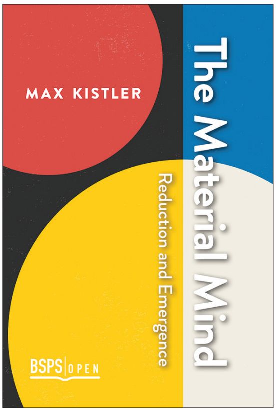

THE MATERIAL MIND: REDUCTION AND EMERGENCE

Max Kistler

ISBN 978-1-77385-607-0

Acknowledgement: *We acknowledge the wording around open access used by Australian publisher,* **re.press**, *and thank them for giving us permission to adapt their wording to our policy* http://www.re-press.org

2

Can Reductive Explanations Be Constructed A Priori?

1. Introduction

The debate on the nature of the mind and its relation to the body in contemporary analytical philosophy starts from a thesis, the truth of physicalism, and an observation, the existence and relative autonomy of psychology. Physicalism is the ontological doctrine according to which (1) everything that exists either belongs to one of the categories studied by physics or is composed entirely of parts that belong to one of these physical categories, and (2) all of the objective properties of the entities recognized in (1) are either properties studied by physics or reducible to them. Moreover, the very existence of scientific psychology seems to show that there are domains of psychological phenomena within which it is possible to discover regularities independently of the underlying physiological or physical phenomena. This apparent autonomy of psychology seems to suggest that it is irreducible to neuroscience and, even more so, to physics. In Chapter 1, I considered an initial influential argument for the irreducibility of psychology developed within the functionalist conception of mental states: Putnam (1967) and Fodor (1974) argued that psychological properties are multi-realizable, whereas reduction requires bridge principles whose existence is incompatible with multiple realizability.

Davidson (1970) advanced a second argument for the irreducibility of psychology to physics: namely, the conceptual framework of intentional psychology is radically heterogeneous with respect to the conceptual framework of physics and physical sciences, insofar as the attribution of psychological predicates to persons and the attribution of physical predicates to the same persons obey incommensurable criteria of correctness. The application of a physical predicate is governed by an experimental procedure and in the simplest cases by a measurement. The attribution of a psychological predicate must obey a constraint of an entirely different kind: it must make the person to whom the predicate is attributed appear to be rational. Similarly, the criteria for evaluating explanations belonging to these two conceptual frameworks, mental and physical, are radically different. The attribution of mental states, and the explanation of actions in psychological terms, are subject to norms of rationality: for an action to be rational, the means chosen must be adequate given the agent's order of preferences and set of beliefs.[1] In contrast, the standards of correctness for the attribution of physical properties, as well as for physical explanations, are essentially agreement with observation and logical validity, within the deductive-nomological model of explanation.

For these two reasons, it has often been taken for granted that there can be no psychophysical laws, and that psychology is irreducible to physics in principle.[2] Yet this conviction seems to be in contradiction to the doctrine of physicalism, according to which all real properties, in contrast, are reducible in principle to physics. The philosophy of psychology is thus faced with the challenge of finding a way to reconcile the acceptance of physicalism with the autonomy of psychology.

The thesis of the supervenience of psychological properties on physical properties seemed to be able to reconcile physicalism with the irreducibility of the mind. Among the many concepts of supervenience that have been explored, strong supervenience has emerged as the most promising to characterize the relationship between psychological and physical properties. For any set of properties $\mathscr{M}$ and any set of properties $\mathfrak{R}$, $\mathscr{M}$ strongly supervenes on $\mathfrak{R}$ if and only if, necessarily, for any property $M \in \mathscr{M}$, for any x, if x is M, then there exists a property $P \in \mathfrak{R}$, such that x is P and, necessarily, for any y, if y

1 A traditional way of analyzing the rationality of an action is in terms of a practical syllogism. It is rational for agent X to do A if and only if X's doing A is the conclusion of a syllogism whose most important premises are (1) X wants B, where B is a desire of X's that X gives priority to under the circumstances, and (2) X believes that performing A under the circumstances is an adequate way to obtain B. The fact that an action is rational in this sense does not prevent the representations of the reasons for performing it from also being the causes of the bodily movement constituting the action. See Davidson (1963); Kistler (2006c). This thesis is opposed to a traditional doctrine according to which the explanation of an action in terms of its reasons belongs to a conceptual framework incompatible with that of causes.

