Nagelian Reduction in Cognitive Science: Philosophy Of Mind and Bridge Law

Department of Philosophy, Logic and Scientific Method

MSc Philosophy of science
Dissertation

Yichen Liu

Supervisor
Dr Marius Backmann

Word Count: 8259

Abstract

This paper attempts a defence of Nagelian reduction in the context of cognitive science. Mind-brain reductionism has been one of the classic issues in philosophy of mind. Instead of a philosophical-theoretical approach that has been traditionally adopted in the literature, I adopt a hybrid approach to the examination of reductionism in this field, by giving a case study that combines empirical research outcomes in neuroscience and classic theoretical debates in philosophy of mind. This paper will consider the case of mental imagery debate regarding contemporary criticisms and defences of Nagelian reduction, the goal will be to defend Nagel's model (Nagel 1961, 1970), by arguing that the Nagelian model is a descriptive adequate account that successfully offers valid explanations to certain phenomena in cognitive science.

Table of Contents

1 Overview

1.1 Introduction

The goal of this paper will be to offer a defence of a moderated model of Nagelian reduction within the domain of cognitive science. It will focus on the descriptive adequacy of the model. I will argue that Nagel's model offers successful explanations and clears out the confusion in the case study of the mental imagery debate. Although the original Nagelian model is indeed flawed, it still potentially has strong explanatory power. With some moderations, it produces valid explanations to phenomena in empirical science, and is indeed what some scientists utilise in their research practices.

This first section offers a general plan of this paper and a brief literature review for background knowledge. In the second section, I will introduce Nagel's original account of his reductive model (Nagel, 1961). I will consider criticisms and defences from both classic literature and contemporary reinterpretation. The third section will be an attempt to apply accommodated model of Nagelian reduction to modern cognitive science. I will then study the case of mental imagery debate and argue that the application of the Nagelian model gives an adequate explanation to the empirical phenomenon. In the last section, I will consider whether my argument is tenable against criticisms, both old and new.

1.2 Background and Thesis Overview

As mentioned above, this paper addresses the application of Nagelian reduction in cognitive science. As a rapidly developing, highly interdisciplinary field, cognitive science addresses theories in multiple scientific domains, typically including psychology, neuroscience, and philosophy. Fascinating subdisciplines of cognitive science such as moral psychology, psycholinguistics, and visualising handle intertheoretical issues due to their integrative nature, they unavoidably encounter one of the central topics in philosophy of science: intertheoretical reduction.

Roughly speaking, scientific reduction is the idea that if a theory in a secondary science reduces to a more basic scientific theory, it means that the reduced theory can be "brought back" to the reducing theory (van Riel et al, 2019) Frequently adopted examples include the reduction of molecular biology by chemistry, thermodynamics by statistical mechanics, as well as controversial ones in social sciences such as the reduction of macroeconomics by microeconomics. In the reductionist debate, the Nagelian model has drawn long-lasting discussions since Ernest Nagel presented this account. (Nagel, 1961, 1974) It received a variety of criticisms in the past few decades, however, contemporary attempts to defend the model offers strong reasons to rethink the validity of Nagelian reduction, considering new adjustments in accordance with the progress in certain growing scientific disciplines such as cognitive science.

In the specific domain of cognitive science, the reductive relation between mental states and physical states is within the concern. In fact, the Mind-body problem is an ancient discourse, starting from Descartes' substance dualism in the 17^{th} century, the relation between the mental and the physical has been theorised from different approaches, among which, reductionist accounts of the mind attempt to reduce psychological properties to more basic properties. The advocators vary in their approaches to the reduction of mental properties, identity theorists such as J.J.C. Smart hold a physicalist position arguing that mental states are brain states (Smart, 1959), while computational theorists may defend a kind of functionalist reduction (van Riel et al, 2019). In this paper, I shall have a closer look at one of the most discussed computational theories of mind, namely Jerry Fodor's the Language of Thought hypothesis (Fodor, 1975, Moreover, I will consider Putnam and Fodor's multiple realisability objection against identity theorist reduction (Putam, 1967, 1975, Fodor, 1968, 1975) I will offer a case study of mental imagery debate to examine Fodor's functionalist reduction and multiple realisability objection. This will set the stage for defending Nagelian reduction.

The reason why reductionism in cognitive science is at the centre of our attention here is that, scientific reduction has shown a convincing ability to explain in fields such as physics and chemistry, it remains an irreplaceable tool in our understandings of intertheoretical relations. However, since its development in the 1970s, multiple realisability is thought to undermine reduction of the mind and threats the credibility of reductive models. Reconstructed and novel reductionist theories were put forward to replace classic accounts of type-identity reduction, including Fodor's token physicalism (Fodor, 1974) But I will argue in this paper that this theory fails in its attempt to undermine the concept of Nagelian bridge law.

2 The Nagelian Reduction Model

2.1 The Original Model

In *The structure of Science*, Nagel gives the following formal definition of reduction:

> A reduction is effected when the experimental laws of the secondary science (and if it has an adequate theory, its theory as well) are shown to be the logical consequences of the theoretical assumptions (inclusive of the coordinating definitions) of the primary science. (Nagel, 1961, 352)

According to Nagel, reduction is essentially an explanation of a theory by another. (Nagel, 1961, 338) More specifically, given the formal definition, it is a deductive-nomological explanation, meaning that by saying one theory reduces to another, it means that the laws in the reduced theory can be logically derived from the laws of the reducing theory. Considering the fact that it is often the case that the primary science and the secondary science do not contain the same set of theoretical vocabulary. Nagel suggests two formal conditions to take this concern into account, they are (1) condition of connectability and (2) condition of derivability.

