. We already know that some a priori *questions* can persist for decades or centuries. The problem is rather that one of the parties wrongly and persistently *dissents* on the matter of the disputed a priori truth, and that both parties accordingly disagree persistently. In the case of Fermat's theorem, Poincare's conjecture, and many other classical conjectures, the historical landscape is starkly different—at least if we attend to the categorical assertions published in peer-reviewed journals and backed by tentative proofs, as opposed to hypothetical assertions expressed in conversation and backed by intuitions. Mathematicians may dissent for a couple of years about whether a particular complex proof of a given conjecture is correct (the recent example of the six-hundred-page-long proof of the *abc* conjecture is a particularly eloquent example; cf. Castelvecchi 2020). When no convincing proof has been published, they may persistently fail to know the truth of the matter, and consequently suspend their (considered) judgments for a long time, but they

ethical conceptions (these background disagreements might be considered as coming from differences in what we have called the starting points of the disputants). One might question, however, whether the latter disagreements need to be a priori; one might argue that they often depend on empirical evidence as well, even if the empirical evidence here is only indirectly relevant to the initial dispute. In response, it should be recalled that if the dispute is indeed virtuous, and if the empirical evidence mentioned is indeed relevant to the dispute (even if only indirectly), both camps should come to scrutinize it and share it, and their disagreement should quickly become independent of it. We believe that this answer is valid. It is fair to acknowledge, however, that it puts strain on the real-world relevance of the notion of a virtuous dispute. Someone skeptical of the claim that persistent disputes quickly become a priori can indeed deny that there are many genuinely virtuous disputes. This is probably what someone who believes that persistent disputes in philosophy are "just hard" to settle, but not a priori (maybe Williamson and Hawthorne?) should do. If she does not want to be accused of mere hand-waiving, she should, however, answer our pessimistic by a detailed analysis of the dynamics of "just hard" disputes showing exactly what kind of cognitive difficulties can make them persist.

doi: 10.48106/dial.v76.i1

do not generally disagree persistently about it.[20] The problem of non-pointless but persistent virtuous disputes is that being reflectively opaque, they seem to imply the existence of a kind of *deceptive* a priori truths; truths, that is, such that rational enough subjects not only fail to know them, but also wrongly believe them to be false (not knowing that they do not know them). We take it that deceptive a priori truths typically represent a kind of a priori truth whose existence will be granted by Rationalists, but denied by Empiricists, and that the challenge of apt a priori disagreements thus goes some way towards explaining why Rationalists, but not Empiricists, tend to be optimists about persistent disputes.

Logical Empiricists notably argued that all a priori truths are analytic and that rational subjects should assent to analytic truths merely in virtue of understanding them (at least if they do not make any performance errors). Assuming that two parties are sufficiently rational, and therefore capable of grasping a prior truth, there should be no room for disagreement about which a priori truths are true. Conversely, if rational subjects disagree about an a priori sentence, it follows that either they understand the disputed sentence differently and the dispute is verbal, or they do not really understand it and it is empty.[21] Logical Empiricists must thus reject the existence of the *deceptive a priori* and deny that persistent virtuous disputes can have a point.

One preliminary conclusion to draw from our discussion is that a theorist who believes that some persistent virtuous disputes have a point is committed to maintaining either that some a priori claims are *synthetic* rather than analytic, or else that some analytic claims are such that understanding them does not suffice to assent to them.

The first option makes ineliminable use of the notion of the synthetic a priori. Plausibly, it entails that persistent virtuous disputes are grounded in a difference in the rational or a priori evidence accessed by both parties. Such a difference would be an additional disagreement factor, one that we have not considered so far but that has the potential, in principle, to explain persistent

20 We should emphasise that our claim here only bears on classical conjectures such as Fermat's theorem, Poincare's conjecture, Goldbach's conjecture and others. We shall see, with the Functions controversy, that there *are* in fact persistent disputes in mathematics, most notably disputes that, unlike these classical conjectures, concern the best way to understand certain mathematical objects, and so the choice of definitions and axioms (what we call 'meta-analytic disputes'). We thank an anonymous referee for pressing us on that point.

21 Unsurprisingly, Hirsch's (2009) argument for the neo-Carnapian view that certain metaphysical disputes are verbal hinges on the fact that parties involved in these disputes regard their claims as "a priori and necessary."

virtuous disputes. The second option has an air of oddity about it. It implies that one could, after decades of reflection, completely change his mind about an analytic claim he understood very well all along.[22] We believe that neither option is completely implausible (we are, in fact, quite sympathetic to the synthetic a priori option). Yet, unless they are fleshed out in more detail, it seems that both strategies can only rename the problem of persistent disputes but not resolve it.

We are now able to sum up the pessimistic challenge to persistent disputes.

First, if persistent disputes have a point, they must involve disputants that are competent enough to settle the dispute. Yet it is difficult to see how such disputes may persist for an inordinately long time since, if they have a point, obstacles hindering the disputants' performances will be gradually overcome. Indeed, it seems that the longer a dispute lasts, the less reasons there are to persist.

Second, since parties in a persistent virtuous dispute swiftly gain access to the same relevant empirical evidence, their disagreement becomes apt and a priori in due course. This means that persistent virtuous disputes involve deceptive a priori truths: a priori truths that sufficiently rational thinkers, who do not err because of performance errors, reject and unknowingly fail to know.

We believe that even perfectly virtuous disputes can persist and have a point; hence, that the pessimistic challenge may be answered—and, indeed, in a rather mundane way. In order to answer this challenge, one need not appeal to any dubious form of rational intuition nor posit cognitive biases, epistemic circles, or asymmetries in the access to the empirical evidence that can only be overcome at an extremely slow pace. One need just acknowledge the existence of a common type of dispute, that we call meta-analytic and that, for reasons we will soon explain, can be extremely long to settle. Our argumentative strategy will rely on a real-world example: a well-known persistent mathematical dispute, which uncontroversially proved to have a point.

22 Williamson (2007, chap. 4) has argued that any purportedly analytic sentence is such that two subjects who understand it could disagree about its truth. His argument does not make it clear, however, that the two subjects could persistently disagree, or even be rational enough to settle their dispute and disagree (see especially Williamson 2007, 91–92). We shall see, in any case, that our proposed solution to the pessimistic challenge makes room for persistent disputes (those that have a point) concerning analytic sentences. Thereby, it does not threaten what we see as an important connection between analytic sentences and assent to such sentences by subjects who understand them. See section 5.4 and especially footnote 27.

doi: 10.48106/dial.v76.i1

5 An Example of Scientific Persistent Dispute: The Functions Controversy

Persistent disputes are not specific to philosophy and may occur, as we have seen, within science as well. Showing that a given scientific controversy that seems persistent really is persistent is far from trivial, however, as it requires showing that it is not covertly fuelled by new empirical discoveries (recalling that we have individuated persistent disputes by the relevant empirical evidence available).

The simplest way to circumvent this problem and to find an uncontroversial example of a persistent scientific dispute is to opt for an illustration coming from a purely formal science, such as pure mathematics. Arguably, in this domain, empirical evidence is irrelevant or at least non-decisive and cannot end a persistent dispute.[23]

Among disputes that have proved persistent, it is also difficult to find one that has uncontroversially proved to have a point. Often enough, *prima facie* persistent disputes do not appear to be clearly or definitively settled. Equally, we believe that the domain of pure mathematics is interesting in virtue of its (approximately) cumulative character (pace Lakatos 1976). In mathematics, the fact that a dispute has been deemed settled for a very long time seems to be a very strong reason to believe that it is indeed settled.