2 Block speaks of the "anti-reductionist consensus" (1997, 107).

is *P*, then *y* is *M*.[3] One consequence of strong supervenience is that, if the psychological properties of a person supervene on the physical properties of her body (and her environment[4]), then it is impossible for there to be two persons whose bodies (and environments) share all of their physical properties but differ in one of their psychological properties. The concept of supervenience has emerged as a promising tool for reconciling the autonomy of psychology with physicalism, insofar as the relation of supervenience, even strong supervenience, is very weak. In particular, the systematic correlation between the underlying properties $\mathfrak{R}$ and the supervening properties $\mathscr{M}$ is compatible with the absence of psychophysical laws.

However, those who hoped that the use of the concept of supervenience would be sufficient to reconcile physicalism with the irreducibility of the mind have been disappointed.[5] As Horgan (1993) and Kim (1993a) have shown, the strong supervenience of the set of mental properties (and the corresponding states of affairs) on the set of physical properties (and the corresponding states of affairs) imposes no constraint on the origin of their correlation; strong supervenience does not guarantee the truth of physicalism. This is clear from the definition of physicalism given above: the reducibility of all properties to those of physics is part of it. However, supervenience is compatible with dualistic and thus anti-physicalist metaphysical theories, notably with parallelism or occasionalism: if God's intervention guarantees a perfect correlation between physical and mental properties, then the latter supervene on the former. This shows that the postulate of supervenience alone does not require the physical to determine the mental, nor does it require the mental to depend on the physical: in some dualist doctrines compatible with supervenience, all properties are determined by God's will and are dependent only on it.

As long as the nature of $\mathscr{M}$ properties is not specified, the necessary correlation of $\mathscr{M}$ properties with $\mathfrak{R}$ properties is compatible with the radical heterogeneity of $\mathscr{M}$ properties with respect to $\mathfrak{R}$ properties, as in classical dualism, reinterpreted in terms of properties. In other words, the existence of a universal correlation between mental and physical properties — even a

3 In symbols: $\Box\, (\forall M \in \mathscr{M})\, (\forall x)[Mx \rightarrow (\exists\, P \in \mathfrak{R})(Px \wedge \Box\, (\forall y)\, (Py \rightarrow My))]$.

4 If the environment is not mentioned, then the thesis becomes that of local supervenience. We will come back to the distinction between local and global supervenience in Chapter 3.

5 I have developed this point elsewhere (Kistler 2004b) and will return to it in Chapter 4.

necessary correlation such as those in strong supervenience — contains no indication of the origin or explanation of this correlation.

The conception developed in this book overcomes this difficulty by conceiving of the relationship between the physical and the mental on the model of nomological determination, in virtue of non-causal laws of composition. Nomological determination thus appears as the metaphysical foundation of supervenience and allows for its explanation.[6]

Based on the inadequacy of the concept of strong supervenience to express the doctrine of physicalism, a number of authors pursue a completely different strategy to achieve a satisfactory conception of the relation between body and mind: they conceive of the connection between physical and psychological truths as even closer than the necessary correlation of strong supervenience.[7] These authors develop the idea that psychological propositions are merely redescriptions of physical states of affairs in another vocabulary. The psychological conceptual framework allows us to redescribe, with different concepts, the same set of states of affairs that appears as physical when described with physical concepts. The relations of "logical supervenience" (Chalmers), "strict implication" (Kirk), or "entailment" (Jackson) are supposed to ground the physicalist determination of psychological states of affairs by physical states of affairs while avoiding the seemingly mysterious necessity that is part of the concept of strong supervenience.

