Andrea Giuseppe Ragno*

Coherence and Reduction

<https://doi.org/10.1515/krt-2021-0031>

Published online December 7, 2021

Abstract: Synchronic intertheoretic reductions are an important field of research in science. Arguably, the best model able to represent the main relations occurring in this kind of scientific reduction is the Nagelian account of reduction, a model further developed by Schaffner and nowadays known as the generalized version of the Nagel–Schaffner model (GNS). In their article (2010), Dizadji-Bahmani, Frigg, and Hartmann (DFH) specified the two main desiderata of a reduction *à la* GNS: confirmation and coherence. DFH first and, more rigorously, Tešić (2017) later analyse the confirmatory relation between the reducing and the reduced theory in terms of Bayesian confirmation theory. The purpose of this article is to analyse and compare the degree of coherence between the two theories involved in the GNS before and after the reduction. For this reason, in the first section, I will be looking at the reduction of thermodynamics to statistical mechanics and use it as an example to describe the GNS. In the second section, I will introduce three coherence measures which will then be employed in the comparison. Finally, in the last two sections, I will compare the degrees of coherence between the reducing and the reduced theory before and after the reduction and use a few numerical examples to understand the relation between coherence and confirmation measures.

Keywords: scientific reduction, coherence, Bayesian networks, confirmation

1 Introduction

Synchronic intertheoretic reductions, namely those reductions between pairs of theories whose domains of applicability largely overlap, are an important field of research in science. Arguably, the best model offered in the philosophical literature which is able to track the main relations occurring in this kind of scientific reduction is the Nagelian account of reduction (Nagel 1961, 1974), a model which was further developed by Schaffner (1967, 1969, 1974, 1977, 1993) and is nowadays known as the GNS, acronym for *generalized version of the Nagel–Schaffner model*. In short, according to the GNS, reducing a theory T_A to another theory T_B is possible

***Corresponding author: Andrea Giuseppe Ragno**, Munich Centre for Mathematical Philosophy, LMU, Ludwigstraße 31, 80539 Munich, Germany, E-mail: Andrea.Ragno@campus.lmu.de

if and only if the laws of T_A are derivable from T_B with the help of bridge laws. Seven classes of criticisms were put forward against this model of reduction. In their article, Dizadji-Bahmani, Frigg, and Hartmann (henceforth ‘DFH’) respond to these attacks by offering some needed modifications to the GNS (Dizadji-Bahmani, Frigg, and Hartmann 2010, 2011). Amongst these clarifications, DFH specify the two main desiderata one might want from the GNS: confirmation and coherence. So, DFH first and, more rigorously, Tešić (2019) later analyse the confirmatory relation between the reducing and the reduced theory in terms of Bayesian confirmation theory. More precisely, evidence confirming one theory also confirms the other theory and vice versa given that after the reduction of one to the other they become connected and share their evidence. The purpose of this article is, instead, to compare the different degrees of coherence between the reducing and the reduced theory before and after the reduction through the Bayesian analysis firstly sketched by DFH and then corrected by Tešić.

For this reason, I will prepare the ground by looking at the classic putative example of a reduction *à la* GNS: the reduction of thermodynamics (henceforth ‘TD’) to statistical mechanics (henceforth ‘SM’) (Section 2.1). Then, I briefly present the Bayesian analysis of the relation between the reducing and the reduced theory given by DFH (Section 2.2) and corrected by Tešić (Section 2.3): this analysis is crucial to probabilistically represent the GNS. Such Bayesian representation will indeed be used in the comparison of the two different degrees of coherence amongst the two theories in question: in fact, the coherence measures which I will take into account for comparing the degrees of coherence are probabilistic. In particular, three coherence measures will be discussed (Section 3); these measures might present counter-intuitive results in certain contexts. The issue here is whether they will report different outcomes in the case of GNS or not, that is, whether one coherence measure might report that the set of theories after the reduction coheres more than the set of theories prior to the reduction, and another reports the opposite. The goal would be having all the coherence measure reporting the same result. In the fourth section, I will show that they actually outline similar results under an assumption formulated by Bovens and Hartmann for their coherence measure (Bovens and Hartmann 2003): when one does considerations on coherence, the several sources reporting evidence should be understood as equally and partially reliable, and taken into account independently because the observations being reported is what is at stake when analysing its coherence. I will therefore compare the degrees of coherence among the reducing and the reduced theory pre- and post-reduction (Section 4). Under the assumptions of the Bayesian analysis provided by DFH and Tešić, and under the assumption of the coherence measure of Bovens and Hartmann, the coherence measures will provide two conditions which the two theories of the GNS likely meet in light of

their conditional dependency, which is in turn due to the reduction of a theory to the other. Finally, in the last section, I present some numerical examples aimed at analysing the relation between coherence measures and confirmation measures in the context of intertheoretic reduction as designed by DFH and Tešić (Section 5).

2 The Generalized Version of the Nagel–Schaffner Model

In this section, I will present the GNS by looking at a putative example of GNS reduction. Then, I will outline the discussion between DFH's and Tešić's Bayesian analysis. The latter does not present knockdown arguments for the former. Rather, it corrects a couple of flaws DFH's analysis has by reviewing the conceptual character of bridge laws.

2.1 An Example of a GNS Reduction

TD is a branch of physics which describes those phenomena observable in macroscopic systems (e.g. solids, liquids, gases, plasma) and their relation to energy, radiation, and properties of matter. The behaviour of these entities can be expressed by the four laws of thermodynamics which make use of macroscopic properties (e.g. pressure, temperature). Yet, such behaviour can also be explained in terms of their microscopic constituents by SM. In fact, based on statistical methods, probability theory and microscopic physical laws, SM explains the behaviour of macroscopic systems in terms of those dynamical laws governing its microscopic quantities (e.g. molecules, particles). The laws of TD can therefore be expressed in terms of the laws of SM. This hints at our first definition of Nagelian reduction, according to which reducing a theory T_A to another theory T_B is possible if and only if the laws of T_A are derivable from T_B with the help of bridge laws, which are empirical facts linking concepts of the reduced theory to terms of the reducing theory. This, however, is not sufficient to properly describe a successful reduction. Consider an example of reduction between TD and SM, namely the Boyle–Charles law (Dizadji-Bahmani, Frigg, and Hartmann 2010, pp. 395–396),¹ to better understand the role of bridge laws. The Boyle–Charles law states that the

¹ In presenting the Boyle–Charles law and its derivation, I closely follow (Dizadji-Bahmani, Frigg, and Hartmann 2010, 2011). For a discussion in greater details, consult Feynman, Leighton, and Sands (1964, chapter 39) and Pauli (1973, pp. 94–103).

temperature T of a gas is directly proportional to the product of the values of its pressure p and the volume V over which it evenly distributed:

$$pV = kT, \quad (1)$$

where k is a constant. This law together with some specific conditions (i.e. gas in thermodynamic equilibrium with the surrounding environment and relatively low pressure) forms the core of the *thermal theory of the ideal gas*. In SM, there is a corresponding theory for ideal gases: the *kinetic theory of the ideal gas*. This theory describes the motion of n particles with mass m of a gas spread over the volume of, for example, a vessel according to Newtonian mechanics. The theory includes two assumptions:

- i. the gas should be ideal to the extent that its molecules, which collide elastically, are point particles;
- ii. the three components of the velocity $\vec{v}$ (v_x, v_y, v_z) should be evenly distributed (i.e. there is no favoured direction).

Following the definition of pressure in Newtonian physics and the first assumption, the gas hitting a wall of the vessel exerts a pressure:

$$p \doteq \frac{mn}{V} \langle v_z^2 \rangle, \quad (2)$$

where $\langle v_z^2 \rangle$ is the average of the square of v_z , a particle's velocity in z -direction with respect to the x - y plane of the wall. After few other calculations² and following the second assumption, the left-hand term of the equation in the Boyle–Charles law can be expressed as:

$$pV \doteq \frac{2n}{3} \langle E_{\text{kin}} \rangle, \quad (3)$$

where $n \langle E_{\text{kin}} \rangle$ is the average kinetic energy of the gas. T can therefore be seen as:

$$T = \frac{2n}{3k} \langle E_{\text{kin}} \rangle. \quad (4)$$

This process shows how to derive the Boyle–Charles law from the laws of Newtonian physics. In fact, first, a particular theory, the *thermal theory of the ideal gas*, (here, eq. (3)), was derived by combining Newtonian physics with the two assumptions of the *kinetic theory of the ideal gas*. Second, eq. (4), that would stand as a bridge law in the GNS, has connected relevant terms, such as T and $\langle E_{\text{kin}} \rangle$, and it

² These calculations will not be mentioned here because their explanation goes beyond the scope of this article. See Feynman, Leighton, and Sands (1964, chapter 39) for a more precise understanding of this derivation.

has yielded a version of the Boyle–Charles law bounded to some conditions. Finally, it has been shown that this particular version of the Boyle–Charles law bounded by particular conditions is strongly analogous (or even coincide) to the standard version of the Boyle–Charles law. Nowadays, scientists consider the reduction of TD to SM successful.³ The reduction of TD to SM is considered to be a synchronic intertheoretic reduction, namely a reductive relation between two coexisting theories which deal with different levels of a largely overlapping domain. In this reduction, the concepts and the laws of one theory can respectively be expressed as the concepts and be derived from the laws of the more fundamental theory. Accordingly, a correct reduction of TD to SM involves the derivations of the laws of TD from the laws governing the microconstituents of macroscopic systems together with probabilistic assumptions. For this reason, DFH suggest that this reductive relation resembles the GNS, which applies to synchronic intertheoretic reductions.⁴

In the reductive relation, TD is the reduced theory $\mathbf{T_P}$ and SM is the reducing one $\mathbf{T_F}$.⁵ According to the GNS, $\mathcal{T_P}$ corresponds to the set of empirical propositions $\{T_P^{(1)}, \dots, T_P^{(n)}\}$ associated with $\mathbf{T_P}$. Sp, $\mathcal{T_F}$ is the set of empirical propositions $\{T_F^{(1)}, \dots, T_F^{(n)}\}$ associated with $\mathbf{T_F}$. Here, the empirical propositions of $\mathcal{T_P}$ and $\mathcal{T_F}$ are the various laws of the theories.⁶ As shown in the case of the Boyle–Charles law, the reduction of $\mathcal{T_P}$ to $\mathcal{T_F}$ would then follow three steps.