We understand that using a mathematical example might bring with it some additional complications. The semantics and ontology of mathematics are often deemed less straightforward than those of (say) geology or biology. We believe that these complications are rather light and largely outweighed by the advantages of mathematics mentioned in the preceding paragraph.

The example we have chosen from pure mathematics is the Functions Controversy. This scientific dispute has the advantage of having been, without a doubt, both persistent and uncontroversially proved to have had a point.

Between the beginning of the 18th century and the end of the 19th century, many controversies arose around different mathematical "results" concerning functions. Some of the controversial results were rather technical, but they included the following two simple claims:[24]

23 Note that this means, moreover, that it is not necessary to show that such a dispute is virtuous in order to show that it is a priori if it persists.

24 The mathematical layman can construe functions as graphs, discontinuities of a function as gaps in its graph, and the points at which it is non-differentiable as those where its graph does not admit a tangent.

(1) Every function is continuous, except possibly at a finite number of points.

(2) Every continuous function is differentiable except possibly at a finite number of points (see Hawkins 1975, 43–44).

Those claims were disputed because mathematicians were seemingly “discovering” some “objects” whose existence appeared inconsistent with their truth. In 1826, Abel, for example, showed that a certain function defined as a convergent series of continuous functions is discontinuous in an infinite number of points, apparently falsifying (1).[25] In 1829, Dirichlet discovered the eponymous “monster” function, which seemed like a function continuous nowhere and thus to falsify (1).[26] In 1872, finally, Weirstrass introduced his own monster, which seemed to be a function that is continuous everywhere but nowhere-differentiable and to falsify (2).[27]

These disputes involved earnest and rational thinkers; indeed, some of the greatest mathematicians of the epoch ranged themselves on both sides of the debate. Yet, the disputes concerning (1–2) were persistent and were not clearly settled until the second decade of the 20th century and the acceptance of modern set theory. The question arises as to how (1–2) could be maintained by many thinkers of quality despite the above counterexamples. It would appear that some proponents of the controversial statements denied that the alleged counterexamples were significant exceptions to the general rule. Others denied that they were genuine *functions* or even that they existed at all.

5.1 *Was the Functions Controversy Verbal?*

It is tempting to make the charge that the Functions disputes were at bottom merely verbal. Indeed, not all disputants understood the term ‘function’ in the same way. Neither did they all define it with an equal degree of rigor and precision. Reviewing the best textbooks in analysis, Hankel noticed in 1870 that among them,

25 That function was $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}\sin(nx)}{n}$, which is discontinuous for every value $(2m+1)\pi$ of x where m is an integer.

26 The Dirichlet monster is $\chi_{\mathbb{Q}} \begin{cases} 1 & if \quad x \in \mathbb{Q} \\ 0 & otherwise \end{cases}$

27 The Weirstrass monster is $f(x) = \sum_{n=1}^{\infty} b^n \cos(a^n \pi x)$ (with a an odd integer, b a real number in $[0,1]$ and $ab > 1 + 3\pi/2$).

doi: 10.48106/dial.v76.i1

> one [text] defines function in the Eulerian manner; the other that *y* should change with *x* according to a rule, without explaining this mysterious concept; the third defines them as Dirichlet; the fourth does not define them at all; but everyone draws from them conclusions that are not contained therein. (Kleiner 1989, 293)

There are, however, decisive reasons to think that even if the mathematicians' understanding of functions and their standards of rigor differed quite significantly, this was not the cause of their disputes. If their dispute had been merely verbal, (i) it would have been defused by the introduction of new undisputed names to refer to different kinds of functions, and (ii) its solution could only have brought about a terminological advance, as opposed to a substantial, genuinely mathematical progress. Neither of these was the case in the event.[28] By 1870, it was already clear to many that one could distinguish between the "algebraic" functions, which are defined by an "analytic expression" (i.e., algebraic formula), the "geometric" functions (i.e., whose curve can be drawn freehand), and the "logical" functions (i.e., arbitrary correspondences between two sets of values). Indeed, those who introduced this revisionary and more encompassing logical definition of function still wondered whether all "logical functions" really existed, and if they did, whether they really were functions. Thus Lakatos (1976, 151) points out that according to Dirichlet himself, the "monster" he had discovered was "an example not of an 'ordinary' real function, but of a function which does not really deserve the name." As late as 1904, Poincaré distinguished between logical functions and analytic functions (locally expandable in power series) and suggested that the former were not legitimate in mathematics (see Poincaré 1952, 125).[29] Even more strikingly, in 1905 Lebesgue, whose works permitted the generalization of the theory of integration to some "monstrous" logical functions, still argued that "*true*" functions are analytically representable (i.e., representable by an algebraic formula) (Lebesgue 1905, 139). Hermite essentially shared this sentiment concerning "this lamentable evil of functions without derivatives" (for Hermite's view, see Kleiner 1989, 294).

Moreover, the lack of rigor and precision found in many of the mathematicians' definitions did not result from inattention or neglect. Hence, the disagreement could not have been solved simply by substituting more precise

28 The introduction of new names to settle a verbal dispute is what Chalmers (2011) calls the 'method of elimination'.

29 Yet Poincaré seemed more open to mere "logical" functions in (1899).

definitions for the imprecise ones. Many mathematicians at that time explicitly rejected our modern standards of rigor. It was common, for instance, to regard theorems as rules and mathematical predicates as not in need of a precise formal definition (see especially Richards 2006, 700–713; and Lakatos 1976, 24). This seems also to have been the conception of Euler himself (Youschkevitch 1976, 67). Rigor and precision could only develop, it was thought, at the cost of fruitfulness. As Maloney (2008, 8) puts it, Lebesgue, for one, "[saw] the more precise and general definition of function, which we essentially use today, as a frivolity at best and a liability at worst."

Ultimately, the solution to these disputes did not stem from terminological advance, but from a substantial mathematical progress. Modern set theory and distribution theory were developed in response to such controversies. These controversies were laid to rest eventually, but not before the emerging new theories had shown their credentials and become entrenched in mathematical practice.

5.2 *Was the Functions Controversy Empty, Relativist or Inscrutable?*

As we explained, the Functions Controversy was not verbal. It did not hinge on the fact that some mathematicians, but not others, used a definition of functions, or true functions, that excluded the "monsters." Rather, it rested on the fact that participants in these debates disagreed on which definition was the best and ought to have been used. At this point, it might be suggested that the dispute was perhaps empty or relativist. There is, however, a straightforward argument to the effect that the dispute was neither empty nor relativist. If it had been empty or relativist, it could not have been settled, and we could not be said to know that (1–2) are in fact false. The same argument, it should be noted, *ipso facto* shows that the dispute was not inscrutable either.