The condition of connectability states that additional assumptions are postulated to assist reduction by connecting the theoretical terms of both theories. Later these assumptions are called bridge laws. The condition of derivability states that with bridge laws, all the laws in the reduced theory must be deductively inferred from the reducing theory. Nagel further explains that there are three possible kinds of these assumptions or bridge laws, they can be (1) logical connections, (2) conventions, or (3) factual or material statements. He uses the example of the reduction of classic thermodynamic to mechanics to show the necessity of bridge laws. In this example, the concept of temperature does not appear in kinetic theory, to explain temperature in terms of the mean kinetic energy of the molecules, a corresponding set of assumptions need to be postulated. Nagel argues that the nature of the assumptions, or bridge laws, in this example is contextual, it could be either coordinating definitions or material assumptions inferred from empirical data. However, the contextuality does not deny the necessity of the condition of connectability and the fact that the nature of the bridge laws falls into the above three categories.

Nagel differentiates two types of reduction in his later work (Nagel, 1974, 912), homogeneous and inhomogeneous reductions. The former is a kind of straightforward reductive explanation involving reduced theory and reducing theory that has the same theoretic terms in their laws. He uses the example of the reduction of Galileo's and Keplerian laws by Newtonian gravitational theory to illustrate this kind of reduction. In homogeneous reduction as such, theoretical terms are shared among the laws, it is thus possible and clear that the reduced laws are logically derived from reducing ones. Nagel thinks this kind of reduction is rather unproblematic. On the other hand, an inhomogeneous reduction is less comprehensive due to differences in theoretical concepts. In this paper, he further specifies the definition of bridge laws (or correspondence rules) and their different categories. He writes:

> Correspondence rules formulate empirical hypotheses --- hypotheses that state certain relations of dependence between things mentioned in the reduced and reducing theories. (Nagel, 1974, 919)

Although bridge laws come in different types and some are not testable through observable means, Nagel insists that they are empirically assessable. Here I shall summarise the nature of the original account of bridge laws as the following rules:

1. They are necessary for inhomogeneous reduction to fulfil the connectability condition.
2. They are essentially empirical claims.
3. They are of different forms and functions, they can be logical, conventional, or factual. They are:

   (1) "Empirical hypotheses concerning the extensions of the predicates mentioned in these correspondence rules --- that is, concerning the classes of individual things or processes designated by those predicates." (Nagel, 1974, 920)

   Or

   (2) Rules that "establish analogous identification between classes of individuals or "entities" designated by different predicates." (ibid)

The reduction of psychology in cognitive science is a typical kind of inhomogeneous reduction, since psychological terms are disciplinarily specific and normally do not appear in the potential reducing theories such as neuroscience. Therefore, in this paper, I will leave homogeneous reduction aside and consider only inhomogeneous reductions and the role of bridge laws.

2.2 Criticisms against the original account

After the model was put forward, it has received numerous criticisms. This paper cannot cover a complete evaluation of all existing criticisms; thus, I shall focus on the ones that are the most relevant to our current discussion of reductionism in cognitive science.

I categorise those criticisms into two kinds, first is a set of *methodological issues* regarding the formal definitions within the model; while the second includes several *applicational issues* regarding the pragmatic problems encountered (or said to have appeared) in the actual scientific practice of reduction. This categorisation aims at clearing out the confusion between the quality of the model and the instrumentality of the model. In a simple and analogous sense, a philosophical model is essentially a cognitive tool to assist our understanding. Same as any other kind of tool, say garden tools, a model can be of bad quality due to its false assumption or logical fallacy, or it can be a perfectly logically valid postulation but without any practicality, i.e., it cannot be applied to anything. In this case, a reductive model can be coherent, but cannot explain any scientific phenomena or is simply not adopted by any scientist in their practice. If we set the threshold higher, a good reductive model should be able to cover as many domains in science as possible, if a model can only explain the reduction in, for example, physics, but fails to explain the reductions in social sciences, we would only be able to call it a model of physical reduction instead of scientific reduction. To this paper's concern, I shall only consider the application of the Nagelian model to cognitive science. First of all, I will specify the issues within our consideration:

2.2.1 Methodological issues

The problems of the formal design of the model mainly lie in the theoretical assumptions Nagel endorses. The Nagelian model inherited the issues of its assumptions, which include the deductive-nomological account of scientific explanation and a logical positivist view of laws. The former assumes the logical derivability of theories, and the latter considers laws in terms of first-order logic, which are all problematic in their regards. Consequently, as the Nagelian model is said to be "a logical relation between theories, where a theory is understood as a system of statements, containing laws, formulated in a first-order formal language" (Brigandt, 2017), it is limited in its scale to only be able to explain reductions in highly logically formalised disciplines that embody mainly mathematic equations, such as physics.

