The Justification Gap Problem for Theory Reduction

Lorenzo Lorenzetti*

August 7, 2024

Abstract

Our most successful and widely adopted models of reduction between scientific theories can be categorized into Nagelian and mathematical approaches. We argue that both accounts are critically incomplete due to what we term the *justification gap problem*. This issue stems from the lack of justification for the mathematical mappings and bridge laws these approaches use. We propose that integrating these models with a functionalist view of theoretical quantities can bridge this gap. Hence Nagelian and mathematical models should be turned into forms of functional reduction, a less common but increasingly relevant alternative approach to theory reduction. This conclusion underscores the superiority of functional reduction, revises how we conceptualise Nagelian and mathematical reduction, and counters recent arguments raised by Knox and Wallace (2023) on functionalism and reduction.

Contents

1 Introduction **2**

2 Justification Gap in Mathematical Reduction **5**

3 Justification Gap in Nagelian Reduction **8**

*University of Bristol, contact: lorenzo.lorenzetti.ac@gmail.com

4 Functionalism and the Justification Gap **13**
4.1 Functionalism in the Philosophy of Science 13
4.2 Functionalism and Mathematical Reduction 15
4.3 Functionalism and Nagelian Reduction 17

5 Implications: a Reply to Knox and Wallace **20**

6 Implications: Rethinking Nagelian Reduction **21**
6.1 How Nagelian and Functional Reduction Work 22
6.2 How are Nagelian and Functional Reduction Related? 23
6.3 Nagelian as a Subtype of Functional 24

7 Conclusion: Future Directions **26**

1 Introduction

Reduction between scientific theories is one of the most central topics in the philosophy of science. Establishing reduction relations between theories or parts thereof is essential to science. Inter-theory reduction plays a role in the justification of the success of the reduced theories; in the acceptance or consolidation of a new theory, once we show that the new theory can be used to reduce an older already accepted theory; and in providing a heuristic guide, since the success of an already accepted to-be-reduced theory poses constraints to the development of a more fundamental reducing theory. Moreover, clarifying the relationship between theories describing the world at different levels is pivotal to understanding how reality is structured.[1]

Several accounts of inter-theory reduction are available. We focus on reduction in the physical sciences, but this discussion is also relevant beyond this area.[2] The two most successful and widely adopted approaches can be categorised as follows.

First, *mathematical reduction*. Within this family of accounts, reduction is formulated in terms of (mostly) mathematical relations between the models of the reducing and reduced theories, usually characterised in terms of mappings, limiting relations, or derivations.[3] Second, *Nagelian reduction*, which includes the original account by

[1]See Palacios (2022, 2023, 2024), van Riel and Van Gulick (2019), Crowther (2018).

[2]Inter-theory reduction plays a crucial role also in the philosophy of mind, where for example Nagelian reduction is widely discussed. See Kim (1998, 2005) and Marras (2005).

[3]Proponents of mathematical approaches to reduction include Rosaler (2015, 2019), Wallace (2022), Ladyman and Ross (2007), Nickles (1973), Fletcher (2019), Balzer et al. (1987).

Ernest Nagel and its modern iterations, as well as the so-called Generalised Nagel-Schaffner model.[4] Within this approach, reduction is formulated as the derivation of the laws of the reduced theory from the laws of the reducing theory with the aid of 'bridge laws' connecting the vocabularies of the two theories. For instance, in the context of the reduction of thermodynamics to statistical mechanics, we can derive the Boyle-Charles law from statistical mechanics' laws given a bridge law stating that 'temperature' is 'mean kinetic energy' (Dizadji-Bahmani, 2021).

In both accounts, reduction is achieved by drawing a formal connection between reduced and reducing theory. Within mathematical reduction the key link is represented by a mathematical mapping between the models, within Nagelian reduction bridge laws play a crucial role in linking the theories. These links connect theoretical elements within the theories or the theoretical models, such as terms, quantities, variables, or structures.

We present a challenge that undermines both major approaches to inter-theory reduction, called the *'justification gap problem'*. The issue takes a different form in each approach. Within mathematical reduction, the challenge is that formulating reduction merely in terms of mathematical relations does not explain why a given relation between mathematical structures qualifies as reduction. That is, why are we justified to claim such mapping counts as a reduction instead of other arbitrary mappings? Within Nagelian reduction, the challenge concerns how bridge laws are obtained.[5] Given a specific bridge law, we can ask: how is this bridge law justified? What makes it not *ad hoc*? Without an adequate explanation to fill these gaps, an account of inter-theory reduction is crucially incomplete.

We argue that adopting a functionalist view about the nature of theoretical elements like quantities and importing it within each account can solve the problem. Functionalism is the thesis that 'to be x is to play the role of x'. Hence theoretical terms, concepts, or quantities are defined by their roles in theories, and properties are characterised in terms of their behaviour. This view has been extensively applied to the philosophy of mind and more recently to the philosophy of science, to solve certain issues. In the former context mental states like pain are analysed in terms of their causal role, in the latter theoretical elements such as spacetime are functionally characterised in terms of their role in scientific theories.[6] Furthermore, functionalism

[4]See Nagel (1961, 1970). For a recent defence of Nagel see van Riel (2011), Butterfield (2011), and Sarkar (2015). On the Generalised Nagel-Schaffner model see Schaffner (1967, 2012) and Dizadji-Bahmani et al. (2010).

[5]This is based on the explanatory gap problem raised by Kim (1998, 2005, 2008).

[6]On functionalism in mind see e.g. Putnam (1967), Fodor (1968), Lewis (1972). On functionalism in science see e.g. Lam and Wüthrich (2018), Knox (2014, 2019), Read (2018), Roberts (2022), but the view traces back to Lewis (1970) and Carnap (1958).

has been used to model instances of reduction, in the form of *functional reduction.* In the philosophy of mind it is used to reduce phenomenal states to brain states, in the philosophy of science to formulate inter-theory reduction, such as reducing spacetime to non-spatiotemporal structures that can play the spacetime role.[7]

Functionalism solves the justification gap problem by providing a justification as to why a given map or bridge law constitutes reduction. Given functionalism, mathematical mappings and bridge laws establish reductive relations in a justified way when they link, e.g., theoretical quantities that play the same theoretical role. Adopting this solution would crucially transform how inter-theory reduction is conceptualised. By combining functionalism with mathematical and Nagelian reduction we turn them into forms of functional reduction. In particular, they become two different specific versions of the general functional approach to reduction originally defended by Kim (1998, 2005). Hence mathematical and Nagelian reduction should not be conceived as independent alternative accounts but rather as subtypes of another broader approach to inter-theory reduction, namely functional reduction.

This conclusion has important implications. The position that mathematical and Nagelian reduction should be conceived as subtypes of functional reduction contrasts the mainstream view that functional reduction is either a third distinct alternative to the other major approaches to reduction or a special version of Nagelian reduction.[8] Additionally, it allows us to respond to recent papers by Wallace (2022) and Knox and Wallace (2023), who adopt mathematical reduction and argue that reduction and functionalism should be divorced. In contrast, we argue that functionalism should be embedded into mathematical reduction.

The paper is structured as follows. Section 2 introduces the justification gap problem against mathematical reduction, Section 3 shows how the justification gap affects Nagelian reduction, and Section 4 shows how adopting functionalism and implementing it into mathematical and Nagelian reduction solves the problem. Sections 5 and 6 explore further implications of the solution we propose. The former replies to Knox and Wallace (2023), the latter proposes a revision of the mainstream views concerning the relationship between Nagelian and functional reduction. Section 7 proposes future research directions prompted by the conclusions of this paper.

[7]See Lewis (1970, 1972), Kim (1998, 2005), Lam and Wüthrich (2018), Huggett and Wüthrich (2021), Butterfield and Gomes (2023, 2022), Robertson (2022), Lorenzetti (2022, 2024).

[8]The former position is defended in different ways by Kim (1998, 2005), Marras (2002, 2005), Fazekas (2009), and Morris (2020)). The latter position has been defended by Lewis (1970), Butterfield and Gomes (2022, 2023). The view that functional reduction is a broader position than Nagelian and functional reduction has been recently advocated by Lorenzetti (2024).