Before moving forward, it is worth pausing on the decisive claim that the Functions controversy has been settled and has, accordingly, a point. We believe that in the present state of mathematics, this claim is uncontroversial. We also believe that it is (almost) uncontroversial that settling this dispute *that way* constituted a mathematical progress (denying this would require developing a revisionary / reactionary view of function that has no serious advocate today). What is less clear, and will be important later, is the normative status of this resolution, this progress, and the point of the dispute. A radical conventionalist might argue that the Functions controversy was solved by the mere acceptance of a stipulation (to the effect that functions

doi: 10.48106/dial.v76.i1

are logical functions) rather than by the discovery of a fact. He will probably concede that this resolution constituted a progress, but only because this stipulation was useful for us (and more useful than other conflicting ones) and insist that we only have practical reasons to consider (1) and (2) as true, not theoretical ones, and that the point of the Functions Controversy was somehow insubstantial or superficial. On the opposite side, Platonists, Kantians, Intuitionists, and even, arguably, Poincaré-style conventionalists will consider that mathematical truths do not depend on mere stipulations but on the structure of the world or of our minds, that mathematical progress is genuinely theoretical and substantial rather than merely pragmatic, and that the point of the Functions Controversy was thus deep or substantial. Let us call the first view of mathematical progress deflationist. We do not need to take a stand on this deflationism vs. non-deflationism debate here. What is important, however, is that non-deflationism is very plausible and clearly the majority view, that many philosophers, attracted by the claim that progress in philosophy is impossible, scarce, or at best pragmatic—and that the point of persistent disputes in philosophy is at best superficial—would be tempted to grant that mathematical progress is common and usually deep and theoretical.

5.3 *The Point of the Functions Controversy*

If the Functions Controversy was neither verbal nor empty, and by the same token, neither relativist nor inscrutable, it follows that it must have had a point. What, then, was its point? One thing that our discussion suggests already is that this controversy did not concern the properties of something (namely, functions), of which the participants had a *common subjective understanding*. Neither did it concern the best way to articulate such a common understanding. There was no such common understanding. Rather, disputants understood functions quite differently, and they accordingly defined them quite differently and accepted conflicting clusters of analytic claims about them. And their dispute was (implicitly) about the best among their rival understandings. Some mathematicians thought that the best understanding was the algebraic or geometric one, and they assessed under its light all claims about functions. Others favored the logical construal, and these latter ended up on the right side of the debate, correctly denying (1–2). Granting that one's understanding of something is reflected in the analytic claims one is disposed to accept concerning that thing, we might say that the point of the

Functions Controversy was not analytic but rather *meta-analytic*. The fact that the Functions Controversy was not verbal shows that a dispute whose parties appeal to very different understandings of the object at issue need not be verbal, provided that it is meta-analytic.

This is not a trivial conclusion. It might even seem problematic. On the standard, neo-Fregean views of concepts (viz., ways of understanding something that determine the reference to that thing in context), different understandings imply different concepts, and if the parties disagree because they use (or preferentially use) different concepts, it seems that their dispute must be verbal after all. Fortunately, recent work in philosophy of language and metaphilosophy focused on related phenomena provides interesting ways out of this problem.

The first line of research in philosophy of language puts forward "relationist" or neo-Gricean semantics that canvass the possibility of successful communication between two subjects that do not share the same concepts.[30] More germane still, the second line of research in metaphilosophy explicitly argues that what we call meta-analytic disputes are not verbal. Some philosophers working in the rapidly developing fields of metalinguistic negotiations, conceptual ethics, and conceptual engineering understand meta-analytic disputes as meta-conceptual but argue that the concepts involved, even if different, still share a common feature that prevents the dispute from lapsing into the verbal. For instance, they are said to be about the same "topic" (Cappelen 2018, 102–103), or are said to play the same role (Thomasson 2020). Others claim that meta-analytic disputes need not be verbal because the disputants share a similar meta-analytic aim. For instance, Belleri (forthcoming) writes of a "semantically progressive inquiry" and asserts that the unity of inquiry is at the bottom teleological. Yet others invoke externalist views of concepts to argue that even though disputants understand the object at issue in inconsistent ways, they might still share the same concepts (Schroeter and Schroeter 2014, 2016). Notably, Ball (2020) has argued that one should construe what we have called meta-analytic disputes as *metasemantic disputes*, that is, as disputes about the way one should "fix the meaning of words as we have used them before." In this article, we remain neutral on the best view of concepts

30 See, e.g., Récanati (2012, chap. 8) on the first-person, "relationnist" semantics according to which successful communication requires mere "coordination" or "de jure coreference" (Fine 2007; Taschek 1995; Pryor 2016), and neo-Gricean views that can likewise grant a form of mutual understanding without concept sharing (Buchanan 2014).

and meta-analytic disputes.[31] We observe, however, that there are many ways to do justice to the non-verbal character of such disputes.

5.4 *The Functions Controversy and the Pessimistic Challenge*

We say that a dispute is *meta-analytic* when it bears on the choice of the best way, among quite different candidates, to understand something, rather than on the attribution of properties to something the disputants understand in the same way, or on the best way to articulate their shared understanding of it.

Interestingly, the meta-analytic reading of the Functions Controversy allows us to provide a simple answer to the pessimistic challenge.

Take the competence problem first. According to the proposed interpretation of the dispute, what prevented disputants from agreeing was that they did not all understand (and hence define) functions in the same way. More deeply, they disagreed about which understanding was the best. But how, one might ask, could they disagree about that if they were all competent enough to find out which understanding is the best, and time cancelled the "usual suspects" for performance errors?

The comparative quality of competing understandings in pure mathematics and elsewhere depends, importantly, on their consistency and relative fruitfulness. It depends, more broadly, on their inferential profiles, that is, on all the inferences one can draw by their means. For finite minds like ours, however, evaluating such an inferential profile is not instantaneous. Each inference takes a very small amount of time to assess, but the number of inferences that need to be assessed is virtually infinite. Assessing the inferential profile is thus an *open-ended process*, that is, a process to which we cannot assign an a priori upper bound in time, be it in terms of years, or even centuries. Moreover, this process may prove surprising, as apparently consistent understandings may sometimes prove inconsistent (think of the naive understanding of sets, for a classical example), and apparently useless re-construals may sometimes prove fruitful. This means that assessing the relative merits of different ways to understand an object will not only be an open-ended process, but also a *non-monotonic* one: a process that may lead from a time when we have most

31 We would like to thank an anonymous referee of this journal for pressing us on the multiple possible interpretations of meta-analytic disputes.

reason to favour one understanding U_1 over the other one, U_2, to a time when we have most reason to favour U_2 over U_1.

For example, Poincaré, Lebesgue and Borel did not know, and they arguably could not have known without years of inquiries and intricate discussions with peers, that the logical understanding of a function would find its place in an important and consistent mathematical theory (set theory), that classical analysis would easily accommodate it, and that it would prove extremely fruitful in many fields (the popular Fractal theory is precisely a theory of "monstrous," supposedly merely logical, functions) and help provide many mathematical insights.[32] It is in fact arguable that they had good reason, at the very beginning of the 20th century, to dismiss merely logical functions as useless curiosities.[33]

The open-ended character of the process of assessing competing understandings successfully explains why it took mathematicians so long to answer the *questions* surrounding (1–2), and consequently to find out which understanding of 'function' was the best. Conjoined with the non-monotonic character of such a process, it furthermore explains how such persistent questions gave rise to persistent *disputes*. Each time a new aspect of the inferential profile was discovered, its assessment necessarily took some time, allowing for the emergence of dissenting views on the questions under scrutiny. In general, as time passes, new results are made public, cognitive biases and performance errors are removed through fruitful dialogues and debates, and experts become able to fully grasp them. But by the time this process reaches completion, new aspects of the competing inferential profiles may have been discovered, whose assessment may once again give rise to dissenting views through additional performance errors, cognitive biases, or simply ordinary delays and difficulties in communication. If assessing the comparative merits of two understandings were a monotonic process, it could be argued that dis-

32 Commenting on the set theoretic paradoxes Poincaré reportedly prophesied: "later generations will regard Mengenlehre (set theory) as a disease from which one has recovered" (but see Gray 1991).