1. Use auxiliary assumptions to help deriving a restricted version of each element $T_F^{(i)}$ of $\mathcal{T_F}$.⁷ Let $\mathcal{T_F}^* \doteq \{T_F^{*(1)} \dots T_F^{*(n_F)}\}$ be the set of the restricted versions.
2. Adopt bridge laws in order to connect the relevant terms which are not share by the vocabularies of the theories involved.⁸ Substituting the terms in $\mathcal{T_F}^*$ with the terms from $\mathcal{T_P}$ shows that the bridge laws yield the set $\mathcal{T_P}^* \doteq \{T_P^{*(1)} \dots T_P^{*(n_P)}\}$.
3. Show that each element of $\mathcal{T_P}^*$ is strongly analogous to the corresponding element of $\mathcal{T_P}$.

3 Although, the foundations of SM remain a controversial topic.

4 DFH are known for having responded to seven classes of criticisms put forward against the GNS by slightly modifying this version. In this article, we shall refer to their conception of the GNS simply as reduction. See Dizadji-Bahmani, Frigg, and Hartmann (2010) for arguments why the GNS corrected by DFH is our best reduction model for mapping several scientific reductive relations. See, instead, Sarkar (2015) for reasons why one would want to stick with Nagel's original model of reduction (Nagel 1961, 1974).

5 While, the index $\mathbf{P}$ stands for phenomenological, the index $\mathbf{F}$ stands for fundamental.

6 This, however, does not mean that $\mathcal{T_A} = \mathbf{T_A}$ since a theory includes more elements than just its laws.

7 See eq. (3).

8 See eq. (4).

If these conditions are satisfied, it is believed that $\mathcal{T}_P$ is reduced to $\mathcal{T}_F$ with respect to the GNS.

2.2 The Bayesian Analysis of DFH

Amongst the desiderata of reductions in science, coherence and confirmation are the main ones (Dizadji-Bahmani, Frigg, and Hartmann 2010, 2011).⁹ Nagel (1961, p. 341)¹⁰ himself sensed that reduction should reconcile two self-consistent and well-confirmed theories whose domains of application (largely) overlap whenever the two sketch a contradictory view of the world. The example of the reduction of TD to SM perfectly fits into this picture. In fact, TD (here, $\mathcal{T}_P$) and SM (here, $\mathcal{T}_F$) should be consistent with each other and evidence confirming TD should support SM, and vice versa. Obviously, both criteria must be met after the reduction occurs. DFH (Dizadji-Bahmani, Frigg, and Hartmann 2011) use Bayesian networks to represent the relation between reduction and confirmation.¹¹ In fact, Bayesian networks, as illustrated in Nagel (1974), exploit (un)conditional independences and dependencies in order to represent large instances in little space by graphical structures and to perform probabilistic inferences in little time. The type of

9 It is important to remark that DFH introduce the desiderata of reductions while defending their account of bridge laws. In general, according to DFH, bridge laws are factual assertions, which fall into two categories: bridge laws which associate entities postulated by two theories, and bridge laws which associate properties. While entity association laws indicate identities and are internal to $\mathbf{T}_F$, property association laws are external to $\mathbf{T}_F$ and are therefore not required to express identities. Following this, a property in $\mathbf{T}_P$ can correspond to several properties in $\mathbf{T}_F$. This principle is known as multiple realisability (MR), a famous thesis in philosophy which claims that a single event can be implemented by different physical properties. DFH gathers four arguments according to which MR undermines GNS together with four counter-arguments, which sum up their position: amongst $\mathbf{T}_F$ and $\mathbf{T}_P$, bridge laws should associate not only entities, but also properties. Here, I briefly mention the first argument with its corresponding counter-argument. Critics of bridge laws believe indeed that it is untenable that properties in $\mathbf{T}_P$ are not wholly contained in $\mathbf{T}_F$. They suggest that properties in $\mathbf{T}_P$ must be identifiable with properties in $\mathbf{T}_F$. The principle of MR does not identify a property in $\mathbf{T}_P$ with a property in $\mathbf{T}_F$ and, thus, it weakens the claims made by the defenders of the GNS. DFH responds to this concern by stating why reductions are desired in science: consistency and confirmation. Explanations and strict identities between properties in $\mathbf{T}_P$ and $\mathbf{T}_F$ are not necessary. Yet, as it will soon be shown, it seems that the requirement of strict identities appear in the Bayesian analysis offered by DFH (Dizadji-Bahmani, Frigg, and Hartmann 2011). This is one of the reasons why Tešić (2019) is skeptic of such analysis and why I present and use his outlook in this article to prove the coherence amongst the two theories involved in the GNS.

10 Nagel included explanation amongst the desiderata of reductions. See Sarkar (2015) for details.

11 Later, I shall do the same for the relation between reduction and coherence.

statistical calculation involved in Bayesian networks is called Bayesian inference, that is, an inference in which Bayes' theorem is one of the main rules used to update the probability for a hypothesis as more evidence and information become available.¹² Bayesian networks are a type of probabilistic graphical model that uses Bayesian inference for probability computations. Confirming an hypothesis H with a piece of evidence E in Bayesian terms means having a conditional probability $P(H|E)$ larger than the prior probability $P(H)$. In other words, an hypothesis is confirmed by E if:

$$P(H|E) > P(H). \quad (6)$$

More precisely, a Bayesian Network is a *directed acyclical graph* (DAG),¹³ which satisfies the Markov condition¹⁴ and whose nodes represent discrete propositional variables, and edges capture their conditional independences and dependencies.¹⁵ To frame GNS in Bayesian terms, DFH introduce few simplifications which I will use throughout the article.

- To simplify the calculations, DFH assume that $\mathcal{T}_F$ and $\mathcal{T}_P$ have only one element, namely T_F and T_P respectively. Their corresponding propositional variables will be T_F and T_P respectively.
- The propositional variables represented by the nodes of a Bayesian network can take two values, i.e. T_F and $\neg T_F$. While the latter means that the proposition T_F is false, the former asserts that it is the case that T_F is true.
- The probability of every node can lie in the open interval $(0, 1)$. I set $\bar{j} = 1 - j$ for all parameters j , unless a parameter is a logical consequences of another variable: in this case, its conditional probability on such variable is 1.

12 Formally, the Bayes rule gives a value of $P(A|B)$:

$$\begin{aligned} P(A|B) &= \frac{P(B|A)P(A)}{P(B)} \\ &= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\neg A)P(\neg A)} \end{aligned} \quad (5)$$

13 Recall graph theory. A *directed graph* consists of an ordered pair (V, E) , where V is the finite set of nodes and E is the set of ordered pair of distinct elements of V , also known as edges. Let $(X, Y) \in E$: then, Y is a parent node of child node X if there is an edge from Y to X . A *directed acyclical graph* is a directed graph where there is no path from a parent node back to itself. Let $(X_1, X_2, \dots, X_K) \in E$ such that $K > 2$: then, there is a path from X_1 to X_K , and X_K is a descendent of X_1 and X_1 an ancestor of X_K .

14 The Markov condition is the assumption that every node is conditionally independent of the set of all its nondescendents, given the set of all its parents.

15 For more details on DAGs see Neapolitan (2003, chapters 1–2). In a DAG, causal relationships are represented by arrows between the propositional variables. Obviously, in the case of the Bayesian networks which will be presented, I am not considering the relationships between the variables representing the laws in the GNS as causal.

Figure 1: The Bayesian Network representing the situation before T_P is reduced to T_F .

Furthermore, three different pieces of evidence supporting the theories in the reduction relation are gained from experimental tests. They are defined with the propositional variables E , E_F , and E_P by DFH (Dizadji-Bahmani, Frigg, and Hartmann 2011, p. 324). The same simplifications introduced for T_F and T_P are applied to $\mathcal{E}$, $\mathcal{E}_F$, and $\mathcal{E}_P$. These three respectively support both theories, only the fundamental theory, and only the reduced theory.

- Evidence supporting only TD, e.g. the Joule–Thomson process.¹⁶
- Evidence supporting only SM, e.g. the dependence of a metal’s electrical conductivity on temperature.¹⁷
- Evidence supporting TD and SM simultaneously, e.g. the second law in TD.¹⁸

According to DFH, the situation before the reduction would then look like the network in Figure 1. Let P_1 be the probability distribution over the variables in such network. The relevant probabilities specifying the network are:

$$\begin{aligned}
 P_1(T_F) &= t_F, & P_1(T_P) &= t_P, \\
 P_1(E_F | T_F) &= p_F, & P_1(E_F | \neg T_F) &= q_F, \\
 P_1(E_P | T_P) &= p_P, & P_1(E_P | \neg T_P) &= q_P, \\
 P_1(E | T_F, T_P) &= \alpha, & P_1(E | T_F, \neg T_P) &= \beta, \\
 P_1(E | \neg T_F, T_P) &= \gamma, & P_1(E | \neg T_F, \neg T_P) &= \delta.
 \end{aligned} \tag{7}$$

Before the reduction occurs, T_F and T_P are probabilistically independent because they do not share the same vocabulary and they are not supported by the same evidence. In fact, E_F is independent of T_P given T_F and, vice versa, E_P is independent of T_F given T_P . Formally:

16 Say that we slow push a gas from a chamber with higher pressure into a chamber with lower pressure so that no heat is exchanged and pressure remains constant in both chamber. We can calculate the amount of cooling of the gas in the second chamber by using the principles of TD. This calculation coincides with the experimental data.