2 Justification Gap in Mathematical Reduction

This section introduces the mathematical approach to inter-theory reduction and then advances a version of the justification gap problem against this approach. To introduce mathematical reduction we present some of the most developed mathematical accounts of inter-theory reduction and then highlight the common core of these accounts and why the justification gap undermines them all.

We can broadly categorise as mathematical reduction those accounts according to which reduction is modelled in terms of mathematical relations between theoretical structures or quantities belonging to different models.[9] A range of different approaches fall under this characterisation, including accounts cashing out reduction in terms of mathematical relations and instantiation relations between models, accounts focused on limiting relations or derivations of specific quantities or variables, and so-called 'structuralist' accounts of reduction.[10] This paper focuses in particular on the accounts defended by Wallace (2022), Nickles (1973), and Rosaler (2015), which are well-developed and representative of mathematical reduction.

Starting with some background, one of the first versions of mathematical reduction has been endorsed by Suppes: "the thesis that psychology may be reduced to physiology would be for many people appropriately established if one could show that, for any model of a psychological theory, it was possible to construct an isomorphic model within physiological theory." (Suppes 1967, p. 59). Although the relation of isomorphism has been considered too strong in the subsequent literature, the notion of reduction as a model-model mathematical relation has remained the hallmark of the approach. Ladyman and Ross (2007) talk about reduction as a link between mathematical structures in terms of structure-preserving mappings or 'morphisms', and Wallace (2022) advocates the following view:

> **Reduction as Instantiation:** "Reduction is [...] the realizing by some substructure of the low-level theory's models of the structure of the higher-level theory's models" where "the lower-level theory instantiates the higher-level one if (roughly) there is a map from the lower-level state space to the higher-level state space that commutes with the dynamics

[9] The approach is usually combined with a semantic or 'maths-first' view of scientific theories, taking scientific theories as constituted by sets of models that are mainly mathematically formulated. See e.g. e.g. Van Fraassen (1980) and Ladyman and Ross (2007, p. 118).

[10] On 'model-based' or 'instantiation-based' see Ladyman and Ross (2007), Rosaler (2015, 2019), and Wallace (2022)); on 'limit-based' and 'mathematical-derivation-based' see Nickles (1973) and the interpretation of Nickle's account by Palacios (2023); on 'structuralism' see Balzer et al. (1987). For a recent introduction see Palacios (2023, 2024) and van Riel and Van Gulick (2019).

and leaves invariant any commonly-interpreted structures (for instance, spacetime structure) in the two theories." (Wallace, 2022, p. 357)

This view characterises reduction as a primarily local relation, that takes place between specific models and hence regards specific parts of the theories. Furthermore, the model-to-model mappings play a key role in formulating reduction.

Another major approach in this family is the one introduced by Nickles (1973) which is referred to as 'Reduction$_2$'. As nicely presented by Palacios (2023):

> **Reduction$_2$**: Let O_i be a set of intertheoretic operations, then a theory T_2 reduces$_2$ to another T_1 iff $\mathrm{O}_i(\mathrm{T}_1) \rightarrow \mathrm{T}_2$, where the arrow represents "mathematical derivation" understood in a broad sense including not only logical deduction but also limiting operations and approximations of many kinds.

Palacios (2023, p. 19) also notes that "mathematical operations such as limits and other approximations are performed not on the theory itself but on functions (or equations) representing physical quantities" and reformulates Nickles' model in a more local way as follows:

> **Reduction*$_2$**: Given a set of intertheoretic operations O_i, a quantity Q_1 of T_1 reduces*$_2$ another quantity $\mathrm{Q}^*{}_1$ of $\mathrm{T}^*{}_1$ iff (i) $\mathrm{O}_i(\mathrm{Q}_1)=\mathrm{Q}^*{}_1$ and (ii) the mathematical operations O_i make physical sense.

Palacios (p. 19) also stresses how the key notion of 'making physical sense' is underspecified. We return to this point when we argue in detail how adopting functionalism improves mathematical models of reduction. For the moment, we highlight how the approach focuses on mathematical connections between theories to characterise reduction, like Ladyman and Ross and Wallace do, as well as Palacios' remark that Nickles' mathematical reduction is first and foremost concerned with drawing mappings (characterised as mathematical operations) between specific functions and quantities within the theories. The latter aspect closely resembles Rosaler's account which we introduce next.

Rosaler (2015, 2019) recently developed another model-based account of reduction which can be classified as a version of mathematical reduction:

> **Rosaler's Model-based Reduction:** "Theory T_h reduces$_T$ to theory T_l iff for every system K in the domain of T_h – that is, for every system K whose behavior is accurately represented by some model M_h of T_h – there

> exists a model M_l of T_l also representing K such that M_h reduces$_M$ to M_l" where a low-level model reduces a high-level model if "the low-level model accounts for the success of the high-level model at tracking the behaviour of the system in question." (Rosaler, 2015, p. 59)

More precisely, "for every physically realistic solution $x_h(s)$ of the high-level model M_h, there exists a solution $x_l(s)$ of the low-level model M_l such that $B(x_l(s))$ approximates $x_h(s)$ to within a margin of accuracy that is at least as small as the margin within which $x_h(s)$ tracks the relevant features of the system K" (Rosaler 2019, p. 293), where B is some function mapping solutions in the low-level state space to solutions in the high-level state space.

Here is an example of reduction drawn from Rosaler (2015, §5.1) and focused on quantum-classical reduction. It illustrates how reduction can be framed within Rosaler's approach as well as the other mathematical approaches presented above.

A semi-classical model for a point-particle system can be mathematically matched with a quantum model of the same system, under the right conditions. Thanks to the Ehrenfest theorem, we can derive Newton's law from the Schrödinger equation for the system if the particle is highly localised in space. This means that, within the quantum model, the centre of the localised wavepacket has a trajectory in configuration space that is (to a high approximation) identical to the trajectory in configuration space of a point particle of mass m within classical mechanics (in the Hamiltonian formulation). Thus, the trajectory of the wavepacket can be practically considered as a solution to the classical dynamic equation for a classical particle, and we can draw a map between the quantum and the classical models defined over the respective state spaces. The conclusion is that, since the mapping can be established and the two models linked, we have established an instance of inter-reduction between quantum and classical mechanics. The example fits very naturally within Wallace's and Rosaler's accounts. It can be framed within Nickles' account by noting that reduction is established by mapping between specific quantities describing the system at different levels, such as position and momentum.

Having presented mathematical reduction, we introduce the justification gap problem against this approach to reduction. Here is the issue in a nutshell:

> **Justification Gap Problem (mathematical reduction):** Formulating reduction merely in terms of mathematical relations between models does not explain why a given mapping between mathematical structures or quantities within those models qualifies as reduction.

Let's analyse how this issue affects each of the three accounts presented so far, starting with 'reduction as instantiation'. The model-based reduction defended by

Knox and Wallace formulates reductive relations as mappings between lower-level and higher-level theoretical models established by mathematical derivations. But, the question is, *why* does a specific mapping establish reduction? Can any mathematical relation between models of different theories constitute a reductive relation? Clearly not any arbitrary mapping can qualify as a reductive relation. The absence of an explanation opens a gap for this account of reduction. Filling this gap requires a principled explanation of what makes a model-to-model mapping a genuine instance of reduction. More precisely, we lack a *justification* for the claim that a given mapping qualifies as a reduction.[11]

Rosaler's account is very close to Wallace's (2022) approach to reduction. Hence we can easily extend the questions raised above to Rosaler's account. Recall the example of mathematical reduction presented above. Rosaler points out that, for certain kinds of physical systems (highly localised quantum systems), we can build mathematical mappings between the state spaces for the same system within the lower and the upper theories' models. That account of reduction is a mathematical reduction in the sense that we provide an asymmetrical inter-level mathematical map between the two models in state spaces. However, it may be asked why the mapping provides a reason to believe that we can recover the classical system from the quantum one. What explains that finding a lower-level model that approximately tracks the behaviour of the upper-level model justifies the reduction claim?