33 It should be noted that in pure mathematics, the comparative quality of two ways of understanding and defining an object is an a priori matter. Arguably, an understanding of an object is better than another if it is mathematically more fruitful and does not lead to contradictions; that is, roughly, if it can yield better mathematical insights. It is true that the claims that most mathematicians prefer a certain understanding, or that they find it more fruitful, are a posteriori, but that is merely a posteriori evidence of an a priori truth (just like the fact that most mathematicians believe last Fermat's theorem has been proved is a posteriori evidence of the a priori truth that its purported proof is valid).

doi: 10.48106/dial.v76.i1

putants should have agreed sooner or later, owing to the gradual cancellation of communicative difficulties, biases, and performance errors. Arguably, they should have inferred, by monotonicity, that the dispute was settled once and for all. Nevertheless, as we explained above, the comparative assessment of two understandings is far from monotonic.

We pointed out at the outset that there is nothing mysterious in a dispute that lasts for a very long time if new relevant empirical evidence arises through continuous discovery. We are now able to make this thought more precise. There is no mystery because the process of assessing a growing body of empirical evidence is open-ended, if the body of evidence grows, and non-monotonic. The Functions Controversy persisted because it is a special kind of rational, non-empirical evidence whose assessment is both open-ended and non-monotonic, similar, in that respect, to the assessment of a growing body of empirical evidence, and unlike the assessment of trivial analytic evidence. The relevant a priori evidence was, in a sense, access*ible* all along to all parties, granted sufficient rationality. Being, however, open-ended and non-monotonic, its assessment took a very long time.

Pessimists grant—or should grant—that new empirical evidence may fuel ongoing debates in such a way that thinkers continue to disagree over the same issue for decades or even centuries. We suggest that their outright rejection of persistent disputes, in which by our definition the empirical evidence is fixed, reveals an unjustified refusal to acknowledge the existence of a type of evidence that is akin to empirical evidence in that its assessment is open-ended and non-monotonic, but that is, like trivial analytic evidence, a priori. This evidence concerns in particular the assessment of different understandings, which is open-ended, non-monotonic, and sometimes a priori. It is meta-analytic.

The meta-analytic reading of the Functions Controversy thus answers the competence problem. It also explains why disputants could disagree on an a priori claim. Thus, it can solve the problem of apt a priori disagreements. Even though the participants in the dispute preferentially resorted to different understandings of the concept of a function, we have seen that they were not talking past each other, and that their disputes were not verbal because they were meta-analytic. The fact that in a meta-analytic dispute two parties can, *without misunderstanding*, understand a disputed sentence in a radically different manner should already dispel the suspicion, associated with certain views of the a priori and the analytic, that any apt a priori disagreement must be verbal. The fact that a priori meta-analytic disputes can be solved shows that apt a priori disagreement need not imply that the disputes are either empty or relativist. More broadly, the meta-analytic reading of the Functions Controversy implies that there are some a priori claims that can only be known and understood by rational subjects' appeal to the best kind of understanding of the subject matter. So, for example, the statement to the effect that "monster functions are genuine functions" can only be known to be true by a subject who understands functions in the right way. While a subject interpreting it in the correct manner will endorse it, one who interprets it in another way is likely to deny it, even though she understands it, hence to fail to know that she does not know that monster functions are true functions.

This explains why, despite lively exchanges, some rational subjects might fail to assent to a given a priori truth or even might dissent from it, unknowingly failing to know that it is true. We have called *deceptive a priori truths* truths on which rational subjects can aptly disagree, and that they can, accordingly, wrongly believe to be false, not knowing that they do not know them. On the meta-analytic reading, the existence of *deceptive* a priori truths is not mysterious. It does not require us to posit unusual or non-standard analytic truths or a *puzzling form of* synthetic a priori. Rather, it stems from the fact that different subjects associate different understandings, and so different analytic truths, with a given term, even though they both understand the term and therefore don't misunderstand each other or talk past each other.[34]

6 The Point of Persistent Disputes

The Functions Controversy allows us to draw the following conclusions: First, the fact that a dispute is persistent, or even persistent and virtuous, does not entail that it is pointless. Second, a good explanation as to why some disputes persist is that they are meta-analytic and that meta-analytic evaluations, being

34 Interestingly, in the course of his argument to the effect that two subjects who understand a purportedly analytic sentence can nevertheless disagree over it, Williamson considers the hypothesis that subjects might disagree because they associate different concepts to the same words (and different thoughts to the same sentences) but rejects it on the ground that it would undermine Frege's requirement of the publicity of senses and that it would render the dispute verbal (Williamson 2007, 114–115). We believe that Frege's requirement is already challenged on other grounds—something Frege (1956, 298) himself seems to acknowledge and that can be accommodated rather well (Récanati 2012, chap. 8)—and that associating different concepts with a word in a disputed claim does not make the dispute verbal if, like in the case of the Functions controversy, the dispute is (implicitly) about the best way to understand the word and its denotation. Cf. footnote 25.

doi: 10.48106/dial.v76.i1

open-ended and non-monotonic, can take decades or even centuries. In order to find the best understanding of a term, one might need to assess the full inferential profile of the latter, which requires much time and can always prove surprising. Finally, and given the plausibility of the non-deflationary view of mathematical progress, the point of persistent meta-analytic disputes can arguably be deeper and more substantial than merely pragmatic.

The mere fact that a dispute is meta-analytic, as the example of the Functions Controversy shows, does not entail that it is pointless. The same could be said about the fact that the dispute is a priori. Even if virtuous persistent disputes become a priori, that does not make them pointless, because some evaluative claims about the comparative quality of different understandings are a priori, yet can yield persistent disputes that have a point.[35]

When the Pessimist proposed an abductive argument to the effect that all persistent disputes are pointless, she may well have been right to suppose that the disagreement factors in the epistemic profile of a dispute (to recall, asymmetries in rationality, in the access to the empirical evidence, and vicious epistemic circles) cannot explain its persistence. The Pessimist was wrong, however, to draw the conclusion that the best explanation of the persistence of a dispute is always that it is pointless. In some cases, the best explanation is that the dispute is meta-analytic and that meta-analytic disputes can involve the open-ended and non-monotonic assessment of priori evidence. In such cases, a persistent dispute need not be pointless. The competence challenge is only challenging for someone who neglects, among the disagreement factors, the difficulty of meta-analytic evaluations.