17 SM can help deriving equations relating the change in the conductivity of certain metals to the prior change in temperature. This involves quantum theory. The relation coincides with the experimental values as well.

18 If a wall dividing a box into two chambers is removed, the gas which was confined to one of the two chambers will spread evenly in the box. SM and TD respectively calculate that the Boltzmann entropy and the TD entropy of the gas increase.

Figure 2: The Bayesian Network representing the situation after the reduction of T_P to T_F , according to DFH.

$$E_F \perp\!\!\!\perp T_P \mid T_F, E_P \perp\!\!\!\perp T_F \mid T_P. \quad (8)$$

The independences in (8) hold because, in the aforementioned Bayesian network, the paths $E_F - T_F - E - T_P$ and $E_P - T_P - E - T_F$ are respectively blocked at T_F and T_P by $\{T_F\}$ and $\{T_P\}$. So, E_F and T_P are *d-separated*,¹⁹ and E_P and T_F as well. Therefore, the conjunction of the prior probabilities of the root nodes T_F and T_P looks like the following:

$$P_1(T_F \wedge T_P) = P_1(T_F)P_1(T_P) = t_F t_P. \quad (9)$$

Already before the reduction, one notices that in the Bayesian network there is a connection between T_P and T_F , namely evidence E . Such link has lead scientists to investigate the intimate relation among those two theories. DFH present then the network for the situation after the reduction (see Figure 2). To reduce one theory to the other, DFH complete three steps: derive T_F^* from T_F together with some auxiliary assumptions; introduce bridge laws which, together with T_F^* , yield T_P^* ; show that T_P^* is strongly analogous to together with T_P . They make important remarks which help defining the values of the conditional probabilities of the three nodes which follow from the only remaining root node T_F . The derivation of T_F^* from T_F and the interpretation of strong analogy between T_P^* and T_P depend on the judgment of the scientists and on the specific context in which the reduction occurs (Dizadji-Bahmani, Frigg, and Hartmann 2011, p. 328). Regarding bridge laws, which are not factual claims in a rigorous sense, T_P^* is a logical consequence of T_F^* , according to DFH. Let then P_2 be the probability distribution over the propositions one has after reducing T_P to T_F . The same simplifications introduced for T_F and T_P

¹⁹ In Bayesian networks, *d-separation* is a property for assessing whether or not a set A of variables is independent of another set B , given a third set C . Formally, $A \perp\!\!\!\perp B \mid C$ if and only if A and B are *d-separated* by $\{C\}$. In other words, in a network, $\{C\}$ blocks all paths between A and B if and only if A and B are *d-separated* by $\{C\}$. For more details, see Neapolitan (2003, chapter 2).

are applied to T_F^* and T_P^* . The relevant probabilities specifying the second network in Figure 2 are:

$$\begin{aligned} P_2(T_F^* | T_F) &= p_F^*, & P_2(T_F^* | \neg T_F) &= q_F^*, \\ P_2(T_P^* | T_F^*) &= 1, & P_2(T_P^* | \neg T_F^*) &= 0, \\ P_2(T_P | T_P^*) &= p_P^*, & P_2(T_P | \neg T_P^*) &= q_P^*. \end{aligned} \quad (10)$$

Following such network, DFH show that, after the reduction, evidence confirming one theory confirms the other and vice versa. In fact:

Theorem 2.1 E_F confirms T_P iff i) $(p_F - q_F) > 0$ and ii) $(p_F^* - q_F^*)(p_P^* - q_P^*) > 0$.

Theorem 2.2 E_P confirms T_F iff i) $(p_P - q_P) > 0$ and ii) $(p_F^* - q_F^*)(p_P^* - q_P^*) > 0$.

The two theorems maintain that in order to have a confirmation flow from E_F to T_P and from E_P to T_F : i) E_F should confirm T_F and E_P should confirm T_P ; ii) T_F should confirm T_F^* and T_P^* should confirm T_P . The two conditions are satisfied because:

- i. one of the original assumptions of the network is the fact that two evidence support their own respective theory;
- ii. T_F likely confirms T_F^* and T_P^* likely confirms T_P , because T_F^* was derived by T_F , and T_P^* is strongly analogous to T_P . Again, such construction of the GNS model is justified by the example in 2.1.

Once one constructs a Bayesian network like Figure 2 and assumes the presence of a confirmatory flow from T_F to T_P via T_F^* and T_P^* , she can prove that, after reducing T_P to T_F , E_F confirms T_P , and E_P confirms T_F .

2.3 The Bayesian Analysis of Tešić

The Bayesian analysis offered by DFH helps representing the reductive relation among T_P and T_F . However, Tešić (2019) points out to two difficulties faced by their analysis. I omit the details of the less relevant difficulty,²⁰ because it does not undercut the main project laid out by DFH.²¹ I focus, instead, on the main critique

20 I am referring to the difficulty faced by DFH in representing the independence between T_P and T_F as conditional.

21 In short, Tešić agree with DFH on the account of (8) prior to the reduction. However, he notices that (8) does not hold after the reduction. I am not focusing on the details here because I will not be discussing the relation between evidence and coherence in this article. Thus, what matters to the goal of my essay is that, ultimately, T_F and T_P are unconditionally independent before the reduction. For a better understanding of this issue, see Neapolitan (2003, p. 1108).

Tešić presents, which regards the value of the probabilities $P_2(T_P^* | T_F^*)$ and $P_2(T_P^* | \neg T_F^*)$. As showed in (10), while $P_2(T_P^* | T_F^*) = 1$, $P_2(T_P^* | \neg T_F^*) = 0$. Tešić claims that thinking the propositional variables T_P^* and T_F^* as interchangeable with each other (Dizadji-Bahmani, Frigg, and Hartmann 2011, pp. 329–330) without explicitly stating that bridge laws are assumed is misleading. In fact, DFH's representation of the bridge law in their Bayesian network suffers from three problems (Tešić 2019, pp. 1108–1111):

- i. Recall eq. (4), that is, the bridge law in the example of the Boyle–Charles law. According to the Bayesian network in Figure 2, one should then have not only $P_2(T_P^* | T_F^*) = 1$, but also $P_2(T_P | T_F^*) = 1$, because $pV = \frac{2n}{3} \langle E_{\text{kin}} \rangle$. Instead of the incorrect entailment $pV = \frac{2n}{3} \langle E_{\text{kin}} \rangle \mid = pV = kT$, Tešić notices that such entailment is only possible by supposing the bridge law B :

$$pV = \frac{2n}{3} \langle E_{\text{kin}} \rangle \mid =_B pV = kT. \quad (11)$$

The problem arising in Figure 2 is that it does not incorporate B in the background, according to Tešić. B therefore needs to be included in the probability function P_2 for the entailment to hold.

- i. The fact that from eq. (4) $P_2(T_P^* | T_F^*) = P_2(T_P^* | T_F^*) = 1$ and $P_2(T_P^* | \neg T_F^*) = P_2(T_P^* | \neg T_F^*) = 0$ follow (Tešić 2019, p. 1118) makes the reduction symmetric. The Bayesian network in Figure 2 seems to imply that the Boyle–Charles law is reduced to the kinetic theory of gases, and vice versa. This clearly goes against the main idea behind every kind of scientific reduction: reduction is anti-symmetric. Furthermore, the interchangeability between T_P^* and T_F^* would prevent partial reductions, which are still important in science, according to DFH (Dizadji-Bahmani, Frigg, and Hartmann 2010, p. 399). In fact, scientists are not always able to connect every term of $\mathbf{T}_P^*$ to $\mathbf{T}_F$ and deduce every law of $\mathbf{T}_P^*$ from $\mathbf{T}_F$ plus bridge laws.
- ii. From eq. (4) it also follows that the marginal probabilities $P_2(T_P)$, $P_2(T_P^*)$ and $P_2(T_F^*)$ are equal (Tešić 2019, p. 1118). It is hard to conceive of them as equal given that the equation $pV = \frac{2n}{3} \langle E_{\text{kin}} \rangle$ is deduced from the kinetic theory of gases together with some auxiliary assumptions, and that the equation $pV = kT$ is then deduce from the equation $pV = \frac{2n}{3} \langle E_{\text{kin}} \rangle$ and the bridge laws. This suggests that $P_2(T_P)$ and $P_2(T_F^*)$ should be left open.

Because of these three problems, Tešić presents an alternative Bayesian network to Figure 2. The main idea behind his network is to explicitly include the propositional variable B representing the bridge law as a root node. Let P_3 be a

Figure 3: The Bayesian Network representing the situation after reducing T_P to T_F , according to Tešić.

probability distribution over the variables in Figure 3. The same simplification applied to T_F and T_P applies to B . Thus, assume that the only element of B is B , and that the two values assignable to the propositional variable B are: B and $\neg B$. Then:

$$P_3(B) = b. \quad (12)$$

Two reasons motivate this explicit specification of the bridge laws (Tešić 2019, p. 1121): a) different scientists (often) have different credence about a particular bridge law (e.g. scientists share in fact different degree of belief on eq. (4)); b) the flow of confirmation depends on the value assigned to the probability of the bridge law. Thus, Tešić assigns the following values to the probabilities $P_3(T_P^* | T_F^*, B)$, $P_3(T_P^* | T_F^*, \neg B)$, $P_3(T_P^* | \neg T_F^*, B)$, and $P_3(T_P^* | \neg T_F^*, \neg B)$:

$$\begin{aligned} P_3(T_P^* | T_F^*, B) &= 1 \\ P_3(T_P^* | T_F^*, \neg B) &= P_3(T_P^* | \neg T_F^*, B) = P_3(T_P^* | \neg T_F^*, \neg B) = a, \end{aligned} \quad (13)$$

where $a \in (0, 1)$. Accordingly, the new probability assignments do not face the three problems noticed by Tešić. In fact, the first problem is avoided because now one has $P_3(T_P^* | T_F^*, B) = 1$ and, thus, $T_F^*, B | = T_P^*$, instead of simply having $T_F^* | = T_P^*$. The second problem is evaded by showing that the values in (13) entail $0 < P_3(T_F^* | T_P^*, B) < 1$: this means the reduction represented by Tešić's network is not symmetric. Finally, the third problem is successfully addressed because it is proved that the prior probabilities $P_3(T_P^*)$ and $P_3(T_F^*)$ can be either different or equal: in fact, it depends on the particular values one assigns to the relevant probabilities. According to this analysis, Theorems 2.1 and 2.2, which have already appeared in DFH's network, follow.