The same argument can be developed towards Nickles' Reduction$*_2$ account as well. In this context, the justification gap problem is closely related to the question of what qualifies a mathematical operation as making 'physical sense', as required by the account. Palacios (2023) indeed criticises the account on the basis that we lack a clear specification of the crucial notion of 'physical sense'. The issue of asking whether an operation makes physical sense and thus qualifies as reduction is the same kind of challenge raised by the justification gap problem. What makes a given transformation supportive of reduction, as opposed to another arbitrary one? Section 4 argues that adopting functionalism and implementing it within each of these accounts of mathematical reduction can solve the justification gap problem.

3 Justification Gap in Nagelian Reduction

This section introduces the Nagelian approach to reduction and then raises the justification gap problem against the account. Within Nagelian reduction we include

[11] A similar worry is raised by Palacios (2023, p. 26) about the analogous *structuralist account* that expresses reduction in terms of transformations between models.

both the standard model by Nagel (1961, 1970) and the so-called Generalised Nagel-Schaffner model by Schaffner (1967, 2012) and Dizadji-Bahmani et al. (2010). This is the mainstream approach to theoretical reduction and it has been adopted, among others, also by van Riel (2011), Butterfield (2011), and Sarkar (2015).[12]

We focus here on Nagel's (1961) classic model of reduction, since the improvements brought by the newer versions of the account do not matter for our discussion. According to Nagel's account, a theory T_P can be reduced to another theory T_F iff the laws of T_P can be deduced from the laws of T_F plus some auxiliary assumptions. In case the two theories do not share their theoretical terms, we need also to postulate bridge laws (also called 'conditions of connectability') which connect the different vocabularies of the two theories. Indeed, since different theories may include different terms in their vocabularies, to derive the laws of the reduced theory from the reducing theory, for each term which occurs in the reduced theory but not in the reducing theory, there must be a connecting statement linking the term with an expression in the reduced theory.

For instance, in the context of the reduction of thermodynamics to statistical mechanics, we can derive the Boyle-Charles law from statistical mechanics' laws given a bridge law stating that 'temperature' means 'mean kinetic energy' (Dizadji-Bahmani, 2021).[13] Consider the Boyle-Charles law of thermodynamics:

$$PV = kT \tag{1}$$

where P is pressure, V is volume, T is temperature, and k is a constant. In statistical mechanics, we can formulate the following law, where E_{kin} is the kinetic energy (defined as $mv^2/2$):

$$PV = (2/3)\langle E_{kin}\rangle. \tag{2}$$

Then, if we associate and replace the thermodynamic quantity T with the statistical mechanical quantity $\langle E_{kin}\rangle$, we can deduce (1) from (2), up to a constant. This connection plays the role of the Nagelian bridge law in this case.

[12]An alleged alternative approach to Nagelian reduction is the so-called "New Wave" approach (Bickle, 1996). We set this account aside here as we share the now widespread opinion that New Wave Reductionism is not an alternative to Nagelian reduction but rather collapses on the latter on closer scrutiny (cf. Dizadji-Bahmani et al. (2010)).

[13]This is a simplistic example which has been criticised as imprecise (cf. Bangu (2011)). However is commonly used to introduce Nagelian reduction as it shows how Nagelian reduction works in a very clear and simple way, hence we adopt it for the sake of this presentation. Nothing substantial within our discussion depends on the physical adequacy of this example.

This example shows how bridge laws work and why they are crucial to the law-deduction process on which the Nagelian account is based. However, a few other questions can be raised to understand the Nagelian view in more detail. First, what are bridge laws, syntactically speaking? That is, how can they be formalised? Second, a family of interrelated questions: *(a)* What kind of connections bridge laws are, beyond their specific syntactic form? *(b)* How are bridge laws established, i.e. how do we come to know a given bridge law? *(c)* Why are we justified to believe in any given bridge law? These questions about the nature of the bridge laws have been extensively discussed and often conflated, but we keep them clearly distinguished here.[14] We start by analysing the first question and then discuss the second group of queries. In particular, we are mainly interested in question (c) concerning the justification of the bridge laws, which raises an justification gap problem similar to the challenge we raised against mathematical reduction.

Concerning the syntactic aspect of bridge laws, three main options are available to describe what kind of connections bridge laws are: *(i)* conditional statements of the form 'for all x, if x is t_F then it is t_P', where t_F denotes a term in the reducing theory and t_P denotes a term in the reduced theory; *(ii)* bi-conditional statements, similarly to the first option; *(iii)* identity statements, identifying the terms belonging to different theories. Different positions have been defended concerning this issue, notably Kim (1998) and van Riel (2011) argue that only bridge laws in the form of identity statements can really support reduction, whereas Dizadji-Bahmani et al. (2010) argue for the minimal view of bridge laws as conditional statements, since "All we need for the deduction is that whenever t_F applies, then t_P applies" (p. 406).

However, specifying the syntactic form bridge laws should take does not settle questions (a)–(c). Consider the example of the Boyle-Charles law within Nagelian reduction. We noted that the derivation of the thermodynamical law (1) from the statistical mechanical (2) is based on the bridge law allowing for the replaceability of temperature and mean kinetic energy. Setting the syntactic nature of the link aside, we can ask: what kind of link is this? Is the bridge law e.g. a factual or an analytic statement? Furthermore, how do we know that given bridge laws hold in the first place? How do we know that there is a co-reference between the terms? Finally, how is that specific bridge law justified? Why do we believe in that bridge law instead of another arbitrary one?

Let's analyze questions (a) and (b) about the content and epistemology of bridge laws and review the answers provided in the literature. These two questions have been addressed by Nagel (1961, ch. 11) himself. He discusses three candidate ways

[14]See Beckermann (1992), Kim (1998), Fazekas (2009), Klein (2009), Dizadji-Bahmani et al. (2010), and van Riel (2011) on the nature of bridge laws.

to describe what kind of links the bridge laws are. Each of them is related to an epistemological account of the bridge laws. These alternatives should be distinguished from the three options reviewed above concerning the *form* of the bridge laws and taken as orthogonal to them. Here are the options:[15]

- Bridge laws are 'logical connections', that is analytic statements or meaning connections, established by grasping the meaning of the linked terms.
- Bridge laws are 'conventions' or stipulations, established by fiat. Hence they have the epistemological status of conventions.
- Bridge laws are factual statements or 'material' hypotheses, empirically established. They can be for example mere *de facto* correlations, nomic connections, or ontological links (e.g. identities or relations among extensions).

These accounts specify what is the content of bridge laws and, relatedly, how we can have access to them. A further question, our main focus, concerns the justification of any given bridge law. This is a kind of justification gap problem:

> **Justification Gap Problem (Nagelian reduction):** Why are we justified to maintain that a given bridge law holds?

A challenge of this kind against Nagelian reduction has been famously proposed by Kim (1998, 2005, 2008). We introduce Kim's objection and elaborate on it, confronting it with the different views on the nature of bridge laws presented so far.

Kim is focused on the application of Nagelian reduction to the philosophy of mind and argues that Nagelian bridge laws are unable to close the gap between phenomenal states and neural states. For instance, he considers a bridge of the form 'Pain occurs to $x \leftrightarrow$ neural state N_1 occurs in x' and argues that such bridge law leaves open questions like "why does pain correlate with N_1 rather than another neural state?; why doesn't itch correlate with N_1?; why does any qualitative experience correlate with N_1?; and so on." (Kim, 2008, pp. 98-99). Hence the bridge law itself needs explanation.

While Kim is concerned with the issue of reducing and explaining the mental, this challenge affects Nagelian reduction in general as an approach to inter-theory reduction, as often recognised. For example, Marras (2005) and Fazekas (2009) describe Kim's challenge as the general issue that Nagelian bridge laws are 'unexplained auxiliary premises' themselves in need of explanation.

[15] Cf. Dizadji-Bahmani et al. (2010, p. 400), van Riel (2011, p. 358), and Fazekas (2009, §2.2).

Can the answers to (a)–(b) presented above provide a way to address the justification gap problem?[16] We argue that they can't or, at best, they are incomplete and thus some details are missing. The next section will then explain why adopting functionalism can fill the gaps and solve the problem.