Another problem of the model concerns the incommensurability of the meaning of the theoretical terms, which was seen in Feyerabend and Kuhn. (Feyerabend, 1963, Kuhn, 1962) The incommensurability argument holds a holistic theory of meaning, stating that the meanings of the terms within a theory are interdependent, therefore there is not a common criterion to compare meanings intertheoretically. Following this assumption of the semantic of theoretic terms, Feyerabend's criticism of the Nagelian model rises a problem from a historical perspective of scientific progress. He argues that Nagelian reduction relies on a consistency condition and a condition of meaning invariance that hinder the acceptance of novel scientific theories. This is because if reduction were indeed logical derivation, these two conditions imply that meanings of the terms would be consistent within theories and invariant across theories, due to the deductive nature of explanation. It leads to the fact that the acceptance of a theory rather than another only relies on arbitrary criteria such as the time order of their appearance, if they include incommensurable vocabularies. If a paradigm shift happens, a novel theory would be wholly replacing the old one without amplifying the body of scientific knowledge. This, to Feyerabend, is a problematic position to hold. (Curd et al., 2013,1010)

2.2.2 Applicational issues

The most cited examples of reduction, as mentioned above, are ones from physics, such as the reduction of thermodynamic by statistical mechanics and Newtonian mechanics to relativity theory. It is less clear how reduction works in other empirical sciences and social sciences, including cognitive science. In these cases, theoretical terms are always radically different in the reduced and reducing theories, therefore the bridge laws can be hard to identify and make sense of.

Issues regarding bridge laws are the dominant concerns of the applicational aspect of the Nagelian reductive model. The notably strong objection from multiple realisability has been considered to undermine the reduction of psychological properties to neurophysiological properties. Other applicational objections worth discussing were nicely summarised by Dizadiji-Bahmani et al., whose defence of the Nagelian model I shall also consider in the following sections (Dizadiji-Bahmani, Frigg, Hartmann, 2010).

Objections from identifying bridge laws

I shall address three issues here regarding bridge laws. The first one concerns the nature of these laws in actual scientific practice. Although Nagel has put forward three different types of bridge laws, namely logical, conventional, and factual laws as listed above, it is said to be criticised because "we do not use bridge-laws when discovering for example that parts of chemistry reduce to atomic physics. Rather, we infer bridge-laws from the discovery of relevant similarities between the two theories." (van Riel et al., 2019) This objection is clearly an empirical assumption about the ways how scientists and philosophers of science approach a reduction in their practice. It asks a question about application; thus, it is empirically examinable through looking at historical cases of scientific research and examine how scientists identify bridge laws.

A second issue is that, since it is argued that bridge laws are *ad hoc* in the sense that they are referred from discoveries, they do not exist independently of the reduced and reducing theory, therefore they are not independently testable without involving a particular context of a reduction.

The third issue with bridge laws is regarding its content. Dizadiji-Bahmani et al. point out that "There is a question about what kind of statements bridge laws are." (Dizadiji-Bahmani, 2010) Furthermore, they argue that it leads to the problem of multiple realisability, which I shall elaborate on next.

Objection from multiple realisability

The multiple realisability argument was initially proposed by Hilary Putnam to reject identity theory in philosophy of mind (Putnam, 1967). The identity theorists argue that every particular mental state is identical to a specific brain state. The example of pain is often used in this discussion, mind-brain identity theorists believe that the feeling of pain is identical to C-fibres firing on a neurological level. The multiple realisability argument rejects this identity by arguing that pain, as a psychological kind, can be realised by different physical kinds, hence pain is multiply realisable. Human brains and, for instance, cat brains differ in their biological structure, while both species can feel pain. Similarly, behavioural evidence shows that octopuses, although biologically wired completely alien compared to mammals, can feel both physical and psychological pains. Given these examples, multiple realisability theorists make a further inference that it may be possible that machines and Martians can be in the same state of pain as well.

This argument threatens the reductionism of psychology because a successful reduction is supposed to show that the reduced theory is nothing over and above the reducing theory, namely that all the properties of the mind should be able to be explained in terms of properties of the brain. It seems that, at least a mind-brain identity theory is blocked by argument from multiple realisability.

Furthermore, Fodor (1975) suggests that multiple realisations not only undermine the reduction within cognitive science, but they also reject reductionism as a whole, particularly the Nagelian model, because it endangers the status of bridge laws. If a higher-level property is multiply realisable, it is correspondent with multiple lower-level properties. A bridge law connecting this higher-level property with its multiple lower-level realisers would be a disjunction that covers all the possible cases of realisation. A further concern is that the number of realisers has an open-ended possibility. The issue goes down to a fundamental question of whether laws can be disjunctions and whether we want to allow an open-ended disjunctive law.

2.3 Modification of the Original Nagelian Model

In the above sections, I identified, summarised, and explained some frequently discussed issues of the Nagelian reductive model. It did not aim to be an exhaustive list but rather a selective one that will be considered in the following sections.

Undoubtedly, the original model itself is flawed and needs modification. In this section, I will address them one by one and consider the suggested modifications in the literature. The goal of this section will be to offer an account of an improved version of the Nagelian model, to examine whether it achieves its aim in the application to the case of the mental imagery debate. However, we would not wish to undermine the essence of the model by modifying it, otherwise, it would be an objection instead of modification. Therefore, I shall point out explicitly that there are three main goals that I hope to achieve in this paper, each relates to one of the major characters, or cores, of the Nagelian model. They are:

(1) A successful reduction essentially offers an explanation. The explanatory power of a model is an important criterion of its evaluation.

(2) Bridge laws are necessary for inhomogeneous reduction. They play a crucial role in the process of reduction.

(3) Derivability is essential to reduction, which means, the conclusion of the reduced theory is a logical consequence of the premises in the reducing theory.

With these three goals beard in mind, I will now proceed to construct a reconsidered model of Nagelian reduction, progressing from its methodological issues to its applicational issues.