For all the reasons mentioned in this section, I will compare the coherence pre- and post-reduction between the reducing theory and the reduced one by following the probability assignments specified in the Bayesian network in Figure 3.

3 Coherence Measures

Coherence measures are probabilistic measures of the degree of coherence of information sets. They are real-valued functions and the value they assign to each set of propositions represents the degree of coherence of such a set. Coherence might not have a fixed meaning and quantifying coherence suffers from this presumed vagueness behind its notion. Because of this, there is not a single coherence measure: different coherence measures try to grasp several conceptions one might have of coherence. Regarding GNS, it might seem obvious that the reduction of T_P to T_F establishes some sort of coherence between the two theories (Sarkar 2015, p. 47) because of the way T_P is logically derived from T_F and the bridge law. In particular, from the perspective of a confirmation-laden coherence measure such as the Shogenji-Schupbach's measure, the condition which may seem to confirm an improved agreement between T_P and T_F after the reduction of one to other is the positive confirmation flow that goes from T_F to T_P via the bridge laws and the auxiliary assumptions. The measure, in fact, treats the coherence of an information set as the mutual support of the propositions in it, which is the view that coherence corresponds to the probabilistic dependence between propositions in a set. As opposed to this view, the Olsson's measure intends coherence as the relative overlap amongst those propositions. The higher their overlap over their conjunction the higher the degree of coherence and the higher the agreement of the propositions. These two coherence measures are the main ones in the literature and each of them corresponds to two different properties, i.e. dependence and agreement. Dependence and agreement cannot however be fulfilled at the same time.²²

In Nagel's words, reduction makes sure that two theories with largely overlapping domains are mutually consistent when they describe the same event. It appears that the properties aforementioned are in this context desirable. Therefore, what one would need from these coherence measures are stable results to the extent that they will all give similar results. Furthermore, while one would expect no coherence at all (or, a very low degree of coherence) prior to the reduction, after the reduction T_F should cohere with T_P .²³ In fact, obtaining the same result through different means count as a valid way to further support Nagel's view on coherence

²² See Koscholke, Schippers, and Stegmann (2018) for more details.

²³ Due to the limited breadth of this article, I am not taking into account the total evidence in the comparison between pre- and post-reduction joint probabilities. As it will be shown below, I am limiting myself to compare the degree of coherence of two information sets containing T_P and T_P . Thus, I consider the prior probability of the propositional variables T_F , T_F^* , T_P^* , T_P . If one wants to focus on assessing the degree of coherence between a theory and the total evidence supporting it,

in scientific reductions. Here, it is important to remark that I am interested in the notion of relative coherence rather than an absolute one; what is relevant is to check that the set containing the theories prior to the reduction is less coherent than the set of theories after the reduction occurs. The three coherence measures (Bovens and Hartmann 2003; Olsson 1999; Schubach 2011) which I will now present, might yield different results in certain contexts:²⁴ hopefully, in the case of the Bayesian network representing the GNS, they will not.

3.1 Schubach's Measure

To introduce Schubach's measure, consider a finite and non-empty information set $\mathcal{S}$, that is a set of ordered pairs:

$$\{\langle R_1, A_1 \rangle \wedge \dots \wedge \langle R_n, A_n \rangle\}, \quad (14)$$

where R_i is a source reporting that A_i , the content of the report R_i , is true. The content of the information system $\mathcal{S}$ is the ordered set of report contents $\langle A_1 \wedge \dots \wedge A_n \rangle$. Let P be the probability distribution over the propositions $\langle A_1 \wedge \dots \wedge A_n \rangle$, where $P(A_i)$ gives the degree of confidence of a rational agent in A_i . A coherence measure is a function C which maps every finite and non-empty set $\mathcal{S}$ of propositions with positive probability to a single real-valued outcome $C(\mathcal{S})$. In other words, the degree of coherence $C(\mathcal{S})$ consists precisely in the degree of coherence of its content $\langle A_1 \wedge \dots \wedge A_n \rangle$ (Bovens and Hartmann 2003). Then, let $\mathbf{S}$ be the set of all finite and non-empty information sets $\mathcal{S}$ of propositions with positive probability. Shogenji (1999) proposes to define the coherence of a set $\mathcal{S} \in \mathbf{S}$ as:

she should discuss coherence measures which resemble confirmation measure. See Bovens and Hartmann (2003, pp. 626–628) and Meijs (2005, chapter 5) for more details.

24 In epistemology, the coherence theory of justification, or coherentism, regards a belief in a set of beliefs or the set of beliefs itself as justified if the former coheres with the belonging set and the latter forms a coherent system. Usually, coherence is used as an epistemic criterion for theory choice even though coherentism does not assess the truth of a proposition or a belief in terms of its coherence. Rather, Moretti (2007) have found out that several coherence measures are confirmation conducive, which means that if evidence confirms a hypothesis, confirmation is transmitted to any hypotheses that are sufficiently coherent with the former hypothesis (Dietrich and Moretti 2005). Furthermore, following (Moretti and Akiba 2007), coherence measures violate intuitive epistemic principles we would want them to respect. This leads every coherence measure producing counter-intuitive results, e.g. the degree of coherence changes by adding or removing logical consequences to the information set, according to every coherence measure. See Olsson (2017) for more details.

$$\mathfrak{J}(S) \doteq \frac{P(A_1 \wedge \dots \wedge A_n)}{\prod_{i=1}^n P(A_i)}. \quad (15)$$

Shogenji conceives of the coherence of S as a mutual support between the propositions in S . The measure is sensitive to the number n of sources in cases of logically equivalent propositions. Given that the joint probability of the propositions of S equals to 1 when all sources reports the same proposition, as n , the number of agreeing reports, tends to infinity, so does the degree of coherence: indeed, "the more coherent beliefs are, the more likely they are together" (Shogenji 1999, p. 338). Instead, if the propositions of a set are independent, the variables representing them will be probabilistically independent and the coherence measure will equal to 1: in this case, the set S is neither coherent nor incoherent. Finally, if the joint probability of the propositions is lower than the probability of their products, the set will be incoherent. A series of criticisms²⁵ has been offered to the way this measure tries to capture the meaning of coherence (Schupbach 2011, pp. 3–5). For this reason, Schupbach suggests to generalise Shogenji's measure by calculating the degree of coherence as a weighted average of the degree of coherence of all subsets of S :

$$\mathfrak{S}(S) \doteq \text{Log} \left[\frac{P(A_1 \wedge \dots \wedge A_n)}{\prod_{i=1}^n P(A_i)} \right]. \quad (16)$$

Schupbach then proposes several definitions which are needed to avoid the problems arising with Shogenji's original measure and, therefore, to measure the degree of coherence in a more precise manner. As I will show in the next section, I am only interested in information sets with cardinality two. If sets have cardinality two, the *subset-sensitive* generalisation of Shogenji's measure offered by Schupbach can be simply represented by the Log-normalised version of the Shogenji's measure (Schupbach 2011, p. 9), namely eq. (16).²⁶ In fact, sets with cardinality two avoid all the problems faced by sets with cardinality strictly larger than two. So, if $n = 2$ and $S = \{A_1, A_2\}$, then:

$$\mathfrak{S}(S) = \text{Log} \left[\frac{P(A_1 \wedge A_2)}{P(A_1) \times P(A_2)} \right]. \quad (17)$$

If one measures the degree of coherence of a set containing probabilistically independent propositions as members, 0 will be the outcome since the joint probability of probabilistically independent variables is their product and therefore the logarithm of the ratio of two same values is 0. The outcome of $\mathfrak{S}(S)$ will tend to

²⁵ E.g. adding an irrelevant proposition to an information set (coherent or not) necessarily does not change the degree of coherence of the set, according to Shogenji's measure.

²⁶ For this reason, I will not show the intricate means Schupbach's measure encompasses to avoid Shogenji's counter-intuitive results.

negative infinity if the set $\mathcal{S}$ is not coherent. Vice versa, it will tend towards positive infinity if $\mathcal{S}$ is coherent.