If we believe that bridge laws are conventions, stipulated *by fiat*, it is clear that a justification is missing. Holding bridge laws as brute facts merely sidesteps the issue of specifying why we are justified in believing a specific bridge law. If we hold bridge laws as 'meaning connections' between terms or 'factual statements' things are trickier. According to the former option, we would be justified in believing a given bridge law in virtue of the meaning of the terms involved. A given bridge law holds if we can replace the higher-level term with the lower-level term in the law because the terms have the same meaning. According to the latter option, we are justified in believing in a given bridge law if e.g. we can establish by experiment that mean kinetic energy and temperature are the same thing.

Here is why both alternatives are incomplete and why they plausibly lead us to adopt functionalism. We sketch the main points which Section 4 will elaborate on.

If we want to maintain that bridge laws are 'meaning connections' and that we can replace one term with another because they share the same meaning, we need an account of the meaning of theoretical terms or quantities that allows us to claim that e.g. temperature and mean kinetic energy have the same meaning. Nagel (1961) himself is sceptical about this route and discards this alternative, focusing on the two other options. On the other hand, we will argue that adopting functionalism as a thesis about the meaning of terms and quantities can complete the bridge-laws-as-meaning-connections account and address the justification gap problem. However, of course, this means that it is functionalism that solves the justification gap problem.

If we want to maintain that bridge laws are factual statements that are empirically determined, we face another issue. As granted by Nagel (1961, p. 356) and acknowledged by Dizadji-Bahmani et al. (2010), the problem is that bridge laws cannot be tested independently. There is no way to empirically test in a direct way that temperature and mean kinetic energy are interchangeable. Rather, we empirically obtain the laws and then deduce the bridge laws from there, as Dizadji-Bahmani et al. put it very clearly:[17]

> It is not the case, as Nagel seems to suggest, that we start with T_F, then write down a bridge law (which we know to be correct!), and finally

[16]For instance, Klein (2009) maintains that those three alternatives actually concern the *justification* of the bridge laws as claims of co-reference.

[17]Marras (2005, p. 343) also grants that Nagelian bridge laws are derived from a comparison between the laws of the theories at different levels.

> deduce T_P. Rather, what happens is that we begin with T_F and T_P and then try to find bridge laws that (modulo small corrections) make T_P derivable from T_F [...] It is not the Boyle–Charles Law that we derive from the kinetic theory plus a bridge law; it is the bridge law that is derived from the Boyle–Charles Law and the kinetic theory. (Dizadji-Bahmani et al., 2010, pp. 407-8)

This view provides a more adequate epistemological account in support of the position that bridge laws are factual statements. Bridge laws are obtained via the laws in which the terms appear. This suggests a possible answer to the justification gap problem: we are justified in believing a bridge law if it allows us to derive the upper-level law from the lower-level law. However, this would clearly make bridge laws *ad hoc* statements and make reduction a trivial process. We argue in the next section that this is not so, provided we implement this position with a functionalist view of theoretical terms and quantities. Hence, once again, it is functionalism that solves the problem and fills the justification gap.

4 Functionalism and the Justification Gap

This section argues that adopting functionalism and combining it with mathematical and Nagelian reduction can solve the problem raised by the justification gap, as it provides a justification for mathematical mappings and bridge laws. Section 4.1 introduces functionalism in the philosophy of science and its connection with inter-theory reduction, Section 4.2 shows how functionalism solves the justification gap problem within mathematical reduction, and Section 4.3 shows how it solves the problem within Nagelian reduction.

4.1 Functionalism in the Philosophy of Science

In the broadest terms, functionalism is the following thesis:[18]

> **Functionalism:** to be x is to play the role of x.

According to this view, theoretical terms, concepts, or quantities are defined by the roles they have in theories, and properties are cashed out in terms of their causal roles or behaviour. This view has been widely applied to the philosophy of mind. Recently it has been extensively implemented into the philosophy of science

[18]Cf. Robertson (2022, p. 988) and Lorenzetti (2023, p. 922).

as well. For example, Knox argues that spacetime should be defined in terms of the spacetime role and "A structure will play the spacetime role in our theories just in case it describes the structure of the inertial frames, and the coordinate systems associated with these." (Knox, 2014, p. 15).

These accounts aim to analyse *what it takes* for a concept or quantity to count as e.g. spacetime, thermodynamic entropy, or time, and do so by looking at their role within our best scientific theories. We can say that analysing what it takes for a given quantity to count as that particular quantity means providing a definition of it. Given that we are dealing with concepts and quantities within theories, functionalism can alternatively be framed as a thesis concerning the definition of theoretical terms, in the sense of theoretical term employed by Lewis (1970, 1972), according to which a theoretical term is a term introduced by a theory. This way to frame functionalism is common in the literature.[19] For the purpose of our paper, nothing crucial depends on this, hence we will interchangeably talk about functionalism about quantities or concepts or functionalism about theoretical terms.

Functionalism also has a major role in the topic of reduction in mind and science, where several functional reductionist approaches have been developed.[20] Our main focus here is on the use of functionalism within inter-theory reduction in science. In this context, functional reduction is primarily used to model cases of reduction between scientific theories. It has been used for instance to model reductive relationships between thermodynamics and statistical mechanics, between classical and quantum mechanics, and between general relativity and quantum gravity theories. According to functional reduction, the primary aim of reduction is to find the right lower-level realisers for the upper-level behaviour: reduction is secured if we find in the lower-level theory some theoretical elements that play the functional roles described by the upper-level theory. For instance, let's say we can functionally define 'temperature' in terms of its role within thermodynamics, and we find out that 'mean kinetic energy' plays the role of temperature: in that case, we can functionally reduce temperature to mean kinetic energy, and this can be regarded as a step in the reduction of thermodynamics to statistical mechanics.

In the most general terms, the core of the functionalist approach to inter-theory reduction can be represented as follows (cf. Kim, 2005, pp. 101-102):[21]

[19]See for example Lewis (1970, 1972) but also Butterfield and Gomes (2023), Huggett and Wüthrich (2013, 2021), Baron (2022), Lorenzetti (2022).

[20]On functional reduction in the philosophy of mind see e.g. Lewis (1972), Kim (1998, 2005), Morris (2020). On functional reduction in science see Palacios (2024) for an introduction, and also Esfeld and Sachse (2007), Lam and Wüthrich (2018, 2020), Huggett and Wüthrich (2013, 2021), Butterfield and Gomes (2022, 2023), Robertson (2022), Lorenzetti (2022, 2023, 2024), Albert (2015).

[21]This is adapted from the original functional reductionist approach by Kim, which is directly

Core Functional Reduction:

- **[Step 1]** The quantity M to be reduced is functionally defined as that quantity that plays the functional role R in the theory T_P.
- **[Step 2]** Find a quantity in the reduction base, i.e. in the domain of another theory T_F, that performs the functional role R.

The next two subsections argue that adding functionalism about theoretical terms or quantities to mathematical and Nagelian reduction can solve the justification gap problem. If one adopts this strategy, mathematical and Nagelian reduction becomes subtypes of core functional reduction, focused on either mathematical relations between models or law derivations.

4.2 Functionalism and Mathematical Reduction

Section 2 reviewed the mathematical approach to reduction by Wallace (2022), Nickles (1973), and Rosaler (2015). We show how adopting functionalism can solve the justification gap problem within each account.

According to Wallace, reduction is formulated in terms of the realisation of structures within higher-level models by structures within lower-level models, where "the lower-level theory instantiates the higher-level one if (roughly) there is a map from the lower-level state space to the higher-level state space that commutes with the dynamics and leaves invariant any commonly-interpreted structures (for instance, spacetime structure) in the two theories." (Wallace, 2022, p. 357). The problem is explaining *why* drawing a mathematical mapping between two state spaces in which the dynamics commute is enough to prove the existence of a reduction relation. What makes this mapping the right one for reduction as opposed to others? How is this not a merely arbitrary mathematical relation?

Consider the example introduced before. A semi-classical model for a point-particle system can be matched with a quantum model of the same system, under the right conditions. The quantum system's trajectory can be approximately considered as a solution to the classical dynamic equation for a classical system, and we can draw a map between the quantum and the classical models over the respective state spaces. Why does this count as a reduction?