I shall start from the theoretical assumptions adopted in the Nagelian model. Recall the interpretation of the model quoted in 2.2.1: “a logical relation between theories, where a theory is understood as a system of statements, containing laws, formulated in a first-order formal language” (Brigandt, 2017) This has been challenged on two grounds: (1) Strict deductive logically derivation, and (2) laws of first-order logic. Nagel responds to the first challenge by loosening the requirement for a strict logical derivation, he acknowledges that “in actual scientific practice, the derivation of laws from theories usually involves simplifications and approximations of various kinds.” (Nagel, 1974, 915) Therefore, he suggests it is correct to leave space for approximations in the derivational process.

On the second issue, Dizadiji-Bahmani and colleagues argue that Nagel did not endorse such a view for his reductive model that requires laws only to be in the form of first-order logic. Moreover, regarding the examples they use to defend the Nagelian model, they write: “Neither did we present a first-order formulation of the theories, nor did we even mention a bifurcation of the vocabulary into theoretical and observational terms. Where first-order logic is too weak, we can replace it with any formal system that is strong enough to do what we need it to do. The bifurcation of the vocabulary plays no role at all.” (Dizadiji-Bahmani, 2010)

Given these analyses, it seems unproblematic to make the first two adjustments of the model, one is to allow approximations in the derivation of the laws; another is to drop the requirement of first-order logic and allow other forms of formal systems, for example, probabilistic theories. By loosening these requirements, the problem with incommensurability is hopefully solved as well. The incommensurability argument points out that due to the nature of deductive reasoning and a holistic view of meaning, the Nagelian model does not allow progress in science and the accumulation of knowledge. However, once approximations are allowed, new knowledge can come in and adjustments can be made through derivations, it would not be problematic to explain scientific progress.

After considering the theoretical issues, we shall look at whether with the modifications, the Nagelian model can be applied to real scientific practice. Before introducing the case study, some comments regarding the bridge laws are worth mentioning. I will consider two reasonable suggestions of the bridge laws: the first one argues that bridge laws are factual; the second one argues that bridge laws are probabilistic.

The factual argument from Dizadiji-Bahmani et al. (2010) dismisses the first two kinds of possible accounts of bridge laws that Nagel initially provides, namely meaning identification and conventions. They argue that meaning equivalence is irrelevant to reduction if one holds a holistic view of meaning: this is an internalist account of meaning that neglects the properties that the terms actually *refer to*, which is an external consideration. They point out that "what matters is whether the properties that the terms in the bridge laws refer to stand in a relevant relation to each other." (ibid) Therefore bridge laws cannot be meaning equivalence. Also, they cannot be arbitrary conventions either, because bridge laws can be true or false about the relations between the reduced and reducing theories. The only option left is a factual account of empirically testable bridge laws.

The testability requirement for bridge laws brings back the question of their contextual dependence, which, as explained above, questions the independent testability of the bridge laws. This problem will be addressed in the case study, because the best counterargument against this objection is to offer an existing case in scientific practice to show that bridge laws can exist independently.

Although Dizadiji-Bahmani et al. give reasons why bridge laws are factual, they did not specify what type of factual statements they could be, they suggest several candidates that "express mere correlations (or Humean regularities), nomic connections involving certain necessities, identity statements, or yet other metaphysical relations" (ibid) It remains an open question what a bridge law can be. In fact, it is logically possible that factual bridge laws can be of multiple kinds, but here I shall consider a probabilistic account offered by Marco J. Nathan and Guillermo Del Pinal (Nathan et al., 2014), which was designed as a replacement for reductive bridge laws. I argue that their account of probabilistically associative bridge laws is better accommodated into a reductive model to make better sense.

This probabilistic account is considered in the context of cognitive science, the authors argue that the success in interdisciplinary research has demonstrated a mapping between neural data and cognitive hypotheses. However, the nature of the bridge laws that connect these two cannot be statements of type-identities or token identities, more specifically because they cannot be reductive, because of the failure of reverse inferences, meaning that we cannot infer neural activation from cognitive states: they argue that "virtually no brain region can be mapped onto cognitive functions via a single bridge law; rather, each brain region is covered by multiple bridge laws which associate it with a variety of cognitive function from the activation of a neural region, the inference falls short of absolute certainty.” (ibid, 646)

The associative bridge laws they suggest “are probabilistic and context-sensitive relations that do not identify their relata, either at the type-level or at the token-level”(ibid, 645) These bridge laws are designed to capture inductive inferences between cognitive states and neural activations, because the mapping between them is not strictly one-to-one determined, but rather probabilistically associated. Therefore, the bridge laws can be evaluated through conditional probability by applying Bayes theorem directly.

Nathan et al. offer an insightful account of probabilistic bridge laws; however, I think bridge laws without reductive nature fail to capture the metaphysical relations between some certain kind of higher-level and lower-level properties. Recall the three features that I hope to defend in this paper, reductive relations are essentially logical derivations, that are mapped by bridge laws, that essentially offer explanations to phenomena. Probabilistic mapping alone cannot cover the essences of reductive relation, as Nathan et al. admit themselves that associative bridge laws are “metaphysically uncommitted” (ibid, 650) that they do not express identities, that they “are often not even stable across individuals” and “can only be used reliably in specific experimental conditions.” (ibid) If these bridge laws are indeed mere probabilistic mappings and presuppose no metaphysical relations, there seems no reason to dub them “bridge laws” instead of just calling them probabilistic mappings. Reduction necessarily requires at least some kind of logical entailment. On the other hand, if the reduced theory is logically entailed by the reducing theory, they are probabilistically mapped as well --- this explains why Nathan et al.’s associative laws are capable of explaining classic cases of reduction, i.e. thermodynamic and statistical mechanics but not vice versa: associative laws do not cover the whole metaphysical relations between thermodynamic and statistical mechanics, because they are more than just associative mapping.