3.2 Olsson's Measure

Olsson (1999) understands coherence as a total agreement between the propositions in an information set $\mathcal{S} \in \mathbf{S}$. In the case of a set $\mathcal{S} = \{A_1, A_2\}$, Olsson then proposes the following coherence measure:

$$\mathfrak{O}(\mathcal{S}) = \frac{P(A_1 \wedge A_2)}{P(A_1 \vee A_2)}, \quad (18)$$

which would look like the following for a set $\mathcal{S} = \{A_1, \dots, A_n\}$:²⁷

$$\mathfrak{O}(\mathcal{S}) = \frac{P(A_1 \wedge \dots \wedge A_n)}{P(A_1 \vee \dots \vee A_n)} \quad (19)$$

This measure ranges over the closed interval $[0, 1]$. $\mathfrak{O}$ indeed assigns 0 to cases of minimal agreement, that is, when the propositions involved are logically inconsistent. Regarding eq. (18), A_1 and A_2 would therefore not overlap in cases of minimal agreement. Whilst, in cases of maximal agreement between the propositions reported, the measure will give 1 as outcome. This means that the propositions in the set are logically equivalent. This measure also presents counter-intuitive results (Dietrich and Moretti 2005, p. 407), such as the possibility of reporting the degree of coherence of two positively dependent propositions as lower than that of two negatively dependent propositions. This remark about $\mathfrak{O}$ is not relevant for the aim of this article: in next section, I will compare the degree of coherence before and after reduction by exploiting (and assuming), respectively, the independence and the positive dependence amongst the two theories.

3.3 Bovens and Hartmann's Measure

Bovens and Hartmann (2003)²⁸ change the course of previous coherence measures. Instead of trying to make the intuition one can have on the notion of coherence precise (i.e. mutual support, total agreement, relative overlap), they focus on the

²⁷ Meijs generalises Olsson's measure into a coherence measure viewed as relative overlap. See Meijs (2005, ch. 3) for more details. Because such measure, regarding this specific context, meets the same conditions of the boosting confidence measure elaborated by Bovens and Hartmann (which will be outlined in the next subsection), I will keep using Olsson's measure.

²⁸ Their position is known as *weak Bayesian coherentism*. See Huemer (2007) for a broader discussion on Bayesianism and coherentist theories of epistemic justification.

role that coherence, as a property of information sets, plays: boosting our confidence that the content of an information set is true *ceteris paribus* once the information is received from independent and partially reliable source. The model Bovens and Hartmann construct aims to measure the degree of confidence in the joint truth of an information set. Such degree is determined by the combination of conditions such as results' expectation, tests' reliability, coherence of the information. Each of these conditions has a specific measure. Respectively, one has i) an expectance measure, which is about the degree of prior expectance of the joint truth of an information set, ii) a reliability measure, which computes the degree of reliability of the sources, and iii) the coherence measure, which is needed to assess to degree of coherence of the set.

Consider again n partially reliable sources i which report proposition A_i , for $i = 1, \dots, n$, so that the information set is $\{A_1, \dots, A_n\}$. The propositional variable A_i is defined for the proposition A_i and it can take on two values: A_i and $\neg A_i$. Similarly, the propositional variable R_i can also take on two values: R_i and $\neg R_i$. If the report mentions that A_i is the case after consulting the proper source, then R_i ; otherwise, $\neg R_i$. Then, let P be a probability distribution over the variables $A_1, \dots, A_n, R_1, \dots, R_n$, which satisfy the constraints of having independent and partially reliable sources. One needs sources to be:

- i. *independent*, or else, sources would present information either by looking at other reports and bring additional information or by conveying what they think a coherent report is. For coherence to play a role as confidence boosting, the sources should rather gather information through and only through their own observations which they will report without biasedly inferring what they think a coherent report would look like;
- ii. *partially reliable*, or else, sources would either be truth-tellers or randomisers. The latter would report the information needed in entirely random manner and assessing the degree of coherence of information reported without any degree of reliability is useless. The former would report only true information making the property of coherence redundant, because it would no longer be important whether the report is coherent or not as it is already true given that its information comes from a fully reliable source.

These two points can be formally translated.

- i. Having independent sources means that R_i should only report that A_i is the case given the she has (likely) observed A_i without her observations being affected by additional facts. Probabilistically speaking, A_i *screens off* R_i from all other variables A_j and R_j . Thus, there is a conditional independence between R_i and $A_1, R_1, \dots, A_{i-1}, R_{i-1}, A_{i+1}, R_{i+1}, \dots, A_n, R_n$, given A_i , for $i = 1, \dots, n$:

$$R_i \perp\!\!\!\perp A_1, R_1, \dots, A_{i-1}, R_{i-1}, A_{i+1}, R_{i+1}, \dots, A_n, R_n | A_i. \quad (20)$$

- ii Partial reliability can be specified with two parameters, namely the true positive rate $P(R_i | A_i) = p$ and the negative positive rate $P(R_i | \neg A_i) = q$. Bovens and Hartmann then assume that all sources are equally reliable, that is, all sources have the same p and the same q . The assumptions is introduced because knowing how much one trusts a source is not relevant to assessing the degree of coherence of an information set. For the reasons stated above, all sources in this model are deemed epistemically imperfect, which means that they are more reliable than randomisers, but less reliable than truth-tellers. Thus, the following constraint is imposed on P :

$$p > q > 0. \quad (21)$$

Bovens and Hartmann then define the degree of confidence in the information set as equal to the posterior joint probability of the propositions in the set after all reports have been collected:

$$P^*(A_1 \wedge \dots \wedge A_n) = P(A_1 \wedge \dots \wedge A_n | R_1 \wedge \dots \wedge R_n). \quad (22)$$

Then, they apply the Bayes rule on eq. (22) and simplify it with respect to the independence constraint (eq. (20)):

$$P^*(A_1 \wedge \dots \wedge A_n) = \frac{P(R_1 | A_1) \times \dots \times P(R_n | A_n) \times P(A_1 \wedge \dots \wedge A_n)}{\sum_{A_1, \dots, A_n} P(R_1 | A_1) \times \dots \times P(R_n | A_n) \times P(A_1 \wedge \dots \wedge A_n)}. \quad (23)$$

While the numerator can be seen as:

$$P(R_1 | A_1) \times \dots \times P(R_n | A_n) \times P(A_1 \wedge \dots \wedge A_n) = p^n \xi_0, \quad (24)$$

where p^n is the positive rate $P(R_i | A_i) = p$ to the n th degree, and ξ_0 is the prior probability of the conjunction of $n - 0$ positive values and 0 negative values of the variables $A_1, \dots, A_n$. The denominator looks like:

$$\sum_{A_1, \dots, A_n} P(R_1 | A_1) \times \dots \times P(R_n | A_n) \times P(A_1 \wedge \dots \wedge A_n) = p^n \xi_0 + qp^{n-1} \xi_1 + \dots + q^n \xi_n. \quad (25)$$

In the denominator, Bovens and Hartmann gather all terms in which the variables $A_1, \dots, A_n$ take, first, n positive values and 0 negative values, then $n - 1$ positive values and 1 negative value and so on, until the term in which those variables take 0 positive values and n negative values is reached. This means that, for instance, ξ_1 is a prior probability where one proposition is false. Finally, if both numerator and denominator are divided by p^n , the posterior probability $P^*(A_1 \wedge \dots \wedge A_n)$ would be:

$$P^*(A_1 \wedge \dots \wedge A_n) = \frac{\xi_0}{\sum_{i=0}^n \xi_i \chi^i}, \quad (26)$$

where $x = \frac{q}{p}$, that is, the likelihood ratio, and $\sum_{i=0}^n \xi_i = 1$. ξ_i is therefore the conjunction of all the joint probabilities of the instances of $n - i$ positive values and i negative values of the variables $A_1, \dots, A_n$.

According to the information collected so far, the three measures of Bovens and Hartmann can be finally presented. The *expectance measure* is defined by the prior joint probability of the propositions in the information set, i.e. the probability before any report was received:

$$\xi_0 = P(A_1 \wedge \dots \wedge A_n), \quad (27)$$

The more ξ_0 increases, the more the degree of coherence of the set increases. The degree of coherence of the information set, i.e. eq. (22), is a monotonically increasing function of r , that is, the *reliability measure*:

$$r \doteq 1 - x, \quad (28)$$

where x is the likelihood ratio. This measure ranges over the open interval $(0, 1)$, because the sources are neither fully reliable ($r \neq 1$), nor entirely unreliable ($r \neq 0$). The last relevant measure is the *coherence measure*. In order to evaluate the coherence of an information set, they measure the proportion of the confidence boost b , defined by the ratio $\frac{P^*(A_1 \wedge \dots \wedge A_n)}{P(A_1 \wedge \dots \wedge A_n)}$, relative to the confidence boost $b^{\max}$: a confidence boost which would have been received if the same information had been received in the form of maximally coherent information. A maximally coherent information set would contain only logically equivalent propositions and it has a specific distribution of ξ :

$$\langle \xi_0, 0, \dots, 0, 1 - \xi_0 \rangle. \quad (29)$$

After calculating the posterior joint probability of a maximally coherent information set and its confidence boost, Bovens and Hartmann compute the coherence measure of an information set $S = \{A_1, \dots, A_n\}$. This measure is functionally dependent on the expectance and the reliability measure (cf. Bovens and Hartmann 2003, p. 612):

$$\mathfrak{B}(S) = \frac{b(S)}{b^{\max}(S)} = \frac{\xi_0 + (1 - \xi_0)(1 - r)^n}{\sum_{i=0}^n \xi_i (1 - r)^i}. \quad (30)$$

According to Meijs (2005), the maximal requirement is what makes Bovens and Hartmann's measure produce counter-intuitive results, because $\mathfrak{B}$ may, in some cases, consider the degree of coherence of a set containing independent proposition as higher of the degree of coherence of a set whose members are positively dependent. This feature of $\mathfrak{B}$ can be threatening for the aim of the article: comparing the degree of coherence pre- and post-reduction among the two theories. Therefore, it will be necessary to check whether or not the results which come

from the comparisons of the degree of coherence pre- and post-reduction, and which are obtained through different coherence measures, are similar. In case the results will contrast with each other, further remarks regarding the choice of the proper coherence measure will have to be made. It would, in fact, become a contextual question of which coherence measure one should prioritise with respect to the epistemic conception of coherence (e.g. mutual support, boosting confidence) they represent.