If we adopt functionalism, we can provide a justification and reply to this challenge. Or, to say the least, explicitly endorsing functionalism exposes an already

focused on properties rather than theoretical quantities. See also Marras (2005, p. 345).

implicit assumption. The functionalist would maintain that 'being a classical system' just means being something that *performs certain roles within the model of classical mechanics*. Then, functionalism justifies why, to reduce a classical model to a quantum model, all we need to do is provide an account of how the quantum model can represent the behaviour described by the classical model. According to functionalism the condition for being a classical system is to play a certain role in the classical models, and the mathematical mapping at stake shows exactly that the quantum system can indeed evolve like the classical one. Hence the quantum system can be regarded as a classical system in the right conditions (the conditions for classicality) given that it performs classical behaviour and to be a classical system is to perform classical behaviour.

An example considered by Knox and Wallace (2023) is the reduction of Newtonian gravity to general relativity within a mathematical approach to reduction. As before, we can see here how the justification for the instance of reduction Knox and Wallace defend can be provided by functionalism about theoretical concepts. The justification gap can be filled by functionalism in the following way. In their paper, Knox and Wallace show that, in the right conditions, a given general relativistic system modelled via Einstein field equations can evolve according to the Newtonian Poisson equation. This allows a mapping between certain models of general relativity and Newtonian models. However, we still need a reason why to *reduce* a Newtonian gravitational system to general relativity all we need to do is show how the general relativity model evolves according to the equation characterising the Newtonian model in the right conditions. If we adopt functionalism we can claim that 'being a Newtonian gravitational system' just means performing certain roles within the models of Newtonian gravitation. In this case, the role is codified by the Poisson equation. That is the functional role that the relativistic system is able to play given the right conditions. This particular mathematical derivation (and *not* any other arbitrary relation) constitutes an instance of reduction because this precise derivation allows us to demonstrate that a class of relativistic systems in the right regime can realise the role which defines Newtonian gravitational systems. Hence they can be regarded as Newtonian gravitational systems in that regime.

Let's now consider how this analysis generalises, starting with Rosaler's account. That view is very close to Wallace's (2022) approach to reduction. As such, we can easily extend the arguments developed above to Rosaler's account. What explains the fact that finding a lower-level model that approximately tracks the behaviour of the upper-level model justifies the reduction claim? A functionalist assumption would be able to fill the justification gap: the justification stems from the fact that being a given upper-level system is to behave in a certain way as represented by the upper-level model.

The same kind of argument can be developed towards Nickles' Reduction$*_2$ account. And, similarly, functionalism can explain the notion of 'physical sense' justifying the mathematical operations. Palacios (2023) indeed criticises the account on the basis that a clear specification of the crucial notion of 'physical sense' is missing from the account. The issue of asking whether an operation makes physical sense and thus qualifies as reduction is the same kind of challenge raised by the justification gap problem. Hence functionalism can overcome the justification gap in this case too, thereby explaining what it means to make physical sense: the appropriate mathematical operations that validate reductive relations within Nickles' Reduction$*_2$ are those allowing us to show that a given lower-level quantity has the same functional role in the theory (in the right context) as the upper-level quantity.[22]

We showed how the justification gap problem applies to mathematical reduction, showing how it can be raised against leading approaches. Explicitly adopting functionalism in each case justifies how the account works, solving the justification gap problem. This is not to say that adopting functionalism is a necessary condition to solve the issue. However, absent a better alternative solution, functionalism provides an ideal strategy and is strongly supported.

Crucially, combining functionalism with mathematical reduction transforms mathematical reduction: the latter approach would become a form of functional reduction, in particular a more specific version of core functional reduction which is focused on mathematical structures and models and which expresses functional roles in terms of roles in the models of a theory.

4.3 Functionalism and Nagelian Reduction

Section 3 reviewed Nagelian reduction and the debate about Nagelian bridge laws. We raised a justification gap problem concerning the justification of bridge laws. This section shows how combining functionalism with Nagelian reduction solves the problem.

Consider the toy example of Nagelian reduction described by equations (1) and (2) and the bridge law involved. Regardless of the syntactic form the bridge law connecting T and $\langle E_{kin} \rangle$ takes and the kind of link the bridge law can be, we can ask why we should believe in such bridge law. What justifies replacing $\langle E_{kin} \rangle$ with T in the lower-level law to derive the higher-level law?

[22]For example, the functional reduction of thermodynamic entropy S_{td} to Gibbs entropy S_g defended by Robertson (2022) is arguably a kind of quantity-to-quantity Nickles-reduction that also employs functionalism to justify why the mapping between these quantities qualifies as reduction.

Here is how adopting a functionalist view about the identity of theoretical quantities can address the challenge. Let's say, for the sake of argument, that temperature T is defined in terms of its role within thermodynamics as specified by the ideal gas law (1). Why should we be justified in stipulating a bridge law connecting it with $\langle E_{kin} \rangle$ and allowing for the replaceability of the latter with the former? If we endorse functionalism, quantities are individuated by their roles in theories. Hence the connection between the quantities is justified by the fact that in our context the two quantities share the same nomic profile, given that they (approximately) instantiate the same relations with other quantities as represented in (1) and (2). Functionalism justifies why we can use T and $\langle E_{kin} \rangle$ interchangeably within that context, vindicating the postulation of the bridge law: T and $\langle E_{kin} \rangle$ play the same role, and thus if a quantity is defined by the role it plays in the theory, then the two quantities can be used interchangeably.[23]

Let's now reconsider the discussion at the end of Section 3. We presented the standard positions concerning the content of bridge laws, according to which bridge laws can be meaning connections, conventions or factual statements. We maintained that endorsing either of those views does not address the justification gap problem, whereas adopting functionalism can bridge the gap. More precisely, we noted that taking bridge laws as conventions does not address the issue in the first place, whereas adopting one of the two other options provides a starting point, although they should be combined with functionalism to solve the justification gap problem. The following shows how functionalism can be combined with believing that bridge laws are meaning connections or factual statements and how it solves the issue.

Consider the view of bridge laws as meaning connections. If you believe that we can replace one term with another because they share the same meaning, then you need an account of the meaning of theoretical terms that allow you to claim that temperature and mean kinetic energy have the same meaning. Functionalism precisely provides such an account. If one holds that the meaning of a term or quantity is fixed by its role in the theory, and in our case believes that T and $\langle E_{kin} \rangle$ play the same role, at least in the right context, in virtue of bearing the same relations with the other quantities respectively present in equations (1) and (2), then the two quantities have the same meaning. Hence if bridge laws are meaning connections that establish that two quantities are interchangeable, a functionalist theory of meaning of theoretical terms solves the justification gap problem.

Consider the view of bridge laws as factual statements. We saw how they cannot be empirically tested independently. The natural solution provided by Dizadji-Bahmani et al. (2010) and others is to say that we empirically obtain the laws and

[23] This is compatible with all the three views concerning the syntactic form of the bridge laws.

then derive the bridge laws by comparing the equations. Functionalism provides a non-*ad-hoc* criterion to justify a given bridge law. It goes beyond saying that a bridge law is justified when it allows us to derive the upper-level law from the lower-level law. If you are a functionalist, you can maintain that comparing equations (1) and (2) allows us to establish that T and $\langle E_{kin} \rangle$ play the same (functional) nomic role. A common view among functionalists is that functional roles are determined via the Ramsey sentence of a theory, which is built from the laws of the theory, hence the functional role of a quantity is precisely determined by its relation with other quantities as represented in the laws of the theory.[24] Since quantities are characterised in terms of their functional roles, we can derive in a justified way the bridge law connecting T and $\langle E_{kin} \rangle$ from the two equations, as opposed to other bridge laws. The bridge law is derived not just because it is the only connection that allows us to derive (1) from (2), but it is justified because the same functional role characterises the quantities.

Hence, regardless of the view one takes about the content of bridge laws, functionalism can be combined with such an account to solve the justification gap problem.