For all we know, there might be persistent disputes that are not pointless, even though they are not meta-analytic. Yet we would like to suggest that our diagnosis is quite general and that many persistent disputes in philosophy, in the sciences, and in public life (i) are meta-analytic and a priori, (ii) persist precisely for this reason, and (iii), crucially, are not necessarily pointless.

7 Meta-Analytic Disputes, Metalinguistic Negotiations, and Deep Disagreements

The view that many persistent disputes are meta-analytic disputes (as we have called them) is not entirely new. Arguably, it has been held under various

35 It is, in fact, tempting to dispel the apparent mysteries of the notion of synthetic a priori by claiming that synthetic a priori claims are simply meta-analytic claims.

guises by many philosophers, in relation to certain scientific and philosophical persistent disputes. Carnap's argument against traditional ontology, for example, relied on the thesis that disputes over meta-analytic questions (which he dubbed 'external questions') are empty, or perhaps relativist (see footnote 5). The view that persistent disputes are meta-analytic may well be at the root of Gallie's (1956) influential take on "essentially contested concepts." It may also be said to inform Williams's (2011, chaps. 7–10) analysis of ethical disputes and, arguably, Kuhn's (2012) understanding of (the disputes surrounding) scientific revolutions.

More recently, Sider (2009) has construed metaphysical disputes as disputes over the best understanding of quantifiers (and the best quantifier concepts). Many works in the field of metalinguistic negotiations, conceptual ethics, and the conceptual engineering literature have argued in a similar vein that philosophical disputes are often metaconceptual (and hence meta-analytic) disputes (Plunkett 2015; Burgess and Plunkett 2013; Cappelen 2018).

Likewise, Fogelin (1985) noticed that many disputes are "deep" in the sense that they stem from "a clash in underlying principles," can accordingly persist even though "the parties [are] unbiased, free of prejudice, consistent, coherent, precise and rigorous" and "by their nature, are not subject to rational resolution." Godden and Brenner (2010) and Shields (2021) have all argued that deep disagreements are in fact meta-conceptual.

Our view that some meta-analytic disputes are both substantive and scrutable and can persist without being pointless is much less widespread, however. Indeed, all these authors, except the most recent (e.g., Sider, Plunkett, Sundell, Burgess, Capellen, Shields), seem to believe that meta-analytic or "metaconceptual" questions are pointless. To our knowledge, even the latter do not put forward, as we do, an explicit argument to the effect that such disputes can be *persistent* and still have a point.[36] More importantly, all of them seem to hold that the point of a meta-conceptual dispute is always somehow pragmatic rather than deep and substantial.[37] The plausibility of

36 Plunkett and Sundell (2013, 241–244) do claim that metaconceptual disputes are *worth having*. Plunkett (2015) argues that much philosophy is (at least implicitly) metaconceptual. However, as we have seen, a dispute can be *interesting* and hence worth having without *having a point* (see 2015, 4). It can even be worth having while being non-empty, non-verbal, and non-relativist but pointless (see 2015, 8).

37 This is connected to the claimed Carnapian inheritance of the conceptual engineering literature and to the claimed Wittgensteinian inheritance of the deep disagreement literature. See especially Shields (2021).

doi: 10.48106/dial.v76.i1

the non-deflationist view of mathematics strongly suggests that they are wrong.

It is also worth noting that we have hinted at an argument for the pervasive character of meta-analytic disputes just above, but that this argument—call it the pervasiveness argument—is quite different from those typically proposed in the metalinguistic and conceptual ethics literature. First, in this literature, meta-analytic disputes are always construed as metaconceptual or metalinguistic. We saw that there are other construals of meta-analytic disputes. Second, the most thorough arguments for the pervasive character of metaconceptual and metalingustic disputes essentially rely on the linguistic data surrounding some (potentially pointless) ordinary as well as philosophical disputes. Plunkett's (2015) important argument in this vein is a case in point, insofar as it is a linguistic argument applied to metaphilosophical questions. Roughly, his argument is that:

(i) Some linguistic data suggest that a given exchange is a dispute whose parties really disagree (i.e., they do not misunderstand each other), but mean different things by the disputed sentence.
(ii) The claim that their dispute is a metalinguistic negotiation can explain these data, and it can explain them more simply than the claim that the dispute is relativist or empty, which relies on complex non-standard semantic frameworks (such as recent brands of expressivism or relativism) (2015, 848–849).

Our argument relies partly on linguistic data as well, to wit, the data surrounding the Functions Controversy. It relies mostly, however, on epistemological and historical considerations to the effect that:

- Some persistent meta-analytic disputes have proved to have a point (the Functions Controversy).
- The meta-analytic reading of a persistent virtuous dispute allows us to defuse the best arguments for the charge that it is pointless, in answer to the pessimistic challenge.

Accordingly, the ground for ruling out the rival relativist or expressivist analyses is not the greater complexity, but the implication of pointlessness carried by these alternative interpretations. One might see our pervasiveness argument as contributing to the metalinguistic negotiation literature by providing an additional, optimistic reason to believe that many scientific and philosophical disputes are implicitly meta-analytic (and thus maybe metalinguistic and metaconceptual) because they persist and have a point. And of course, our main argument strengthens the interest of such disputes, as it shows that they can have a point even though they are persistent.

8 Conclusion

In this article, we examined and rejected the widespread imputation that persistent disputes are pointless. Thus, we characterized pointless disputes, put forward a typology, and reconstructed the strongest pessimistic argument against the claim that persistent disputes might have a point. To defuse the pessimistic argument, we proposed a meta-analytic reading of a concrete example: the illustrious "Functions" controversy. In general, when a dispute is meta-analytic, disputants disagree about which understanding or set of analytical truths among different candidates is the best one. The epistemic difficulty of settling the disagreement at this level is what renders their dispute persistent. Significantly, however, it does not render it pointless, as this collective task is achievable in principle.

If this is true, then one should not have unnecessarily sanguine expectations of the time it takes to settle such a dispute. To paraphrase Hegel, who might here be classified as one of the greatest optimists in the history of philosophy, one should trust the "power of the negative," for, in some instances, the very negativity of a sustained disagreement may strengthen the natural power of reason.*

Alexandre Billon
0000-0001-8905-0298
Université de Lille
alexandre.billon@polytechnique.org

* We began working on this project a long time ago, and given our differences in occupations, the first author could contribute way more to the paper. We benefited from the insightful comments of two referees of *Ergo* and four referees of this journal. Without their help the paper would have been weaker and much less interesting. This paper has also been supported by the European Center for Humanities and Social Sciences (MESHS-Lille, France) and by the Regional Council Hauts-de-France.

doi: 10.48106/dial.v76.i1

Philippe Vellozzo
0009-0007-2358-9283
Université de Bordeaux
ph.vellozzo@gmail.com

References

Alston, William P. 1991. *Perceiving God: The Epistemology of Religious Experience*. Ithaca, New York: Cornell University Press, doi:10.7591/9780801471254.

—. 2005. *Beyond "Justification": Dimensions of Epistemic Evaluation*. Ithaca, New York: Cornell University Press, doi:10.7591/9781501720574.

Ayer, Alfred Jules, ed. 1959. *Logical Positivism*. Glencoe, Illinois: Free Press.