These authors may as well argue that the relation between cognitive properties and neural properties is not reductive as a further step, but it is misguided to think that associative laws can undermine or replace reductive laws in scientific practices in general, neither is it a rejection against reductive models. On the other hand, reductive bridge laws can be probabilistic in nature as well once we allow approximations in the reductive process. Associative laws have weaker explanatory power compared to reduction in the sense that there is no metaphysical constraint between two theories, thus the relation is associative and possibly arbitrary, which would not be able to explain law-like relations.

To sum up, I have argued in line with contemporary defenders of the Nagelian model that bridge laws are factual and possibly probabilistic, and therefore empirically testable. Factual bridge laws combined with a moderated account of the reductive model can potentially better apply to scientific practice. In the next section, I shall introduce the case of the imagery debate to examine whether this adjusted model achieves this goal.

3 The Imagery Debate and Reduction

3.1 The false debate

Mental imagery is a kind of quasi-perceptual experience we have when we imagine visual scenarios without simultaneously receiving visual stimuli from the physical world. If you are asked to close your eyes and memorise or imagine what your childhood home looked like, you may "see" it vividly in your head with your "mind's eye". The nature of mental imagery caused an ongoing debate since the 1960s, cognitive scientists split into two groups in the 1990s, one of pictorialists and another of descriptionalists. The pictorialists believe that mental imagery possesses the spatial properties that pictures do, while the descriptionalists hold that mental imagery is language-like and are represented in the mind through manipulating mental language following grammar-like rules.

I will argue in this section that the original imagery debate in cognitive science is a false dilemma. These two positions talk across each other on a conceptual level, however, the mistake they made happens to demonstrate the crucial role played by bridge laws in scientific reduction. I shall start by introducing an experiment of mental rotation, which is considered a central and classic experiment in the imagery debate. I will then look at arguments and evidence from both sides.

In the initial mental rotation experiment by Shepard et al. in 1971, the subjects are asked to look at pairs of three-dimensional objects (figure 1). The pairs are either identical objects rotated by different angles or both rotated and mirrored. The task for the subjects is to tell whether they are rotated or mirrored and rotated, the subjects' reaction time and the rotated angle are measured. The result of the experiment shows that there is a linear correlation between the responding time of the subject and the angles rotated of the object. The pictorialists argue that this result supports their hypotheses because the best explanation for the linear correlation is that the subjects rotate the objects in their mind in the same way that they would rotate a physical object in real-world: the large the angles are that they need to rotate, the longer time they need to do so.

The descriptionalists argue against this account by pointing out there is no “mind’s eye” to observe the picture, and when we observe the brain (through brain scan), we would not be able to find a picture within that has the spatial features required to explain mental rotation. Amy Kind points out that “In order to avoid dualism, the, the pictorialist seems forced to suppose that these pictures I the head are located in the brain.” (Kind, 2020) On the other hand, a representation does not need to be in the form of the thing it represents, for example, arbitrary symbols may represent physical objects in the form of language, graphics may represent a statistical model. The descriptionalists do not suggest that the representation of pictures is in the form of natural language, but rather could be in the form of computational binary code or something similar. The mind manipulates these mental languages like computers manipulate binary language to represent pictures.

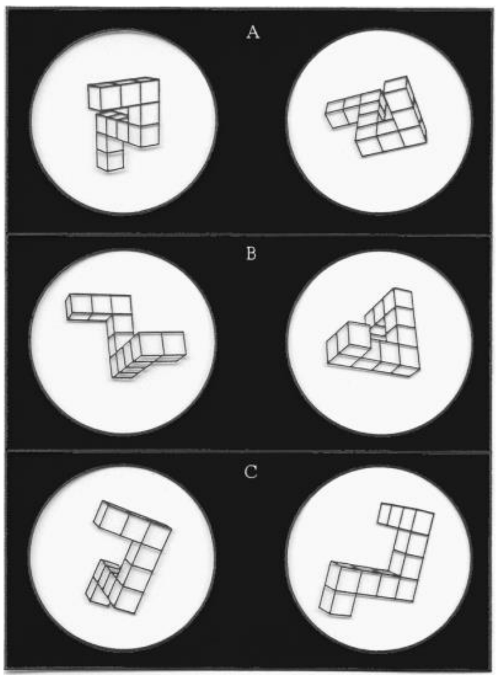

Figure 1

These two positions seem to create a dilemma on the appearances; however, I argue that they are not necessarily opposite viewpoints. This is because both the pictorialists and the descriptionalists are fundamentally reductionists, they adopted different additional assumptions in their reductive processes, and these assumptions are bridge laws that embedded in their position in philosophy of mind. However, I do not wish to defend either of these positions --- they could be both wrong. Instead, I hope to show that it is possible that scientists and philosophers identify false bridge laws and subsequently create confusion in their discourse, this shall support the importance of the functionality of bridge laws and Nagelian reduction as a whole.

3.2 Two reductionist positions

3.2.1 Pictorialist type-identity reduction

In the mental imagery debate, pictorialists argue that mental imagery is picture-like and have spatial properties, the evidence they seek is the location of the picture in the physical brain: this seems to be an impossible task. Later pictorialists developed a quasi-pictorial theory to avoid having to find literal pictures in the brain (Kosslyn et al., 1977, 1978) But they insist on certain pictorial features of mental imagery, for example, mental imagery should cause the same visual experience as if there were physical stimuli to be perceived, consequently, the same neural activations should be seen in both scenarios. They proceed with the search in empirical neuroscience for evidence.