As in Section 4.2, we stress that we are not claiming that functionalism is a necessary condition for the establishment of Nagelian bridge laws. Still, this proves that functionalism can be included within Nagelian reduction to bridge a justification gap and to provide an account of the epistemology of bridge laws. Hence, absent a better non-functionalist explanation of why we are justified to draw any given bridge laws in a principled way, this supports the claim that functionalism should be included within Nagelian reduction. One crucial implication of this discussion is that, just like for mathematical reduction, if we accept that functionalism should be combined with Nagelian reduction to fill the justification gap, then Nagelian reduction turns out to be a subtype of core functional reduction. Section 6 elaborates on this point. Two other considerations should be pointed out before concluding.

First, we note that this argument works only if the role those quantities bear with other quantities as represented in equations (1) and (2) can provide at least a partial definition or essential characterisation of what it means to be temperature or mean kinetic energy. This depends on the correctness of functionalism in this specific case, i.e. as applied to these specific quantities. Functionalism is not a trivial thesis but rather its adequateness depends on the quantity at stake and on how we specify the role. However, this does not harm our claim that the thesis of functionalism is a natural way to solve the justification gap problem.

Second, this section is consistent with positions discussed in the literature. For

[24]See Lewis (1970, 1972), Butterfield and Gomes (2023), Huggett and Wüthrich (2013, 2021), Baron (2022).

example, Beckermann (1992) implicitly suggests a functionalist motivation for bridge laws. Furthermore, our position agrees with Lewis's (1970) conclusion that, if we combine Nagelian reduction with a functionalist understanding of theoretical terms, bridge laws (as identities) can be functionally obtained and thus deduced from the equations, as opposed to postulated as additional empirical hypotheses. We can call this view *'Lewisian functional reduction'*. However, differently from us, Lewis takes this view as a special type of Nagelian reduction. On the other hand, in light of the justification gap problem, we maintain that Nagelian reduction is a subtype of functional reduction.

5 Implications: a Reply to Knox and Wallace

This section builds on the results of Section 4 and responds to recent papers by Knox and Wallace (2023) and Wallace (2022) on inter-theory reduction and functionalism in physics. They defend mathematical reduction and distinguish between two kinds of functionalism: *causal-role functionalism* and *constitutive functionalism*. They argue that (1) functionalism has no role to play in reduction if one endorses mathematical reduction, hence (2) functional reduction can only be formulated in terms of Nagelian reduction. We contrast both claims in light of the justification gap problem.

Causal-role functionalism is the standard Lewisian functional reduction, as discussed in Section 4.3. It is inherently reductive. It considers functionalism as a tool for inter-theory reduction and embeds it within Nagelian reduction. It is focused on finding lower-level realisers for upper-level functional roles: one first functionally defines an upper-level theoretical term and then looks for a lower-level realiser. It works by building term-wise definitions and establishing statements of co-extensivity or identities, in a bridge-laws manner. This is a version of 'core functional reduction' as defined in Section 4.1.

Constitutive functionalism is closer to Dennett's (1991) functionalism and is focused on providing functional definitions of terms or quantities. It concerns the interpretation of theoretical terms via their roles in theories, in particular within the mathematical models: "extracting a predicate description from mathematical models" (Knox and Wallace, 2023, p. 9). The approach only concerns the first step of causal-role functionalism: it concerns the functional definition but not the search for a lower-level realiser. Constitutive functionalism is neutral about the realiser, hence it is also neutral about reduction. This view corresponds to 'functionalism' as presented in Section 4.1.

Knox and Wallace maintain that mathematical reduction should be divorced from functionalism (constitutive functionalism): within the mathematical account reduction concerns mathematical mappings between models of different theories and has nothing to do with functionalism, which works as an interpretation tool.

They take liquids as an example. Within constitutive functionalism, a functionalist approach to liquids is only focused on defining their role, which is fixed by the Navier-Stokes equations: what it is to be a liquid is to obey them. Constitutive functionalism is not interested in the microphysical underpinning. The same could be said about the property of viscosity, which they claim is fixed by Navier-Stokes equations. Although this approach to functionalism is not associated with reduction, one could still ask what explains in microphysical terms that the system obeys the Navier-Stokes equations. They argue that this reductive question is tackled by looking into statistical mechanics, while functionalism does not play any role. Hence there is no space for a mathematical-reduction version of casual-role functionalism:

> Once we combine (a) the constitutive-functional analysis that tells us what it is to be a liquid of a certain viscosity is to satisfy certain equations, and (b) the derivation that certain collective degrees of freedom of microphysically characterized systems indeed do satisfy those equations, no residual reductive work remains. (Knox and Wallace, 2023, p. 7)

Based on the results obtained so far, we argue against their conclusions. As a form of mathematical reduction, the model of reduction they endorse is challenged by the justification gap problem. We need to justify why showing that certain microphysical degrees of freedom satisfy the higher-level equations should constitute a successful reduction. As articulated in Section 4.2, importing functionalism within mathematical reduction addresses this challenge. Hence (1) functionalism can play a key role in mathematical reduction and we have a strong reason not to divorce these two positions. This also implies (2) that the reductionist approach underwriting causal-role functionalism should not be restricted to Nagelian reduction only, as it can be embedded within mathematical reduction too.

6 Implications: Rethinking Nagelian Reduction

This section argues that we should rethink the relationship between Nagelian and functional reduction. Two positions have been endorsed in the literature. On one view they are regarded as distinct alternatives, on another view functional reduction is taken as a special kind of Nagelian reduction. The conclusions drawn in the previous sections support the claim that both views are misguided in the following sense. Given the justification gap problem, if we adopt functionalism as a solution, Nagelian reduction is a subtype of functional reduction.

This conclusion bears important implications. Nagelian reduction is often taken as the standard, broadest, and most minimal approach to reduction in science.[25] Reframing Nagelian reduction as a special case of a broader account would urge a revision of the debate around inter-theory reduction.

This section is structured as follows: §6.1 takes a step back and considers in more detail how functional and Nagelian reduction work as applied to the same case; §6.2 shows how the relationship between Nagelian and functional reduction is structured in the literature; §6.3 argues why Nagelian reduction should be conceived as a special version of functional reduction involving the derivation of laws.

6.1 How Nagelian and Functional Reduction Work

Within Nagelian reduction, reduction between theories is focused on the derivation of higher-level laws from lower-level laws. The Boyle-Charles law of thermodynamics (1) $PV = kT$ can be derived by the statistical mechanical law (2) $PV = (2/3)\langle E_{kin}\rangle$ provided a bridge law associating the thermodynamic quantity T with the statistical mechanical quantity $\langle E_{kin}\rangle$.

To start analysing the relationship between Nagelian and functional reduction, let's see how functional reduction would model the same case of theoretical reduction, reducing temperature to mean kinetic energy within the context of ideal gasses. Following the two-step process of core functional reduction, and assuming for the sake of the argument that (1) specifies the functional role of temperature:

- *Functional characterisation of thermodynamic quantity T:* Temperature is that quantity T that, in the right regime and up to a constant, is directly proportional to P and V. More precisely, it obeys the relation $PV = kT$.

- *Finding the realiser:* The statistical mechanical quantity $\langle E_{kin}\rangle$ denoting average kinetic energy that, in the right regime and up to a constant, is directly proportional to P and V. More precisely, it obeys the relation $PV = (2/3)\langle E_{kin}\rangle$.

Hence we infer that T is functionally realised by $\langle E_{kin}\rangle$ in the relevant context, since $\langle E_{kin}\rangle$ can play the functional (nomic) role of T. Reduction is therefore achieved by showing that we can understand higher-level phenomena involving temperature in terms of the lower-level theory, in virtue of the fact that the lower-level quantity realises the appropriate upper-level behaviour in the right regime.

[25] Butterfield (2011), van Riel and Van Gulick (2019), Dizadji-Bahmani (2021), Batterman (2023).

We also point out that, within the functional reductionist model, we can derive the Boyle-Charles law from the statistical mechanical law as a result of the functional reduction. Given (i) the statistical mechanical law $PV = (2/3)\langle E_{kin}\rangle$; (ii) that kinetic energy plays the role of temperature (up to a difference in the constant); and (iii) functionalism about temperature, we can substitute temperature for kinetic energy in the statistical mechanical equation and obtain $PV = kT$. This last step is not an essential part of the functional reductionist account. However, we can distinguish between two conceptions of functional reduction. The first is simply *'core functional reduction'* and amounts to the bare two-step approach presented above. The second, *'expanded functional reduction'*, denotes a more specific functional reductionist account that includes the law-derivation step. Although this distinction is not standard, the classification just presented will be useful in clarifying the relationship between Nagelian and functional reduction later.