4 Comparing the Degree of Coherence

In this section, I will compare the degree of coherence of the two theories before and after the reduction. To calculate the degree of coherence with the measures aforementioned, consider the information set $\mathcal{S} = \{T_F, T_P\}$ containing the two theories before the reduction, and the information set $\mathcal{S}' = \{T_F, T_P\}$ containing the two theories after the reduction. Then, let $\succeq$ be a quasi-ordering relation over the set $\mathbf{S} = \{\mathcal{S}, \mathcal{S}'\}$, which denotes the binary relation *at least as coherent as*, such that if $\mathcal{S}' \succeq \mathcal{S}$, then $\mathcal{S}'$ will be *at least as coherent as* $\mathcal{S}$. To compare $\mathcal{S}'$ and $\mathcal{S}$ is important to use the assumption Bovens and Hartmann makes for the sources: they need to be partially and equally reliable, and independent. So, I am implying that, for example, a scientist has the same credence for $P_1(E_F|T_F)$, $P_3(E_F|T_F)$, $P_1(E_P|T_P)$, $P_3(E_P|T_P)$, and so on.²⁹ Having such an assumption regarding the evidence is important as neither Schupbach nor Olsson consider the role coherence plays for the evidence supporting theories.

4.1 Schupbach's Measure

Recall Schupbach's coherence measure. Before the reduction, T_P and T_P are probabilistically independent. Hence, according to Schupbach, the value of coherence will be 0:

$$\mathfrak{S}(\mathcal{S}) = \text{Log} \left[\frac{P_1(T_F \wedge T_P)}{P_1(T_F) \times P_1(T_P)} \right] = \text{Log} \left[\frac{P_1(T_F) \times P_1(T_P)}{P_1(T_F) \times P_1(T_P)} \right] = 0. \quad (31)$$

²⁹ I am calculating the degree of coherence by following the assumptions Bovens and Hartmann make regarding the importance of having independent and equally (and partially) reliable sources. In practice, this means that I am assuming the three types of evidence for TD and SM have the same marginal probability and have been independently tested.

The situation changes after T_P is reduced to T_P as the two become probabilistically dependent:

$$\mathfrak{S}(\mathcal{S}') = \text{Log} \left[\frac{P_3(T_F \wedge T_P)}{P_3(T_F) \times P_3(T_P)} \right]. \quad (32)$$

Therefore, according to Schubach, the two theories will be consistent with each other after the reduction if and only if what is inside the logarithm is strictly higher than 1, that is, if and only if the joint probability of T_F and T_P is higher than the product of their prior probabilities. Thus, one can have the following theorem (see Appendix A):

Theorem 4.1 $\mathfrak{S}(\mathcal{S}') = \mathfrak{S}(\mathcal{S})$ iff $p_F^* = q_F^*$ or $p_P^* = q_P^*$. $\mathfrak{S}(\mathcal{S}') > \mathfrak{S}(\mathcal{S})$ iff $p_F^* > q_F^*$ and $p_P^* > q_P^*$.

The first part of theorem means that if T_P and T_P^* are independent or T_F and T_F^* are independent, therefore T_P and T_F remain independent after the reduction and their coherence does not improve. The second part of the theorem, instead, means that coherence between T_P and T_F is gained after the reduction if and only if: i) the probability of the conjunction of T_P and T_P^* is higher than the conditional probability of T_P on $\neg T_P^*$; and ii) the probability of the conjunction of T_F and T_F^* is higher than the conditional probability of T_F on $\neg T_F^*$.

4.2 Olsson's Measure

The same theorem will appear with the comparison made with Olsson's measure. Before T_P is reduced to T_F , one have to maintain the independence between their variables. Thus, the coherence measure formulated by Olsson would look like the following:

$$\mathfrak{O}(\mathcal{S}) = \frac{P_1(T_F \wedge T_P)}{P_1(T_F \vee T_P)} = \frac{P_1(T_F) \times P_1(T_P)}{P_1(T_F) + P_1(T_P) - P_1(T_F) \times P_1(T_P)}. \quad (33)$$

Once the reduction happens, $\mathfrak{O}$ for $\mathcal{S}'$ is:

$$\mathfrak{O}(\mathcal{S}') = \frac{P_3(T_F \wedge T_P)}{P_3(T_F) + P_3(T_P) - P_3(T_F \wedge T_P)}. \quad (34)$$

Here, one would need to fix the prior probability $P_1(T_P)$, i.e. $P_1(T_P) = P_3(T_P)$, to meaningfully compare the two measures. Then, the sufficient and necessary condition for $\mathcal{S}' \succeq \mathcal{S}$ to hold is that the denominator of $\mathfrak{O}(\mathcal{S}')$ is less than or equal to the denominator of $\mathfrak{O}(\mathcal{S})$, because the numerator of $\mathfrak{O}(\mathcal{S}')$ is greater than the

numerator of $\mathfrak{D}(S)$.³⁰ Once the values are substituted in the coherence measure $\mathfrak{D}$,³¹ the following theorem is obtained:

Theorem 4.2 $\mathfrak{D}(S') = \mathfrak{D}(S)$ iff $p_F^* = q_F^*$ or $p_P^* = q_P^*$. $\mathfrak{D}(S') > \mathfrak{D}(S)$ iff $p_F^* > q_F^*$ and $p_P^* > q_P^*$.

This is the same condition one finds with Schupbach's measure. The higher the difference $P_3(T_F^* | T_F) - P_3(T_F^* | \neg T_F)$ and the difference $P_3(T_P | T_P^*) - P_3(T_P | \neg T_P^*)$, the higher the degree of coherence.

4.3 Bovens and Hartmann's Measure

To evaluate the degree of coherence pre- and post-reduction in light of the analysis of Bovens and Hartmann, the reliability measure they formulate should not play any role.³² In fact, $\mathfrak{B}$ construct a quasi-ordering relation over the set $\{S, S'\}$ and this binary relation is formally independent of the reliability measure. Reconsider the two sets S and S' which have the same size, i.e. 2. Recall that while P_1 is the joint probability distribution for the pre-reduction propositions T_F and T_P , P_3 is the joint probability distribution over the post-reduction propositions T_F and T_P . By calculating the weight vectors $\langle \xi_0, \xi_1 \rangle$ for P_1 and $\langle \xi'_0, \xi'_1 \rangle$ for P_3 , the following difference function can be constructed:

$$f_r(S', S) = \mathfrak{B}(S') - \mathfrak{B}(S). \quad (35)$$

The relation which induces a quasi-ordering over the set of the two information sets S and S' is then defined as:

$$\text{For } S, S', S' \succeq_S \text{ iff } f_r(S', S) \geq 0 \text{ for all values of } r \in (0, 1). \quad (36)$$

Finally, to determine whether $f_r \geq 0$, one needs to assess the conditions under which the sign of f_r is actually positive for all values of $r \in (0, 1)$. For this reason, Bovens and Hartmann calculate the following condition:

- i. $\xi_0 \leq \xi'_0 \wedge \xi_1 \geq \xi'_1$, or
- ii. $\xi_0 \geq \xi'_0 \wedge \xi_1 \leq \xi'_1$.

This condition is necessary and sufficient for $S' \succeq_S$ to hold. First, consider the first condition. $\xi_0 \leq \xi'_0$ resembles what has been showed above: $P_1(T_F) \times P_1(T_P) \leq P_3(T_F \wedge T_P)$. This is our usual condition, which reports:

³⁰ See Appendix A for details.

³¹ See Appendix B for details.

³² I have been using this assumption so far for Schupbach's and Olsson's measure.

$$\bar{a} b t_F \overline{t_F} (p_F^* - q_F^*) (p_P^* - q_P^*) > 0. \quad (37)$$

Then, $\xi_1 \geq \xi'_1$ means that:

$$P_1(T_F \wedge \neg T_P) + P_1(\neg T_F \wedge T_P) \geq P_3(T_F \wedge \neg T_P) + P_3(\neg T_F \wedge T_P). \quad (38)$$

Once $P_1(T_P)$ is fixed and made equal to $P_3(T_P)$ (for details, see Appendix C), a similar theorem to the ones seen above follows:

Theorem 4.3 $\mathfrak{B}(\mathcal{S}') = \mathfrak{B}(\mathcal{S})$ iff $p_F^* = q_F^*$ or $p_P^* = q_P^*$. $\mathfrak{B}(\mathcal{S}') > \mathfrak{B}(\mathcal{S})$ iff $p_F^* > q_F^*$ and $p_P^* > q_P^*$.

These are the same conditions one has with Olsson and Schupbach's measures. The second part of the condition mentioned by Bovens and Hartmann is ruled out because it has been showed earlier that the probability of the conjunction of the two propositions T_F and T_P is higher after the reduction rather than before. Thus, the second part of Bovens and Hartmann's condition does not hold.

5 Coherence and Confirmation

The conditions upon which the coherence measures report that $\mathcal{S}'$ is more coherent than $\mathcal{S}$ are the same reported by the confirmation measure used by DFH (Dizadji-Bahmani, Frigg, and Hartmann 2011) which say that evidence confirming T_P also confirms T_F and vice versa. This raises an important issue about the nature of coherence measures:³³ do the coherence measures employed here simply track the positive flow of information in the model of the intertheoretic reduction designed by DFH and Tešić? To answer this question, I now consider a couple of numerical examples for Schupbach's and Olsson's coherence measures. These examples are aimed at not only understanding the relationship between confirmation and coherence, but also the different shapes of the functions representing the two coherence measures. In these examples, I assume that there is a positive flow of confirmation from T_F to T_P , that is, $p_F^* > q_F^*$ and $p_P^* > q_P^*$.

The first two examples show the different degrees of coherence of $\mathcal{S}'$ according to Schupbach's and Olsson's coherence measure as a function of the amount of confirmation between T_F and T_P^* . The examples have p_F^* and q_F^* set as variables and t_F , a , b , p_P^* , and q_P^* as parameters. The value assigned to t_F , a , b and p_P^* is 0.5, while

³³ I would like to thank the second reviewer for raising this point.