The general strategy of the proponents of "conceptual reduction," as I propose to call it, is to ground the physicalist determination of the mental by the physical, no longer in a form of natural necessity (compatible with dualism) but in a necessity of conceptual origin. According to Kim (1998), there are no mental properties, only psychological concepts, which are second-order concepts; according to Chalmers and Jackson (2001), psychological concepts are such that one can determine a priori which states of affairs (formulated in physical terms) they apply to, provided that one possesses

6 See Chapters 3 and 4. Broad (1925) attributes to emergence a characteristic often taken — wrongly, as we have just seen — to be an essential component of the concept of supervenience: the dependence of supervenient properties on the properties in their base (or the determination of supervenient properties by their base properties). The definition of supervenience does not, in fact, guarantee such dependence or determination.

7 This strategy forms the common thread of otherwise different conceptions of the mind in nature that have been proposed by Yablo (1992, 1997); Chalmers (1996); Jackson (1998); Kim (1998); Chalmers and Jackson (2001); Kirk (2001); and Esfeld and Sachse (2011).

a complete microphysical description of the actual world. Rather than putting the problem of understanding the relationship of the mental and the physical in terms of different kinds of properties, they conceive of it in terms of the relationship between true propositions (or "truths") expressed in mental vocabulary and true propositions expressed in physical vocabulary. According to this approach, the link between the physical and the psychological is not a natural link but a conceptual one. This implies that it is possible, in principle, to obtain knowledge of any non-physical state of affairs (e.g., a mental one), from a complete knowledge of physical states of affairs, without further empirical investigation (i.e., in a purely a priori manner).

A Laplacian demon[8] that knows the set *P* of all physical states of affairs could extract the set of all other — in particular mental — states of affairs, solely via a priori conceptual analysis regarding *P*. According to Chalmers, this is possible even if *P* contains only microphysical states of affairs: "Laplace's demon, say, who knows the location of every particle in the universe — would be able to straightforwardly 'read off' all the biological facts, once given all the microphysical facts" (1996, 35). According to these authors, the fundamental thesis of physicalism is that the set of all physical states of affairs determines the set of all states of affairs, including, in particular, the set of mental states of affairs. Jackson expresses the thesis by saying that "the psychological account of our world is entailed by the physical account of our world" (1998, 24).[9] To use Kim's metaphor, having created the set of physical states of affairs, God had no further work to do in creating all mental states of affairs.[10]

8 A hypothetical being with unlimited reasoning and memory capabilities that allow it to know an exhaustive description of the world at the microphysical level, and to calculate from this description, as well as from the laws of nature, both the future and the past is called a "Laplacian demon":

> An intelligence which, at a given moment, would know all the forces of which nature is animated and the respective situations of the beings which compose it . . . would embrace in the same formula the movements of the largest bodies in the universe and those of the lightest atom; nothing would be uncertain for it, and the future, like the past, would be present to its eyes. (Laplace 1825, 32–33)

Here I am concerned not so much with the power to calculate the future and the past as with the power to derive a description of a state of affairs in macroscopic terms from its description in microscopic terms, at the same instant.

9 Chalmers also speaks of the "logical supervenience" (1996, 33) of the set of all states of affairs on the set of physical states of affairs, whereas Kirk (1996, 2001) says that the latter "strictly implies" the former.

10 This metaphor is often used, for example by Chalmers (1996, 35).

Here is how Jackson defines this "minimal physicalism": (J) "Any world which is a *minimal* physical duplicate of our world is a duplicate *simpliciter* of our world" (1998, 12). By a "minimal physical duplicate of our world," Jackson means a world that is perfectly similar to our own in all physical respects. The qualifier "minimal" means that such a world contains nothing more than what is necessary given its physical constitution.[11]

The purpose of this chapter is to question the possibility of deducing non-physical truths a priori from a description of the world in microphysical terms. The thesis of a priori deducibility from the complete microphysical description *P* is meant to hold for all non-microphysical truths. It can therefore be challenged, without entering into controversies about the specificity of truths about the mind, by focusing on common-sense truths, such as "Water covers most of the Earth" (Jackson 1998, 73).