Ball, Derek. 2020. "Revisionary Analysis without Meaning Change (Or, Could Women Be Analytically Oppressed?)." in *Conceptual Engineering and Conceptual Ethics*, edited by Alexis Burgess, Herman Cappelen, and David Plunkett, pp. 35–58. Oxford: Oxford University Press, doi:10.1093/oso/9780198801856.003.0002.

Basham, Lee. 2001. "Living with the Conspiracy." *Philosophical Forum (Boston)* 32(3): 265–280, doi:10.1111/0031-806X.00065.

Bayne, Tim[othy John]. 2007. "Hypnosis and the Unity of Consciousness." in *Hypnosis and Conscious States: The Cognitive Neuroscience Perspective*, edited by Graham A. Jamieson, pp. 93–109. Oxford: Oxford University Press, doi:10.1093/oso/9780198569794.003.0006.

Bayne, Tim[othy John] and Pacherie, Elisabeth. 2004. "Experience, Belief, and the Interpretive Fold." *Philosophy, Psychiatry, and Psychology* 11(1): 81–86, doi:10.1353/ppp.2004.0034.

Belleri, Delia. forthcoming. "Downplaying the Change of Subject Objection to Conceptual Engineering." *Inquiry*, doi:10.1080/0020174X.2021.1908161.

Bennett, Karen. 2009. "Composition, Colocation, and Metaontology." in *Metametaphysics: New Essays on the Foundations of Ontology*, edited by David J. Chalmers, David Manley, and Ryan Wasserman, pp. 38–76. Oxford: Oxford University Press, doi:10.1093/oso/9780199546046.003.0002.

Block, Ned. 2005. "Two Neural Correlates of Consciousness." *Trends in Cognitive Sciences* 9(2): 46–52. Reprinted in Block (2007a, 342–361), doi:10.1016/j.tics.2004.12.006.

—. 2007a. *Consciousness, Function, and Representation. Collected Papers, Volume 1*. Cambridge, Massachusetts: The MIT Press, doi:10.7551/mitpress/2111.001.0001.

—. 2007b. "Consciousness, Accessibility, and the Mesh Between Psychology and Neuroscience." *Behavioral and Brain Sciences* 30: 481–548, doi:10.1017/s0140525x07002786.

—. 2011. "Perceptual Consciousness Overflows Cognitive Access." *Trends in Cognitive Sciences* 15(12): 567–575, doi:10.1016/j.tics.2011.11.001.

Buchanan, Ray. 2014. "Reference, Understanding, and Communication." *Australasian Journal of Philosophy* 92(1): 55–70, doi:10.1080/00048402.2013.788526.

Burgess, Alexis and Plunkett, David. 2013. "Conceptual Ethics I." *Philosophy Compass* 8(12): 1091–1101, doi:10.1111/phc3.12086.

Cappelen, Herman. 2018. *Fixing Language: An Essay on Conceptual Engineering*. Oxford: Oxford University Press, doi:10.1093/oso/9780198814719.001.0001.

Carnap, Rudolf. 1931. "Überwindung der Metaphysik durch logische Analyse der Sprache." *Erkenntnis* 2: 219–241. Reprinted and translated in Ayer (1959, 60–81), doi:10.1007/bf02028153.

—. 1947. *Meaning and Necessity: A Study in Semantics and Modal Logic*. 1st ed. Chicago, Illinois: University of Chicago Press. Second edition: Carnap (1956).

—. 1950. "Empiricism, Semantics, and Ontology." *Revue Internationale de Philosophie* 4(11): 20–40. Reprinted in Carnap (1956, 205–221).

—. 1956. *Meaning and Necessity: A Study in Semantics and Modal Logic*. 2nd ed. Chicago, Illinois: University of Chicago Press. Enlarged edition of Carnap (1947).

Castelvecchi, Davide. 2020. "Mathematical Proof That Rocked Number Theory Will Be Published." *Nature* 580(7802): 177, doi:10.1038/d41586-020-00998-2.

Chalmers, David J. 2011. "Verbal Disputes." *The Philosophical Review* 120(4): 515–566, doi:10.1215/00318108-1334478.

Christensen, David. 2007. "Epistemology of Disagreement: The Good News." *The Philosophical Review* 116(2): 187–217, doi:10.1215/00318108-2006-035.

Dehaene, Stanislaus and Naccache, Lionel. 2001. "Towards a Cognitive Neuroscience of Consciousness: Basic Evidence and a Workspace Framework." *Cognition* 79: 1–37, doi:10.1016/s0010-0277(00)00123-2.

Elga, Adam Newman. 2007. "Reflection and Disagreement." *Noûs* 41(3): 478–502, doi:10.1111/j.1468-0068.2007.00656.x.

Enoch, David. 2011. *Taking Morality Seriously: A Defense of Robust Realism*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199579969.001.0001.

Feldman, Richard H. 2003. *Epistemology*. Foundations of Philosophy Series. Upper Saddle River, New Jersey: Prentice-Hall, Inc.

—. 2006. "Epistemological Puzzles about Disagreement." in *Epistemology Futures*, edited by Stephen Cade Hetherington, pp. 216–236. Oxford: Oxford University Press, doi:10.1093/oso/9780199273317.003.0013.

Feltz, Adam and Cokely, Edward T. 2013. "Predicting Philosophical Disagreement." *Philosophy Compass* 8(10): 978–989, doi:10.1111/phc3.12076.

Fine, Kit. 2007. *Semantic Relationism*. Oxford: Blackwell Publishers, doi:10.1002/9780470690826.

Flocke, Vera. 2020. “Carnap’s Noncognitivism about Ontology.” *Noûs* 54(3): 527–548, doi:10.1111/nous.12267.

Fogelin, Robert J. 1985. “The Logic of Deep Disagreements.” *Informal Logic* 7(1): 1–8, doi:10.22329/il.v7i1.2696.

Frances, Bryan. 2014. *Disagreement*. Cambridge: Polity Press, https://archive.org/details/disagreement0000fran.

Frege, Gottlob. 1918. “Der Gedanke. Eine logische Untersuchung.” *Beiträge zur Philosophie des Deutschen Idealismus* 1: 58–77. Ed. by Arthur Hoffmann and Horst Engert; Erfurt: Verlag der Keyserschen Buchhandlung, reprinted in Frege (1993).

—. 1956. “The Thought: A Logical Inquiry.” *Mind* 65(259): 289–311. Translation of Frege (1918) by Anthony Quinton and Marcelle Quinton, doi:10.1093/mind/65.1.289.

—. 1993. *Logische Untersuchungen*. Göttingen: Vandenhoeck & Ruprecht. Edited by Günther Patzig.

Gallie, Walter Bryce. 1956. “Essentially Contested Concepts.” *Proceedings of the Aristotelian Society* 56: 167–198, doi:10.1093/aristotelian/56.1.167.

Godden, David M. and Brenner, William H. 2010. “Wittgenstein and the Logic of Deep Disagreement.” *Cogency. Journal of Reasoning and Argumentation* 2(2): 41–80.

Goodman, Nelson. 1978. *Ways of Worldmaking*. Indianapolis, Indiana: Hackett Publishing Co.

Gottlieb, Joseph. 2018. “Verbal Disputes in the Theory of Consciousness.” *Ergo* 5(12): 319–347, doi:10.3998/ergo.12405314.0005.012.