On the other hand, the descriptionalists dismiss the use of neuroscientific evidence, Zenon Pylyshyn writes: “I claim that when such questions as whether images are depictive or spatial are formulated more clearly, the evidence does not provide support of the picture-theory over a symbol-structure theory of mental imagery. Even if all the empirical claims were true, they do not warrant the conclusion that many people have drawn from them: that mental images are depictive or are displayed in some (possibly cortical) space.” (Pylyshyn, 2002) This is because the philosophers arguing against a pictorialist account of mental imagery are largely functionalists, or at least endorse some kind of computational theory of mind: they hold that the mind computes on a language-like system to produce representations like mental imagery.

The above analysis of the debate shows that both sides hold a kind of reductive account of mental imagery: which entails an assumption of the mind, this assumption plays the exact role of a bridge law. The former, the pictorialist, hold a kind of type-identity theory of the mind. In their attempts of relying on brain activation for evidence of a cognitive state, it shows that they believe demonstrating same regional activation of neurons of two cognitive process equals demonstrating these two cognitive states are identical, or mental imagery are picture-like: they assume that cognitive states are identical to brain states. The opponents of the pictorial theory are functionalists, among them, Fodor offers an explicit argument against type-identity theory through the argument of multiple realisability, which we shall discuss in the next section.

3.2.2 The Language of Thought Hypothesis and multiple realisability

In section 2.2.2, I have briefly mentioned the multiple realisability argument from Fodor, it claims that type-identity theory that attempts to reduce a natural kind to another encounters the problem of multiply realisable higher-level properties. Multiple bridge laws would be needed to connect a higher-level property and its lower-level realisers. He argues that a type-identity theory is essentially a disjunction of token-identities, which undermines the concept of kind, because it is an open question of what tokens a natural kind includes if it is multiple realisable and have an uncertain number of realisers.

Fodor thus concludes that a reduction of psychological properties to neurological properties is untenable due to its problematic bridge laws. Fodor argues that "The only psychological models of cognitive processes that seem even remotely plausible represent such processes as computational." (Fodor, 1975, 27) and "The available models of cognitive processes characterize them as fundamentally computational and hence presuppose a representational system in which the computations are carried out" (ibid, 99) He proposes a functionalist account of the mind, namely the language of thought hypothesis, which considers mental states as the computational roles within a system of mentalese, that is, a hypothesised language that the mind operates on, similar to natural language and computer language, it has linguistic features, namely semantics and syntaxes.

The attempt to characterise cognitive properties in terms of computational properties is itself an attempt of reduction. It is a reduction of mental states by its functional causal role played within our cognitive processes. Moreover, the language of thought hypothesis implicitly adopted a Nagelian model of reduction, an example from its position on the mental imagery can present this point, which I shall elaborate on.

To consider whether the functionalists are reductionists, I will consider if their reasoning fit into the three core features of Nagelian reduction that I hope to defend. More specifically, I will consider Fodor's account of mental imagery to identify its deductive derivation, its bridge laws, and its explanatory power.

Regarding the nature of Mental imagery, Fodor believe mental images are embedded in descriptive processes, he writes: "For insofar as mental images are constructed from descriptions, the descriptions can function to determine what the images are images of, and how their properties are to be interpreted." (ibid, 192) As he presupposes the existence of a representational system, he argues that "one of our psychological faculties functions to construct images which accord with descriptions. That is, we have access to a computational system which takes a description as input and gives, as output, an image of something that satisfies the description." (ibid) This account of mental imagery is a move of reducing the cognitive state of having a mental image to a functional state of a postulated computational system. Furthermore, it is an account of reduction in the Nagelian sense, because it tries to derive a cognitive state from computational states through a formal system: the language of thought. The language of thought *qua* language-like, have the logical features of natural languages. But unlike natural languages, the language of thought is an empirical hypothesis, because there is a factual state of how the mind actually works, it might function on a language of thought, or it might as well not.

Therefore, we have a very clear example of philosophers using factual assumptions as bridge laws to logically derive a secondary science (psychology) from a more basic scientific theory (computation), in order to give explanation of a mental phenomenon. The language of thought hypothesis is a bridge law that exists before its application to the mental imagery discussion, and it is independently testable from the mental imagery debate or any specific examples in cognitive science.

3.3 The wrong dilemma

As argued in the last two sections, in this imagery debate in cognitive science, we have two reductive accounts, namely a type-identity reduction of mental states to brain states, and a functionalist reduction of mental states to computational states.

I have analysed why Fodor's language of thought hypothesis constructed an account of Nagelian reduction. It is worth making a further comment to its opponent to make their position equally clear. The pictorialists, as I argued, are identity theorists of the mind who hold that mental states can be reduced to brain states. What they do in practice is to look at empirical data in neuroscience and try to derive what is going on in the subjects' minds: they infer from the linear correlation of the reaction time and the rotation angle that the subjects have images in their head to rotate spatially. However, correlation in neural data does not imply identity, there is a step missing in their reasoning per se. To reach the conclusion of identity through deductive reasoning, they take a step of abductive reasoning, namely the inference to the best explanation, which takes the identity theory as the best explanation of the perfect correlation. Consequently, the identity theory plays a role of a bridge law in this case.

Having explained these two reductive positions, Nagelian reductive model helps to reveal the true nature of the tension between the pictorialists and the descriptionalists in the mental imagery debate. It is fundamentally not only a debate about what mental imagery really is, but rather more of a debate between identity theorists and functionalists, who are both reductionists but hold different views of the mind. It would not make much sense to discuss the nature of a particular state of mind before philosophers and cognitive scientists figure out what the true nature of the mind is, because the assumptions of the mind play a major role in interpreting empirical data and experimental evidence, as seen in the imagery debate.