Figure 4: The x-axis and the y-axis in the two pictures are respectively the amount of confirmation between T_F and T_F^* (i.e. $P_3(T_F^*|T_F) - P_3(T_F)$) and the degree of coherence of $\mathcal{S}'$ (in the first graph, there is $\mathfrak{Q}(\mathcal{S}')$, while in the second there is $\mathfrak{S}(\mathcal{S}')$). The variables p_F^* and q_F^* take any value that goes from 0.01 to 0.99. When the difference $p_F^* - q_F^*$ tends to 1, the area showed in the graphs converges towards a single value since only two values, i.e. $p_F^* = 1$ and $q_F^* = 0$, can give $p_F^* - q_F^* = 1$. The highlighted red lines are for $p_F^* = 0.5$ and they show the positive correlation between coherence and confirmation.

q_p^* has 0.1.³⁴ Interestingly, both graphs reveals the degree of coherence of the two theories after the reduction and the amount of confirmation between T_F and T_F^* are positively related (Figure 4).

The other two examples in Figure 5 show, instead, the different degrees of coherence of $\mathcal{S}'$ according to Schupbach's and Olsson's coherence measure as a function of the amount of confirmation between T_p^* and T_p . Here, I set p_p^* and q_p^* as variables and t_F , a , b , p_F^* , and q_F^* as parameters. Like in the previous example, the value assigned to t_F , a , b and p_F^* is 0.5, while q_F^* has 0.1. As opposed to the two graphs showed beforehand, the graphs display a negative correlation between the degree of coherence of the two theories after the reduction and the amount of confirmation between T_p^* and T_p : the more T_F and T_p are coherent, the less T_p^* confirms T_p .

Finally, the graphs in Figure 6 reveal the positive correlation between the prior probability of the bridge laws and the degree of coherence of $\mathcal{S}'$. As expected, a higher assignment of b contributes to a higher degree of coherence. Bridge laws, however, do not provide a major contribution for the coherence of the set containing the two theories after the reduction.

³⁴ The values assigned to these parameters can be potentially given by scientists and their disagreement about prior probabilities and likelihoods could be represented in another matter (e.g. with imprecise probabilities). Hence, here I assign arbitrary values simply to better understand the relation between confirmation and coherence.

Figure 5: The x-axis and the y-axis in the two pictures are respectively the amount of confirmation between T_p^* and T_p (i.e. $P_3(T_p|T_p^*) - P_3(T_p)$) and the degree of coherence of S' (in the first graph, there is $\mathfrak{D}(S')$, while in the second there is $\mathfrak{S}(S')$). The variables p_p^* and q_p^* take any value that goes from 0.01 to 0.99. Also in these examples, when the difference $p_p^* - q_p^*$ tends to 1, the area showed in the graphs converge towards a single value since only two values, i.e. $p_p^* = 1$ and $q_p^* = 0$, can give $p_p^* - q_p^* = 1$. The highlighted red lines are for p_p^* set as 0.5 and they show the negative correlation between coherence and confirmation.

Figure 6: The x-axis and the y-axis in the two pictures are respectively the different probability assignments for b and the degree of coherence of S' (in the first graph, there is $\mathfrak{D}(S')$, while in the second there is $\mathfrak{S}(S')$). The variable b take any value that goes from 0.01 to 0.99. Here, t_F , a , p_F^* , p_p^* , 1_F^* , and q_p^* are set as parameters. Like in the previous example, the value assigned to t_F , a , p_F^* , and p_p^* is 0.5, while q_F^* and q_p^* have 0.1.

6 Conclusions

In this article, I have showed that, according to three coherence measures and under the assumption elaborated by Bovens and Hartmann,³⁵ the degree of

³⁵ Sources and, thus, evidence is independent, and partially and equally reliable.

coherence amongst two self-consistent and well-confirmed theories with largely overlapping domains of application which are involved in a reduction relation *à la* GNS is higher than the degree of coherence of the same theories which, before one is reduced to another, sketch a contradictory view of the world. To do this, I first presented the classic putative example where the GNS can be applied to described the main relations occurring in a scientific reduction: the reduction of TD to SM. Then, I looked at two attempts of modelling the GNS in probabilistic terms. Specifically, DFH and Tešić offer two Bayesian analyses, which mostly disagree on the probabilistic account of bridge laws, a crucial relation in a GNS. In fact, Tešić undermines the view shared by DFH that the propositional variables T_P^* and T_F^* are interchangeable. This view faces three problems: i) it proposes an incorrect entailment amongst the laws derived from the fundamental theory through the bridge laws and the fundamental theory itself; ii) it makes the reduction symmetric; iii) it omits the role of auxiliary assumptions and boundary conditions in deriving laws strongly analogous to the ones involved in the reduction. The necessity of explicitly assuming the bridge law **B** in order to overcome these three challenges leads Tešić then to formulate a new probability distribution P_3 and therefore new probability assignments, such as $P_3(T_P^* | T_F^*, B) = 1$, $P_3(T_P^* | T_F^*, \neg B) = P_3(T_P^* | \neg T_F^*, B) = P_3(T_P^* | \neg T_F^*, \neg B) = a$, and $P_3(\neg T_P^* | T_F^*, B) = 0$. This new purported probabilistic representation of GNS has been used to assess the degree of coherence between the two theories in question as the three coherence measures proposed are expressed in probabilistic terms. I briefly mentioned which epistemic notion of coherence these measures try to grasp. Schupbach's measure understands coherence as mutual support amongst the propositions included in an information set. Olsson's measure conceives coherence as an agreement between the propositions involved. Bovens and Hartmann, instead, focus on the role that coherence, as a property of an information set, should play whenever one wants to assess whether or not the set is true. Coherence, according to this framework, boosts one's confidence that the set is true *ceteris paribus* once the information is received from independent and partially reliable source. Due to the limit breadth of the article, I did not explain in greater the counter-intuitive results these measures may offer in certain scenarios. Rather, I tried to heuristically show that, in the case of GNS constructed in function of Tešić's suggestions, they do not report different outcomes. In the fourth section, I showed that they actually outline stable results. The three theorems point out that a set containing T_F and T_P have a higher degree of coherence after is reduced to than before the reduction if and only if $P_3(T_F^* | T_F) > P_3(T_F^* | \neg T_F)$ and $P_3(T_P | T_P^*) > P_3(T_P | \neg T_P^*)$. These two conditions are likely met by the theories involved in the GNS. Thus, under the assumptions of

the Bayesian analysis provided by Tešić and the one provided by Bovens and Hartmann, the coherence measures seem to report that the GNS make the two theories involved cohere with each other in light of their positive dependency. Those two inequalities highlighted in the theorems have to hold in order to consider two theories involved in the reductive relation coherent. In the last section, I finally introduced six numerical examples aimed at better understanding the relation between coherence and confirmation as well as their respective measures with respect to the context of intertheoretic reduction as designed by DFH and Tešić. Interestingly, while bridge laws and the confirmation flow between T_F and T_F^* are positively related to degree of coherence of S' , the confirmation flow between T_P^* and T_P is not.

Further projects can still be proposed at the intersection of coherence and GNS. First of all, regarding Tešić's probability assignment of the bridge law and, in general, regarding every probability assignment in the Bayesian network (see Figure 3), it would be interesting to use tools borrowed from imprecise probability (IP) to better represent the credence scientists might have towards prior probabilities and likelihoods. In fact, usually scientists disagree about the value assignments of prior probabilities and likelihoods. IP is a generalisation of probability theory, which is applicable to cases where it is hard to identify a unique probability distribution, because evidence might be scarce, vague, or conflicting. Thereby, the goal of IP is represent the available knowledge more accurately, instead of focusing on a single precise outcome. Would IP be able to represent the situation before and after the reduction and assess whether or not the two post-reduction theories would cohere amongst each other? Second, more coherence and confirmation measures should check the results I showed above and further discussion should be focused on their relation. The three coherence measures might be extensionally equivalent with the respect to the GNS, because, even if their frameworks and their main epistemic notions of coherence differed, they all confirmed the same hypothesis and provided stable results. What would happen with other coherence measures? Would they report an increase or a decrease in coherence if considered as a function of the amount of confirmation? Third, the coherence relation between evidence and theories requires a proper investigation. One might want to start this investigation from the work of Meijs (2005) and then apply it to the case of GNS. Will the evidence of a theory be coherent with the other theory, and vice versa? Fourth and finally, the third question opens up to the diatribe between foundationalism and coherentism, and to the role coherence should play in assessing a scientific theory. I showed and worked with respect to the assumption made by Bovens and Hartmann. Some philosophers have actually

been skeptical of their suggestions. This means that not only new measures, but also other assumptions regarding the sources reporting crucial information (i.e. the evidence) might be employed to construct novel coherence measures. Would considerations on coherence still play a role in deciding to accept a scientific theory if one dropped the conception that comparing the degree of coherence of two (or more) information sets should assume that sources reporting the propositions contained in the sets are equally and partially reliable, and independent? These four questions highlight the fact that philosophers of science and formal epistemologists should work closely with natural scientists.

Appendices

In what follows, I shall refer to Neapolitan (2003) in order to compute the probabilities of the values of random variables in Bayesian networks.