Gray, Jeremy J. 1991. “Did Poincaré Say ‘Set Theory Is a Disease’?” *The Mathematical Intelligencer* 13(1): 19–22, doi:10.1007/BF03024067.

Hawkins, Thomas. 1975. *Lebesgue’s Theory of Integration: Its Origins and Development*. New York: Chelsea Publishing Company.

Hawthorne, John. 2009. “Superficialism in Ontology.” in *Metametaphysics: New Essays on the Foundations of Ontology*, edited by David J. Chalmers, David Manley, and Ryan Wasserman, pp. 213–230. Oxford: Oxford University Press, doi:10.1093/oso/9780199546046.003.0007.

Hirsch, Eli. 2009. “Ontology and Alternative Languages.” in *Metametaphysics: New Essays on the Foundations of Ontology*, edited by David J. Chalmers, David Manley, and Ryan Wasserman, pp. 231–259. Oxford: Oxford University Press, doi:10.1093/oso/9780199546046.003.0008.

—. 2016. “Three Degrees of Carnapian Tolerance.” in *Ontology after Carnap*, edited by Stephan Blatti and Sandra Lapointe, pp. 105–121. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199661985.003.0006.

Hohwy, Jakob and Rosenberg, Raben. 2005. “Unusual Experiences, Reality Testing, and Delusions of Alien Control.” *Mind and Language* 20(2): 141–162, doi:10.1111/j.0268-1064.2005.00280.x.

Hume, David. 1758. *An Enquiry Concerning Human Understanding*. 2nd ed. London: Andrew Millar of the Strand.

van Inwagen, Peter. 2017. *Thinking about Free Will*. Cambridge: Cambridge University Press, doi:10.1017/9781316711101.

James, William. 1902. *The Varieties of Religious Experience: A Study in Human Nature*. London: Longmans, Green & Co.

Kahneman, Daniel, Slovic, Paul and Tversky, Amos, eds. 1982. *Judgment under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press, doi:10.1017/cbo9780511809477.

Kelly, Thomas. 2005. "The Epistemic Significance of Disagreement." in *Oxford Studies in Epistemology*, volume I, edited by Tamar Szabó Gendler and John Hawthorne, pp. 167–196. Oxford: Oxford University Press, doi:10.1093/oso/9780199285891.003.0007.

—. 2010. "Peer Disagreement and Higher-Order Evidence." in *Disagreement*, edited by Richard H. Feldman and Ted A. Warfield, pp. 111–174. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199226078.003.0007.

Kleiner, Israel. 1989. "Evolution of the Function Concept: A Brief Survey." *The College Mathematics Journal* 20(4): 282–300, doi:10.2307/2686848.

Kouider, Sid, Dehaene, Stanislaus, Jobert, Antoinette and Le Bihan, Denis. 2007. "Cerebral Bases of Subliminal and Supraliminal Priming During Reading." *Cerebral Cortex* 17(9): 2019–2029, doi:10.1093/cercor/bhl110.

Kouider, Sid, Sackur, Jérôme and Gardelle, Vincent de. 2012. "Do we Still Need Phenomenal Consciousness? Comment on Block (2011)." *Trends in Cognitive Sciences* 16(3): 140–141, doi:10.1016/j.tics.2012.01.003.

Kriegel, Uriah. 2013. "The Epistemological Challenge of Revisionary Metaphysics." *Philosophers' Imprint* 13(12), http://hdl.handle.net/2027/spo.3521354.0013.012.

Kuhn, Thomas S. 1962a. *The Structure of Scientific Revolutions*. Chicago, Illinois: University of Chicago Press. Separate book publication of Kuhn (1962b); second edition: Kuhn (1970).

—. 1962b. "The Structure of Scientific Revolutions." *International Encyclopedia of Unified Science* 2(2). Reprinted in Neurath, Carnap and Morris (1970, 53–272).

—. 1970. *The Structure of Scientific Revolutions*. 2nd ed. Chicago, Illinois: University of Chicago Press. Enlarged, second edition of Kuhn (1962a).

—. 2012. *The Structure of Scientific Revolutions*. Chicago, Illinois: University of Chicago Press. 50th Anniversary Edition, with an introductory essay by Ian Hacking; original publication: Kuhn (1962a).

Lakatos, Imre. 1976. *Proofs and Refutations: The Logic of Mathematical Discovery*. Cambridge: Cambridge University Press. Republished as Lakatos (2015).

—. 2015. *Proofs and Refutations: The Logic of Mathematical Discovery*. Cambridge: Cambridge University Press. Republication of Lakatos (1976), edited by John Worrall and Elie Zahar, doi:10.1017/cbo9781316286425.

Lamme, V. A. F. 2004. “Separate Neural Definitions of Visual Consciousness and Visual Attention: A Case for Phenomenal Awareness.” *Neural Networks* 17(5/6): 861–872, doi:10.1016/j.neunet.2004.02.005.

LaPorte, Joseph. 2004. *Natural Kinds and Conceptual Change*. Cambrige Studies in Philosophy and Biology. Cambridge: Cambridge University Press, doi:10.1017/cbo9780511527319.

Lebesgue, Henri. 1905. “Recherches sur la convergence des séries de Fourier.” *Mathematische Annalen* 61(2): 251–280, doi:10.1007/BF01457565.

Levy, Neil. 2019. “Due Deference to Denialism: Explaining Ordinary People’s Rejection of Established Scientific Findings.” *Synthese* 196(1): 313–327, doi:10.1007/s11229-017-1477-x.

Locke, John. 1689. *An Essay concerning Human Understanding*. London: Eliz. Holt, for Thomas Basset, at the George in Fleet Street.

—. 1975. *An Essay concerning Human Understanding*. Oxford: Oxford University Press. Edited with a foreword by Peter H. Nidditch, doi:10.1093/actrade/9780198243861.book.1.

MacFarlane, John. 2014. *Assessment Sensitivity: Relative Truth and Its Applications*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199682751.001.0001.

Maloney, Mame. 2008. “Pathological Functions in the 18th and 19th Centuries.” Unpublished manuscript, https://historyofmathematics.org/wp-content/uploads/2013/09/2008-Maloney.pdf.

Naccache, Lionel. 2005. “Visual Phenomenal Consciousness: A Neurological Guided Tour.” *Progress in Brain Research* 150: 185–195, doi:10.1016/S0079-6123(05)50013-X.

Neurath, Otto, Carnap, Rudolf and Morris, Charles W., eds. 1970. *Foundations of the Unity of Science: Toward an International Encyclopedia of Unified Science. Volume II*. Chicago, Illinois: University of Chicago Press.

Nguyen, C. Thi. 2020. “Echo Chambers and Epistemic Bubbles.” *Episteme* 17(2): 141–161, doi:10.1017/epi.2018.32.

Plantinga, Alvin. 2000. *Warranted Christian Belief*. Oxford: Oxford University Press, doi:10.1093/0195131932.001.0001.

Plunkett, David. 2015. “Which Concepts Should we Use? Metalinguistic Negotiations and the Methodology of Philosophy.” *Inquiry* 58(7/8): 828–874, doi:10.1080/0020174X.2015.1080184.