We will see that the truth of physicalism is not sufficient to guarantee the possibility of deducing macroscopic common-sense truths a priori from *P*, because knowledge of *P* is not sufficient for the construction of their reductive explanation; indeed, such a construction has an a posteriori part that goes beyond knowledge of *P*.

2. A Priori Reduction in the Framework of Two-Dimensional Semantics

Consider this macroscopic fact expressed with common-sense concepts:

(*) Water covers most of the Earth.

According to Chalmers and Jackson, facts of this kind can be inferred a priori from two premises:

(1) a complete description of the state of the world in microphysical terms, and

11 Jackson points out that (J) expresses contingent global supervenience: the truth of the physicalist thesis is contingent insofar as it bears only on the actual world, not on all possible worlds. It is compatible with physicalism that other worlds contain non-physical substances. Kirk proposes another definition of minimal physicalism in terms of the "strict implication" of all states of affairs by the set of physical states of affairs. See Kirk (1996, 246; 2001, 544–45).

(2) an analysis of the concepts used to express the fact in question.

Such an inference produces what Chalmers and Jackson call a reductive explanation. According to them, "there is an a priori entailment from microphysical truths to ordinary macrophysical truths" (2001, 316). This means that it is possible to know a priori that the material conditional $P \supset M$ is true, where P denotes "the conjunction of microphysical truths about the world" and M a common-sense truth about macroscopic objects and properties, such as water, for example (*), or life: "There are many living things" (317). Their thesis is that a priori conceptual analysis is all that is required to know that $P \supset M$. In Jackson's terms, "physicalism is committed to the in principle a priori deducibility of the psychological from the physical" (1998, 83). In other words, these authors argue that conceptual analysis makes "armchair metaphysics" possible: according to Jackson (1994), conceptual analysis — which can be carried out "in one's armchair" (i.e., without recourse to experience) — is indispensable and fundamental to metaphysics. To use Horgan's (1984) expression, "cosmic hermeneutics" allows all truths to be derived a priori from a (hypothetical) complete description of the world in microphysical terms.

Chalmers and Jackson seek to establish their thesis within the conceptual framework of two-dimensional semantics (Chalmers 1996, 2004; Jackson 1998). We must be content here with a brief presentation of the fundamental concepts that they use in their argument. Primary intension plays a key role. Generally speaking, the extension of a predicate is the set of objects to which it applies; its intension is a function that determines the extension of the predicate in each possible world. Two-dimensional semantics was originally developed in the context of the semantic analysis of statements containing indexical expressions, such as the words *I* and *here* (Stalnaker 1978; Lewis 1980). In the case of such terms, the intension is a function determined by two factors: the context of utterance and the context of evaluation. When I utter the word *I* on a given occasion, the context of utterance determines, together with the lexical meaning of the word (often called, following Kaplan [1989], the "character" of the word), the reference of the word: namely, in the case of *I*, the speaker. It is therefore the speaker who figures in the content of the proposition expressed. Now let us consider the context of utterance as given. The proposition expressed is therefore well determined. We can then ask ourselves about the modal status of this proposition: is it contingent or necessary? The answer depends on the truth value of the proposition in the set of possible worlds. We therefore need to know the extension (or reference) of the terms contained in the proposition in other possible worlds. "I" is a rigid term (Kripke 1972): that is, given a context of utterance, the reference of the term is the same in all possible worlds where the proposition can be evaluated. Other expressions, especially definite descriptions such as "the fastest man over 100m in 2022," are not rigid and denote different individuals in different possible worlds. For an indexical term, the two factors that determine its extension in other possible worlds — the context of utterance and the possible world in which the proposition is evaluated — are therefore independent; this is why we can speak of two "dimensions" of intension.