A Theorem 4.1

$$P_3(T_F \wedge T_P) = P_3(T_F) \sum_{T_P^*, T_F^*, B} P_3(T_P | T_P^*) P_3(T_P^* | T_F^*, B) P_3(B) P_3(T_F^* | T_F) \quad (39)$$

$$= t_F \left[p_P^* (b p_F^* + a \bar{b} + a b \bar{p}_F^*) + q_P^* (\bar{a} \bar{b} + \bar{a} b \bar{p}_F^*) \right]$$

$$\begin{aligned} P_3(T_P) &= \sum_{T_P^*, T_F^*, B, T_F} P_3(T_P | T_P^*) P_3(T_P^* | T_F^*, B) P_3(B) P_3(T_F^* | T_F) P(T_F) \\ &= p_P^* (b p_F^* t_F + b q_F^* \bar{t}_F + a \bar{b} + a b \bar{p}_F^* t_F + a b \bar{q}_F^* \bar{t}_F) \\ &\quad + q_P^* (\bar{a} \bar{b} + \bar{a} b \bar{p}_F^* t_F + \bar{a} b \bar{q}_F^* \bar{t}_F) \end{aligned} \quad (40)$$

$$\begin{aligned} P_3(T_F \wedge T_P) &> P_3(T_F) \times P_3(T_P) \\ t_F &\left[p_P^* (b p_F^* + a \bar{b} + a b \bar{p}_F^*) + q_P^* (\bar{a} \bar{b} + \bar{a} b \bar{p}_F^*) \right] \\ &> t_F \left[p_P^* (b p_F^* t_F + b q_F^* \bar{t}_F + a \bar{b} + a b \bar{p}_F^* t_F + a b \bar{q}_F^* \bar{t}_F) \right. \\ &\quad \left. + q_P^* (\bar{a} \bar{b} + \bar{a} b \bar{p}_F^* t_F + \bar{a} b \bar{q}_F^* \bar{t}_F) \right] \end{aligned} \quad (41)$$

$$\bar{a} b t_F \bar{t}_F (p_F^* - q_F^*) (p_P^* - q_P^*) > 0$$

B Theorem 4.2

$$\begin{aligned}
 P_1(T_F \wedge T_P) &= t_F t_P \\
 &= t_F \left[p_P^* (b p_F^* t_F + b q_F^* \bar{t}_F + a \bar{b} + a b \bar{p}_F^* t_F + a b \bar{q}_F^* \bar{t}_F) \right. \\
 &\quad \left. + q_P^* (\bar{a} \bar{b} + \bar{a} b \bar{p}_F^* t_F + \bar{a} b \bar{q}_F^* \bar{t}_F) \right] \quad (42)
 \end{aligned}$$

$(P_3(T_P) \text{ instead of } P_1(T_P))$

$(P_3(T_F) \text{ instead of } P_1(T_F))$

$$\begin{aligned}
 P_3(T_F) + P_3(T_P) - P_3(T_F \wedge T_P) &\leq P_1(T_F) + P_1(T_P) - P_1(T_F \wedge T_P) \\
 - P_3(T_F \wedge T_P) &\leq -P_1(T_F \wedge T_P) \\
 P_3(T_F \wedge T_P) &\geq P_3(T_F) \times P_3(T_P) \\
 t_F \left[p_P^* (b p_F^* + a \bar{b} + a b \bar{p}_F^*) + q_P^* (\bar{a} \bar{b} + \bar{a} b \bar{p}_F^*) \right] &\quad (43) \\
 \geq t_F \left[p_P^* (b p_F^* t_F + b q_F^* \bar{t}_F + a \bar{b} + a b \bar{p}_F^* t_F + a b \bar{q}_F^* \bar{t}_F) \right. \\
 &\quad \left. + q_P^* (\bar{a} \bar{b} + \bar{a} b \bar{p}_F^* t_F + \bar{a} b \bar{q}_F^* \bar{t}_F) \right] \\
 \bar{a} b t_F \bar{t}_F (p_F^* - q_F^*) (p_P^* - q_P^*) &\geq 0
 \end{aligned}$$

C Theorem 4.3

Recall Bovens and Hartmann's condition. In order to assess ξ_1 and ξ'_1 , one needs the sum of the joint probabilities of one theory and the negation of the other for our two information sets before and after reduction. First, let me show ξ_1 , which the pre-reduction parameter:

$$\xi_1 = P_1(T_F \wedge \neg T_P) + P_1(\neg T_F \wedge T_P) = t_F(1 - t_P) + t_P \bar{t}_F, \quad (44)$$

which holds in virtue of the probabilistic independence of T_F and T_P . Second, I compute ξ'_1 :

$$\xi'_1 = P_3(T_F \wedge \neg T_P) + P_3(\neg T_F \wedge T_P) \quad (45)$$

Then, I substitute the following values in this equation.

$$\begin{aligned}
 P_3(T_F \wedge \neg T_P) &= P_3(T_F) \sum_{T_P^*, T_F^*, B} P_3(\neg T_P | T_P^*) P_3(T_P^* | T_F^*, B) P_3(B) P_3(T_F^* | T_F) \\
 &= t_F \left[\bar{p}_P^* (b p_F^* + a \bar{b} + a b \bar{p}_F^*) + \bar{q}_P^* (\bar{a} \bar{b} + \bar{a} b \bar{p}_F^*) \right] \quad (46)
 \end{aligned}$$

$$\begin{aligned}
 P_3(\neg T_F \wedge T_P) &= P_3(\neg T_F) \sum_{T_P^*, T_F^*, B} P_3(T_P | T_P^*) P_3(T_P^* | T_F^*, B) P_3(B) P_3(T_F^* | \neg T_F) \\
 &= \bar{t}_F \left[p_P^* (b q_F^* + a \bar{b} + a b \bar{q}_F^*) + q_P^* (\bar{a} \bar{b} + \bar{a} b \bar{q}_F^*) \right]
 \end{aligned} \tag{47}$$

$$\begin{aligned}
 P_3(T_F \wedge \neg T_P) + P_3(\neg T_F \wedge T_P) &\leq P_1(T_F \wedge \neg T_P) + P_1(\neg T_F \wedge T_P) \\
 P_3(T_F \wedge \neg T_P) + P_3(\neg T_F \wedge T_P) &\leq P_3(T_F) \times P_3(\neg T_P) + P_3(\neg T_F) \times P_3(T_P) \\
 &\quad (P_3(T_F) \text{ instead of } P_1(T_F)) \\
 &\quad (P_3(T_P) \text{ instead of } P_1(T_P)) \\
 \bar{a} b (p_F^* - q_F^*) (p_P^* - q_P^*) &\geq 0
 \end{aligned} \tag{48}$$

References

- Bovens, L., and S. Hartmann. 2003. "Solving the Riddle of Coherence." *Mind* 112: 10.
- Dietrich, F., and L. Moretti. 2005. "On Coherent Sets and the Transmission of Confirmation." *Philosophy of Science* 72 (3): 403–24.
- Dizadji-Bahmani, F., R. Frigg, and S. Hartmann. 2010. "Who's Afraid of Nagelian Reduction?" *Erkenntnis* 73 (3): 393–412.
- Dizadji-Bahmani, F., R. Frigg, and S. Hartmann. 2011. "Confirmation and Reduction: A Bayesian Account." *Synthese* 179 (2): 321–38.
- Feynman, R. P., R. B. Leighton, and M. Sands. 1964. *The Feynman Lectures on Physics*. Reading: Addison-Wesley.
- Huemer, M. 2007. "Weak Bayesian Coherentism." *Synthese* 157 (3): 337–46.
- Koscholke, J., M. Schippers, and A. Stegmann. 2018. "New Hope for Relative Overlap Measures of Coherence." *Mind* 128 (512): 1261–84.
- Meijs, W. 2005. "Probabilistic Measures of Coherence." PhD thesis. Erasmus University, Rotterdam.
- Moretti, L. 2007. "Ways in Which Coherence is Confirmation Conducive." *Synthese* 157 (3): 309–19.
- Moretti, L., and K. Akiba. 2007. "Probabilistic Measures of Coherence and the Problem of Belief Individuation." *Synthese* 154: 73–95.
- Nagel, E. 1961. *The Structure of Science*. London: Routledge.
- Nagel, E. 1974. *Teleology Revisited*. New York: Columbia Press.
- Neapolitan, R. 2003. *Learning Bayesian Networks*. Upper Saddle River: Prentice-Hall.
- Olsson, E. J. 1999. "Cohering with." *Erkenntnis* 50 (2/3): 273–91.
- Olsson, E. 2017. "Coherentist Theories of Epistemic Justification." In *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, edited by E. N. Zalta, Stanford University, spring 2017 edition.
- Pauli, W. 1973. *Pauli Lectures on Physics: Thermodynamics and the Kinetic Theory of Gases*. Cambridge: MIT Press.
- Sarkar, S. 2015. "Nagel on Reduction." *Studies in History and Philosophy of Science Part A* 53: 43–56.
- Schaffner, K.F. 1967. "Approaches to Reduction." *Philosophy of Science* 34 (2): 137–47.

- Schaffner, K. F. 1974. "Reductionism in Biology: Prospects and Problems." In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 613–32.
- Schaffner, K. F. 1977. "Reductionism, Values, and Progress in the Biomedical Sciences." In *Logic, Laws, and Life*, edited by R. Colodny, 143–71. Pittsburgh: University of Pittsburgh Press.
- Schaffner, K. F. 1993. *Discovery and Explanation in Biology and Medicine*. Chicago: Chicago University Press.
- Schupbach, J. N. 2011. "New Hope for Shogenji's Coherence Measure." *British Journal for the Philosophy of Science* 62 (1): 125–42.
- Schaffner, K. F. 1969. "The Watson–Crick Model and Reductionism." *British Journal for the Philosophy of Science* 20 (4): 325–48.
- Shogenji, T. 1999. "Is Coherence Truth Conducive?" *Analysis* 59 (264): 338–45.
- Tešić, M. 2019. "Confirmation and the Generalized Nagel–Schaffner Model of Reduction: A Bayesian Analysis." *Synthese* 196 (3): 1097–129.