3.4 Summary

To summarise section 3, I have examined the two counter positions in the imagery debate in cognitive science. I argued that they are both cases of Nagelian reduction, they both show the features of the Nagelian model, they both endorse a logical derivation to explain the higher-level properties in terms of the lower-level properties; they both have bridge laws, and interestingly, from philosophy of mind; they both managed to give an explanation to the phenomena of the mind. It is important to notice that the success to offer an explanation does not require a successful explanation, in fact, that our current scientific theories still cannot guarantee the truth of the explanation they offer, philosophers of science must accommodate approximations and fallibility of scientific theories into their theories of scientific explanation. On the other hand, in this particular case, with the tool of Nagelian reduction, we were able to better make sense of scientists' reasonings and help to identify the core issues of a discussion like the imagery debate, which shows that the model has strong explanatory power.

4 Resolving the issues

In the case study of the mental imagery debate, I showed an example of Nagelian reduction in real scientific practice. In this section, I will analyse how this example gives a solution to the issues of the Nagelian model that have been listed previously.

4.1 Resolving the methodological issues

In the previous sections, I have identified issues within the theoretical assumptions of the Nagelian model and given moderation regarding these issues. The theoretical assumptions I mainly dealt with were requirements for (1) deductive inferences, and (2) first-order logic. I have argued (1) in line with later Nagel that the logical derivation of reduction has to allow approximations, and (2) in line with Dizadiji-Bahmani et al. that the first-order logic is never necessary for Nagelian reduction, any formal system can do the job. Based on these suggestions, two adjustments have been made to loosen the two requirements.

As we have seen in the reasoning of the pictorialist reduction, to achieve a type-identity reduction, the pictorialists have to commit to an abductive inference that assumes the truth of their best available theory of mind. This move is embedded in the deductive reasoning as a premise, with the form as follows:

Premise 1. Empirical data of experimental results show that there is a linear correlation between the reaction time of the subjects' and the rotation angle.

Premise 2. Abductive reasoning is a valid form of logical inference, meaning that the most probable conclusion is one that is the simplest and likely, i.e. the inference to the best explanation.

Premise 3. The best explanation for the linear correlation is a type-identity theory of the mind

Conclusion 1. The mental state of mental imagery is identical to certain brain states.

Conclusion 2. The subjects rotate pictorial mental imagery in their mind that has spatial features.

In this argument, we see a clear structure of deductive reasoning that entails (1) auxiliary assumptions acting as bridge laws, and (2) non-deductive premises that is an approximation and probabilistic in nature. Given what I have argued above that bridge laws can be probabilistic and yet reductive, this account fits perfectly into a Nagelian framework in reduction.

On the other side, the descriptionalist position gives another good example of adopting other forms of formal system in reduction. One may argue that in the pictorialist account of mental imagery, the proponents derived laws of cognitive states (mental rotation of picture-like mental imagery) from behaviour, empirical, and experimental data, which is a mathematical process and rather close to a first-order logic reasoning. While the descriptionalists, who endorses theories like the language of thought hypothesis, postulated a new formal system that is based on computation and linguistic rules. This postulation functions as bridge laws connecting the reduced properties (mental states) to reducing properties (functional states). The functionalist reduction offers a real example in practice of how philosophers and cognitive scientists utilise the Nagelian reduction model as a reasoning tool and endorse the kind of moderation that I considered in this paper.

To sum up, the moderated Nagelian model with loosened requirements in its theoretical assumptions does offer explanations in actual scientific practices, these explanations are formed and have entered further long-lasting debates in cognitive science, which demonstrated the explanatory power of both sides. With the adjustments of theoretical assumptions, this model may have overcome the classic criticisms.

4.2 Resolving the applicational issues

In the previous discussion, I pointed out that this paper does not defend either of the theories in the mental imagery debate, or any account of philosophy of mind. This paper is rather an analysis of the descriptive adequacy of the Nagelian model, meaning that it hopes to demonstrate that the model captures what is actually happening in scientific practices when scientists perform an intertheoretical reduction. Therefore, it is crucial for the goal of this paper to look at whether the applicational issues of the model have been solved.

Regarding the nature of the bridge laws, on one hand, I have argued that they are factual and empirically testable; on the other hand, I have mentioned a few times that the reductionists in the imagery debate implicitly adopt the assumptions in philosophy of mind as auxiliary bridge laws to complete their deductive reasoning. It might seem counterintuitive that philosophy of mind theories can function as bridge laws, since philosophical theories are not normally considered empirical claims. However, unlike theories in political philosophy or thought experiments, the kind of philosophy in cognitive science (or other kinds of special science) often does make empirical claims that carry truth value. The computationalist claim of “the mind computes” is a proposition that can be true or false. And cognitive science itself is an empirical discipline to test whether the claim in philosophy of mind is true or false. Therefore, I argue it is unproblematic to think that philosophy of mind theories work as bridge laws in the reduction within cognitive science.

Consequently, as we have seen, in the case of mental imagery debate, cognitive scientists and philosophers do appeal to independent bridge laws to conduct reduction, instead of inferring bridge laws from existing discoveries of intertheoretical relations. However, these bridge laws may be hidden and unrecognised when the reduction happens, and only identified later through reflection, but this does not mean that they did not play the role in the reductive process. An example to illustrate this point is that, an amateur artist who never learned the theory of art may not realise that she is using an impressionist technique, but it doesn’t change the fact that when she is using short strokes of paint, she is indeed using impressionist technique. Analogously, the reductionists in the imagery debate may have not realised that they use certain assumptions as bridge laws, or they even might have not been aware that they are reductionists, but it does not change the fact that their reasoning is reductive and fits into a Nagelian model.