6.2 How are Nagelian and Functional Reduction Related?

Within the literature, it is either maintained that Nagelian and functional reduction are distinct and alternative approaches or that functional reduction is a subtype of Nagelian reduction. We briefly review here the debate concerning the relationship between the two accounts, which is particularly centred around bridge laws.

First of all, Kim (1998, 2005) considers functional reduction as a completely distinct alternative to Nagelian reduction. He claims that, as opposed to Nagelian reduction, the functional model delivers reduction without appealing to bridge laws, a concept that he found problematic in the first place. Kim thus considers functional reduction as an alternative to Nagelian reduction, which is focused on showing how quantities across different domains can realise the same roles and thus be reduced one to another. In this picture, reduction is not involved with law-derivation and we do not need ex-post principles of connectability telling us how to relate different quantities, as the link is already provided by the match between the functional roles.

Marras (2002, 2005) and Fazekas (2009) argue instead that Kim's functional model of reduction requires bridge laws too, in the sense that it postulates connectability conditions between theoretical terms across theories in the same way as the Nagelian approach, and thus does not avoid the issues raised against bridge laws. More specifically, once we have functionally analysed the quantities at the different levels, the inter-theoretic link between the quantities ultimately constitutes a bridge law of the Nagelian kind. However, they still regard the two accounts as alternative approaches to reduction.

Yet another point of view is expressed by Morris (2020) and Butterfield and

Gomes (2022), who argue that, even if bridge laws have a place in the functional model as well, they are not added to the account in the same way as in the standard Nagelian account, and are thus less problematic.[26] Most notably, Butterfield and Gomes, following Lewis (1970, 1972), claim that functional reduction is a special form of Nagelian reduction that improves the standard version of the latter view, as bridge laws – in form of identities – can be deduced without the need to postulate them as extra principles. This is the position mentioned in §4.3.

What matters for the present topic is that two kinds of stances on the relationship between Nagelian and functional reduction can be extracted:

A. Functional reduction is a different approach than Nagelian reduction, that either (i) does not require bridge laws, differently from Nagelian reduction (Kim (1998, 2005)); or (ii) does require some kind of bridge laws (Marras (2002, 2005), Fazekas (2009), Morris (2020)).

B. Functional reduction is a kind of Nagelian reduction that employs bridge laws (in form of identities) that are added to the account in a special way (Lewis (1970), Butterfield and Gomes (2023)).

Hence it is maintained that either (A) Nagelian reduction and functional reduction are disjoint alternatives, i.e. $Fr \cap Nr = \emptyset$, or that (B) functional reduction is a particular form of Nagelian reduction, where the latter is a broader and more general framework, i.e. $Fr \subset Nr$.

We argue for a different view. If functionalism needs to be implemented within Nagelian reduction to solve the justification gap problem, then both options (A) and (B) are misguided. In fact, Nagelian reduction would turn out to be a subtype of functional reduction, i.e. $Nr \subset Fr$. We note that this is a conditional thesis. As stressed earlier, we do not claim in this paper that it is necessary to adopt functionalism to solve the issue. However, as long as functionalism is the only available solution, there is strong support for the conclusion that Nagelian reduction is indeed a special kind of functional reduction.

6.3 Nagelian as a Subtype of Functional

We show that both (A) and (B) are incorrect if functionalism about theoretical quantities is to be included within Nagelian reduction by default. Consider the reduction of Boyle-Charles law within the standard Nagelian model. Once we compare equation (1) and equation (2), if we postulate a bridge law connecting temperature and

[26] Cf. also Esfeld and Sachse (2007).

mean kinetic energy, we can use that to replace mean kinetic energy with temperature, thereby showing that (1) can be recovered from (2). However, if we assume functionalism about theoretical quantities, the condition of connectability is entailed by functionalism about theoretical quantities itself. Given the thesis of functionalism about theoretical quantities, the upper-level and the lower-level quantities would be understood in terms of their nomic roles in the context, and given the match between the roles as represented by equations (1) and (2), that same thesis would *automatically* deliver the condition of connectability without the need to add a bridge law as an *additional postulate*. Hence Nagelian bridge laws would become functionally induced by default, rendering Nagelian reduction equivalent to expanded functional reduction, and a subtype of core functional reduction. Consider how this conclusion affects views (A) and (B).

Philosophers in the (A) camp take functional and Nagelian reduction as distinct and competing accounts. Given the discussion in this section, and assuming functionalism about theoretical quantities, they should revise their position and maintain that Nagelian reduction is a special type of functional reduction in which we start from the two-step functionalist process to obtain the required conditions of connectability and then derive the laws of the reduced theory from the reducing one. Nagelian reduction thus boils down to expanded functional reduction. This is why our approach diverges crucially from Kim's (2005). Kim believes that the explanatory gap invalidates the Nagelian approach and this urges the formulation of a new alternative approach, i.e. functional reduction. Instead, we conclude that Nagelian reduction should be combined with functionalism and considered as a more specific version of functional reduction focused on law derivation.

Philosophers in the (B) strand take Nagelian reduction as a broader account than functional reduction, where 'functional reduction' refers in their view to a special kind of expanded functional reduction in which the quantities at the different levels are functionally *identified*. In fact, they maintain that we can exhaustively define the entities of each theory in functional terms, and thus draw identity relations between those quantities at different levels that instantiate the same functional description. This is Lewisian functional reduction. This view is regarded by them as a special form of the more general Nagelian account of reduction, since we are deriving the reduced-theory laws from the reducing-theory laws plus identity-based bridge laws that are specifically functionally obtained. That step is absent from the Nagelian approach, as they assume the bridge laws to be postulated in the standard Nagelian view. Given that Nagelian reduction is equivalent to expanded functional reduction (under the assumption of functionalism about theoretical quantities), which is a subtype of core functional reduction, philosophers on the (B) side should agree that

Nagelian reduction is a subtype of core functional reduction and maintain that their own conception of functional reduction is a special case of expanded functional reduction. Our conclusion diverges from Lewis, Butterfield, and Gomes's take because we disagree with how the Nagelian approach works in the first place.

We can represent this relationship more compactly. In place of the bare '$Nr \subset Fr$', we can state the relationship between Nagelian and functional reduction more precisely as follows, where CFr is core functional reduction, EFr is expanded functional reduction, LFr is Lewisian functional reduction, and INr is identity-based Nagelian reduction:

$$CFr \supset (EFr = Nr) \supset (LFr = INr).$$

7 Conclusion: Future Directions

We raised a crucial problem for the two main approaches to reduction between scientific theories. The challenge concerns the justification of mathematical mappings and bridge laws, which play a central role within those accounts. We argued that introducing functionalism within both views and turning them into forms of functional reduction solves the problem. Functionalism solves this pivotal issue and this urges us to rethink the relationship between Nagelian, mathematical, and functional reduction as it has been conceived in the literature. For instance, this means that Nagelian reduction should not be seen as the mainstream default approach to reduction, since plain Nagelian reduction is an incomplete account. Mathematical reduction is similarly incomplete without the functionalist component. This also undermines Knox's and Wallace's views on functionalism and reduction.

Looking forward, we present two research directions elicited by the arguments of this paper, extending beyond the implications considered so far. First, the conclusions of this paper can have far-fetched implications which need to be explored in future works. Consider the topic of scientific explanations. Models of explanation stand at the core of the debate on scientific reduction. For instance, Kim (2005) considered the role of explanations in the two accounts as what made functional reduction preferable to Nagelian reduction. Nagelian reduction is taken to embed a nomological-deductive account of explanation, whereas functional reduction is argued to embody a functionalist model of explanation. However, if the two accounts are not disjoint alternatives, but rather Nagelian reduction is a subtype of functional reduction, we would need to revise the role of explanation within both views.