Plunkett, David and Sundell, Timothy. 2013. “Disagreement and the Semantics of Normative and Evaluative Terms.” *Philosophers’ Imprint* 13(23), http://hdl.handle.net/2027/spo.3521354.0013.023.

Poincaré, Henri. 1899. “L’œuvre mathématique de Weierstrass.” *Acta Mathematica* 22(1): 1–18.

—. 1952. *Science and Method*. Mineola, New York: Dover Publications.

Priest, Graham. 1987. *In Contradiction: A Study of the Transconsistent*. Nijhoff International Philosophy Series n. 39. Den Haag: Martinus Nijhoff Publishers.

—. 2006. *In Contradiction: A Study of the Transconsistent*. Oxford: Oxford University Press. First edition: Priest (1987), doi:10.1093/acprof:oso/9780199263301.001.0001.

Pryor, James. 2016. "Mental Graphs." *The Review of Philosophy and Psychology* 7(2): 309–341, doi:10.1007/s13164-015-0280-1.

Récanati, François. 2012. *Mental Files*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199659982.001.0001.

Rescher, Nicholas. 1985. *The Strife of Systems: An Essay on the Grounds and Implications of Philosophical Diversity*. Pittsburgh, Pennsylvania: University of Pittsburgh Press.

Richards, Joan L. 2006. "Historical Mathematics in the French Eighteenth Century." *Isis* 97(4): 700–713, doi:10.1086/509951.

Rosenthal, David M. 2002. "How Many Kinds of Consciousness?" *Consciousness and Cognition* 11(4): 653–665, doi:10.1016/S1053-8100(02)00017-X.

Schellenberg, J. L. 2007. *The Wisdom to Doubt: A Justification of Religious Skepticism*. Ithaca, New York: Cornell University Press, doi:10.7591/9780801462399.

Schlick, Moritz. 1930. "Die Wende der Philosophie." *Erkenntnis* 1(1): 4–11. Reprinted as "The Turning Point in Philosophy" in Ayer (1959, 53–59), doi:10.1007/bf00208605.

Schroeter, Laura and Schroeter, François. 2014. "Normative Concepts: A Connectedness Model." *Philosophers' Imprint* 14(25), http://hdl.handle.net/2027/spo.3521354.0014.025.

—. 2016. "Semantic Deference versus Semantic Coordination." *American Philosophical Quarterly* 53(2): 193–210.

Sergent, Claire and Dehaene, Stanislaus. 2005. "Neural Processes Underlying Conscious Perception: Experimental Findings and a Global Neuronal Workspace Framework." *Journal of Physiology* 98(4/6): 374–384, doi:10.1016/j.jphysparis.2005.09.006.

Shields, Matthew. 2021. "On the Pragmatics of Deep Disagreement." *Topoi* 40(5): 999–1015, doi:10.1007/s11245-018-9602-0.

Sider, Theodore. 2009. "Ontological Realism." in *Metametaphysics: New Essays on the Foundations of Ontology*, edited by David J. Chalmers, David Manley, and Ryan Wasserman, pp. 384–423. Oxford: Oxford University Press, doi:10.1093/oso/9780199546046.003.0013.

Sosa, Ernest. 2010. "The Epistemology of Disagreement." in *Social Epistemology*, edited by Adrian Haddock, Alan Millar, and Duncan Pritchard, pp. 278–297. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199577477.003.0014.

Stoljar, Daniel. 2017. *Philosophical Progress: In Defence of a Reasonable Optimism*. Oxford: Oxford University Press, doi:10.1093/oso/9780198802099.001.0001.

Taschek, William W. 1995. "Belief, Substitution, and Logical Structure." *Noûs* 29(1): 71–95, doi:10.2307/2215727.

Thomasson, Amie L. 2009. "Answerable and Unanswerable Questions." in *Metametaphysics: New Essays on the Foundations of Ontology*, edited by David J. Chalmers, David Manley, and Ryan Wasserman, pp. 444–471. Oxford: Oxford University Press, doi:10.1093/oso/9780199546046.003.0015.

—. 2020. "A Pragmatic Method for Conceptual Ethics." in *Conceptual Engineering and Conceptual Ethics*, edited by Alexis Burgess, Herman Cappelen, and David Plunkett, pp. 435–458. Oxford: Oxford University Press, doi:10.1093/oso/9780198801856.003.0021.

Wedgwood, Ralph. 2019. "Moral Disagreement and Inexcusable Irrationality." *American Philosophical Quarterly* 56(1): 97–108, doi:10.2307/45128646.

Williams, Bernard Arthur Owen. 1985. *Ethics and the Limits of Philosophy*. London: Fontana Press.

—. 2011. *Ethics and the Limits of Philosophy*. London: Routledge. With a commentary on the text by A.W. Moore and a foreword by Jonathan Lear; original publication: Williams (1985), doi:10.4324/9780203828281.

Williamson, Timothy. 2000. *Knowledge and Its Limits*. Oxford: Oxford University Press, doi:10.1093/019925656X.001.0001.

—. 2007. *The Philosophy of Philosophy*. Oxford: Blackwell Publishers. Second edition: Williamson (2022), doi:10.1002/9780470696675.

—. 2022. *The Philosophy of Philosophy*. 2nd ed. The Blackwell / Brown Lectures in Philosophy. Hoboken, New Jersey: Wiley-Blackwell. First edition: Williamson (2007), doi:10.1002/9781119616702.

Wittgenstein, Ludwig. 1929. "Some Remarks on Logical Form." *Proceedings of the Aristotelian Society, Supplementary Volume* 9: 162–171, doi:10.1093/aristoteliansupp/9.1.162.

Yablo, Stephen. 2009. "Must Existence-Questions Have Answers?" in *Metametaphysics: New Essays on the Foundations of Ontology*, edited by David J. Chalmers, David Manley, and Ryan Wasserman, pp. 507–526. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199266487.003.0013.

Youschkevitch, Adolph-Andrei Pavlovich. 1976. "The Concept of Function up to the Middle of the 18th and the 19th Century." *Archive for the History of the Exact Sciences* 16(1): 37–85, doi:10.1007/BF00348305.

Zeki, Semir and Ffytche, Dominic H. 1998. "The Riddoch Syndrome: Insights in the Neurobiology of Conscious Vision." *Brain: A Journal of Neurology* 121(1): 25–45, doi:10.1093/brain/121.1.25.

Published by *Philosophie.ch*

Verein philosophie.ch
Fabrikgässli 1
2502 Biel/Bienne
Switzerland
dialectica@philosophie.ch

https://dialectica.philosophie.ch/

ISSN 0012-2017

ISBN 1234-5678

Dialectica is supported by the Swiss Academy of Humanities and Social Sciences.

Abstracting and Indexing Services

The journal is indexed by the Arts and Humanities Citation Index, Current Contents, Current Mathematical Publications, Dietrich's Index Philosophicus, IBZ — Internationale Bibliographie der Geistes- und Sozialwissenschaftlichen Zeitschriftenliteratur, Internationale Bibliographie der Rezensionen Geistes- und Sozialwissenschaftlicher Literatur, Linguistics and Language Behavior Abstracts, Mathematical Reviews, MathSciNet, Periodicals Contents Index, Philosopher's Index, Repertoire Bibliographique de la Philosophie, Russian Academy of Sciences Bibliographies.

Contents