The final question is one of multiple realisability. Recall that Fodor argues that bridge laws of mind-brain reduction are disjunctive and therefore not actual laws. I shall borrow from Dizadiji-Bahmani et al.'s counterarguments one more time. They give three reasons why this is an invalid objection, from where I will take a further step. They write:

> "First, as Sober points out, it is not clear where to draw the line between disjunctive and non-disjunctive laws, since what is non-disjunctive in one formulation could turn out to be disjunctive in another one and vice versa. Second, even if it is true that 'proper' laws of nature (whatever these are) cannot be disjunctive, there is no need for bridge laws to be laws of nature in that sense. Bridge laws can be of a different kind and have to satisfy less stringent demands than other laws of nature. All we require from bridge laws is that they serve the purposes of reduction, and disjunctions pose no problem for these. Third, it is not clear why laws of nature cannot have disjunctive form." (Dizadiji-Bahmani et al., 2010)

In the case of the mental imagery debate, these three points make perfect sense. Because I hope to point out that, Fodor uses multiple realisability to argue against type-identity theory of mind, but his functionalist reduction cannot avoid the problem of the same nature. In the same sense that pain can be realised by human brains and Martian brains, a cognitive property can also be realised by the language of thought or any other kind of computational system, this gives a state of mind multiple functional realisers. Therefore, a functionalist has to either admit that their account is vulnerable to the objection that they raised to undermine their competitor, or they can take an easy way out by admitting that bridge laws can be disjunctive. However, as I have argued, bridge laws are factual claims, it cannot be the case that all the available bridge laws are true. In the case of cognitive science, human pain and Martian pain may turn out to be different kinds of pain, and the brain can only work on one computational system, it is either a turning machine or a connectionist neural network, it cannot be both. The empirical question is left open because we are calling for new breakthroughs in neuroscience and cognitive science in general, multiple realisability only seems threatening at this point because we do not have enough empirical evidence to rule out or confirm certain hypotheses, which makes the disjunction look massively open-ended.

4.3 Summary

In section 4, I made the point that the moderated model of Nagelian reduction is able to overcome the problems we mentioned at the beginning of this paper. Furthermore, the moderated account of Nagelian reduction has already been employed by actual scientific practice.

5 Potential Objections and Final Remarks

In this paper, I offered a defence of the Nagelian model of intertheoretical reduction. I conducted a case study of the mental imagery debate in cognitive science and argued that the Nagelian reduction is an adequate descriptive account of scientific reduction, although it needs a few adjustments of its theoretical assumption and applicational means. I argued in detail that the pictorialists and the descriptionalists are both reductionists in the mental imagery debate, and their reasoning processes are paradigmatic examples of the application of the Nagelian model in scientific practices.

However, I left it unanswered that whether the Nagelian model is normatively efficient, which might cause possible objections to my argument. But I believe this is another issue and calls for a separate discussion. The goal of my argument is to show that, as a matter of fact, the Nagelian model manages to cover and explain the reduction within scientific practice, more specifically, within the research of cognitive science. Whether the model tells how a reduction needs to be conducted is a different question to investigate.

In conclusion, the Nagelian model as a classic reductive model gives very good insights into the reductive processes in scientific research. The use of a case study has the potential to explore and examine the descriptive aspect of the model. However, I did not wish to generalise my conclusion from one case to the general practice of scientific research. This case may be limited by its historical and temporal factors, i.e. both type-identity theorists and functionalists in the imagery debate may have been influenced by logical positivism, as well as Nagel himself. But the point has been made is that Nagelian reduction does provide explanations to phenomena as it aims to.

Bibliography

Brigandt, Ingo and Alan Love, 2017, "Reductionism in Biology", *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2017/entries/reduction-biology/>.

Curd, Martin and Cover, J.A., 1998, "Commentary" *In Philosophy of science: The central issues* (2013), W.W. Norton & Company, Inc. : 1003-1044

Dizadji-Bahmani, Foad, Roman Frigg, and Stephan Hartmann. "Who's afraid of Nagelian reduction?." *Erkenntnis* 73.3 (2010): 393-412.

Fodor, Jerry A. 1968, "Psychological explanation: An introduction to the philosophy of psychology."

Fodor, Jerry A. 1975, *The language of thought*. Vol. 5. Harvard university press.

Nagel, E., 1961, *The Structure of Science. Problems in the Logic of Explanation*, New York: Harcourt, Brace & World, Inc.

Nagel, E., 1974, "Issues in the Logic of Reductive Explanations." In *Philosophy of science: The central issues* (2013) Curd, Martin and Cover, J.A., W.W. Norton & Company, Inc. : 911-926

Nathan, Marco J., and Guillermo Del Pinal., 2016, "Mapping the mind: bridge laws and the psycho-neural interface." *Synthese* 193.2: 637-657.

Pylyshyn, Zenon W., 2002, "Mental imagery: In search of a theory." *Behavioral and brain sciences* 25.2 : 157-182.

van Riel, Raphael and Robert Van Gulick, 2019, "Scientific Reduction", *The Stanford Encyclopedia of Philosophy* (Spring 2019 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2019/entries/scientific-reduction/>.

Shepard, Roger N., and Jacqueline Metzler. "Mental rotation of three-dimensional objects." *Science* 171.3972 (1971): 701-703.