Second, the two main contributions of this essay are the formulation of the justification gap problem and the proposal of a unified functionalist solution. However, we did not argue that functionalism is necessarily the only viable answer to the challenge. Are there alternative ways to bridge the crucial justification gap? This question remains open and has the potential to initiate a new research agenda within the debate on scientific reduction.

References

Albert, D. Z. (2015). *After physics*. Cambridge: Harvard University Press.

Balzer, W., C. U. Moulines, and J. D. Sneed (1987). *An architectonic for science: The structuralist program*, Volume 186. Springer Science & Business Media.

Bangu, S. (2011). On the role of bridge laws in intertheoretic relations. *Philosophy of Science 78*(5), 1108–1119.

Baron, S. (2022). Spacetime: Function and approximation. *Synthese 200*(2).

Batterman, R. (2023). Intertheory Relations in Physics. In E. N. Zalta and U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Winter 2023 ed.). Metaphysics Research Lab, Stanford University.

Beckermann, A. (1992). Supervenience, emergence, and reduction. In A. Beckermann, H. Flohr, and J. Kim (Eds.), *Emergence or Reduction?: Essays on the Prospects of Nonreductive Physicalism*, pp. 94–118. W. de Gruyter.

Bickle, J. (1996). New wave psychophysical reductionism and the methodological caveats. *Philosophy and Phenomenological Research 56*(1), 57–78.

Butterfield, J. (2011). Emergence, Reduction and Supervenience: A Varied Landscape. *Foundations of Physics 41*(6), 920–959.

Butterfield, J. and H. Gomes (2022). Geometrodynamics as functionalism about time. In C. Kiefer (Ed.), *From Quantum to Classical: Essays in Honour of H.-Dieter Zeh*, pp. 135–178. Springer International Publishing.

Butterfield, J. and H. Gomes (2023). Functionalism as a species of reduction. In C. Soto (Ed.), *Current Debates in Philosophy of Science: In Honor of Roberto Torretti*, pp. 123–200. Springer Verlag.

Carnap, v. R. (1958). Beobachtungssprache und theoretische sprache. *Dialectica 12*(3-4), 236–248.

Crowther, K. (2018). What is the point of reduction in science? *Erkenntnis*, 1–24.

Dennett, D. C. (1991). Real patterns. *The journal of Philosophy 88*(1), 27–51.

Dizadji-Bahmani, F. (2021). Nagelian reduction in physics. In *The Routledge Companion to Philosophy of Physics*, pp. 499–511. Routledge.

Dizadji-Bahmani, F., R. Frigg, and S. Hartmann (2010, 11). Who's afraid of nagelian reduction? *Erkenntnis 73*, 393–412.

Esfeld, M. and C. Sachse (2007). Theory reduction by means of functional sub-types. *International Studies in the Philosophy of Science 21*(1), 1–17.

Fazekas, P. (2009). Reconsidering the role of bridge laws in inter-theoretical reductions. *Erkenntnis 71*(3), 303–322.

Fletcher, S. C. (2019). On the reduction of general relativity to newtonian gravitation. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 68*, 1–15.

Fodor, J. A. (1968). *Psychological Explanation: An Introduction to the Philosophy of Psychology*. Ny: Random House.

Huggett, N. and C. Wüthrich (2013). Emergent spacetime and empirical (in)coherence. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 44*(3), 276–285.

Huggett, N. and C. Wüthrich (2021). Out of nowhere: The emergence of spacetime in quantum theories of gravity. *http://philsci-archive.pitt.edu/18612/*.

Kim, J. (1998). *Mind in a Physical World: An Essay on the Mind–Body Problem and Mental Causation*. MIT Press.

Kim, J. (2005). *Physicalism, or something near enough*. Princeton University Press.

Kim, J. (2008). Reduction and reductive explanation : Is one possible without the other? In J. Hohwy and J. Kallestrup (Eds.), *Being Reduced: New Essays on Reduction, Explanation, and Causation*. Oxford University Press.

Klein, C. (2009). Reduction without reductionism: A defence of nagel on connectability. *Philosophical Quarterly 59*(234), 39–53.

Knox, E. (2014). Spacetime structuralism or spacetime functionalism? *manuscript*.

Knox, E. (2019, 8). Physical relativity from a functionalist perspective. *Studies in History and Philosophy of Science Part B - Studies in History and Philosophy of Modern Physics 67*, 118–124.

Knox, E. and D. Wallace (2023). Functionalism fit for physics. *manuscript. https://philsci-archive.pitt.edu/22655/*.

Ladyman, J. and D. Ross (2007). *Every Thing Must Go*. Oxford University Press.

Lam, V. and C. Wüthrich (2018). Spacetime is as spacetime does. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 64*, 39–51.

Lam, V. and C. Wüthrich (2020). Spacetime functionalism from a realist perspective. *Synthese*.

Lewis, D. (1970). How to Define Theoretical Terms. *The Journal of Philosophy 67*(13), 427.

Lewis, D. (1972). Psychophysical and theoretical identifications. *Australasian Journal of Philosophy 50*(3), 249–258.

Lorenzetti, L. (2022). Functionalising the wavefunction. *Studies in History and Philosophy of Science Part A 96*, 141–153.

Lorenzetti, L. (2023). Functionalism, reductionism, and levels of reality. *Philosophy of Science 90*(4), 922–936.

Lorenzetti, L. (2024). Two forms of functional reductionism in physics. *Synthese 203*(2), 1–27.

Marras, A. (2002). Kim on reduction. *Erkenntnis 57*(2), 231–257.

Marras, A. (2005). Consciousness and reduction. *British Journal for the Philosophy of Science 56*(2).

Morris, K. (2020). Does functional reduction need bridge laws? a response to marras. *The British Journal for the Philosophy of Science*.

Nagel, E. (1961). *The Structure of Science: Problems in the Logic of Scientific Explanation*. New York, NY, USA: Harcourt, Brace & World.

Nagel, E. (1970). Issues in the logic of reductive explanations. In H. E. Kiefer and M. K. Munitz (Eds.), *Mind, Science, and History*, pp. 117–37. Albany, NY: SUNY Press.

Nickles, T. (1973). Two concepts of intertheoretic reduction. *The Journal of Philosophy 70*(7), 181–201.

Palacios, P. (2022). *Emergence and Reduction in Physics*. New York, NY: Cambridge University Press.

Palacios, P. (2023). Intertheoretic reduction in physics beyond the nagelian model. In *Current Debates in Philosophy of Science: In Honor of Roberto Torretti*, pp. 201–225. Springer.

Palacios, P. (2024). Intertheory Relations in Physics. In E. N. Zalta and U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Spring 2024 ed.). Metaphysics Research Lab, Stanford University.

Putnam, H. (1967). Psychological predicates. In W. H. Capitan and D. D. Merrill (Eds.), *Art, Mind, and Religion*, pp. 37–48. University of Pittsburgh Press.

Read, J. (2018). Functional gravitational energy. *British Journal for the Philosophy of Science 71*(1), 205–232.

Roberts, B. W. (2022). *Reversing the Arrow of Time*. Cambridge University Press.

Robertson, K. (2022). In search of the holy grail: How to reduce the second law of thermodynamics. *The British Journal for the Philosophy of Science 73*(4), 987–1020.

Rosaler, J. (2015). Local reduction in physics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 50*, 54–69.

Rosaler, J. (2019). Reduction as an a posteriori relation. *British Journal for the Philosophy of Science 70*(1), 269–299.

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), 137–147.

Schaffner, K. F. (2012). Ernest nagel and reduction. *The Journal of Philosophy 109*(8/9), 534–565.

Suppes, P. (1967). What is a scientitic theory? *Philosophy of science today*, 55–67.

Van Fraassen, B. (1980). *The scientific image.* Oxford University Press.

van Riel, R. (2011). Nagelian reduction beyond the nagel model. *Philosophy of Science 78*(3), 353–375.

van Riel, R. and R. Van Gulick (2019). Scientific Reduction. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2019 ed.). Metaphysics Research Lab, Stanford University.

Wallace, D. (2022). Stating structural realism: Mathematics-first approaches to physics and metaphysics. *Philosophical Perspectives 36*(1), 345–378.