Here is the definition of the primary intension of a term: it is the function that associates an extension to each context considered as both context of utterance and context of evaluation. This notion is relevant because the speaker is often unaware, at least in part, of the context of utterance. The speaker might be unaware of certain aspects of the context of utterance that determine the content of the indexical terms and thus of the proposition expressed: she might not know where she is when she says *here* or what time it is when she says *now*. However, insofar as she knows the lexical meaning (the character) of the term that she uses, this does not prevent her from knowing the primary intension of the term (and of the proposition expressed) a priori. We can express the primary intension of the word *now* by a series of conditionals: if the word is uttered on Monday at noon (context of utterance), then it denotes, at the same world (context of evaluation), Monday at noon; if the word is uttered on Tuesday at 10 a.m., then it denotes, at the same world, Tuesday at 10 a.m. In each conditional, the antecedent is a world that could, for all the speaker knows, be the one that the speaker is in, its consequent being the reference of the word in that world.

It is crucial for Jackson and Chalmers' argument to assume that the two-dimensional analysis of intension can be applied to other than indexical terms. Kripke (1972) and Putnam (1975a) have suggested that natural kind terms, such as "water," also have an indexical aspect. This suggestion was later developed by Stalnaker (1993) and Haas-Spohn (1995, 1997) as well as by Chalmers and Jackson. According to this hypothesis, terms referring to natural kinds such as "water" that are not overtly indexical nevertheless possess a "hidden indexicality." Insofar as we are partly unaware of the nature of water, the actual world in which we find ourselves acts as the context of utterance: the actual world determines, together with the lexical meaning of the term "water," the reference of each utterance of the term. Let us say that there are three epistemic possibilities of what the content of the term "water" might be: either our world is such that within it water is H_2O, or more precisely that within it water consists of macroscopic samples composed overwhelmingly of H_2O molecules,[12] or that in it water is XYZ or ABC. Like indexical expressions, I can be unaware of which of these worlds I am in yet know the primary intension of the word a priori: if the actual world is such that water is H_2O (context of utterance), then the extension of the term "water" in that world (context of evaluation) is H_2O. Conversely, if the actual world is such that water is XYZ, then the extension of the term "water" in this world (context of evaluation) is XYZ.

The secondary intension is the function that assigns an extension to a term in every possible world (where these worlds are all taken to be counterfactual, except the actual world), where the content of the term is assumed to be determined either by the linguistic meaning alone or by the meaning together with the context of utterance. Kripke (1972) argued that natural kind terms, like proper names, are rigid. This means that their secondary intension is constant: if the reference of the term "water" in the actual world is H_2O, then it has the same reference in all possible worlds. In other words, even when we consider counterfactual worlds in which certain states of affairs concerning water differ from the actual world, we are still talking about the substance that fills the oceans in the actual world.

Jackson and Chalmers' argument proceeds as follows. We have seen that the primary intension of common-sense concepts, such as water, is accessible to us a priori, through conceptual analysis. The primary intension of such a term corresponds to its "character": it is the linguistic meaning, known a priori to all competent speakers. In the case of "water," this meaning can be abbreviated as "the watery stuff we are actually acquainted with" (Jackson 1998, 75). This linguistic meaning determines, together with the context of utterance, in particular the actual world, the content of an utterance of the term. The primary intension of a term consists of a set of application criteria, meaning that it can be expressed by a set of conditionals: each has as its antecedent the description of a world taken to be actual and as its consequent the

12 This precision will be implied henceforth.

extension of the term in that world. Let PH_2O, PXYZ, and PABC be complete descriptions of all the microphysical states of affairs of three (epistemically) possible worlds that differ only in the composition of the aqueous substance. To know the primary intension of "water" is to know the following conditionals: if PH_2O, then water is H_2O; if PXYZ, then water is XYZ; and if PABC, then water is ABC. The a priori knowledge of the primary intension is essentially conditional in that it is a function, which associates to each context (or world) of utterance an extension in the world of evaluation identical to the world of utterance. To know the value of the function (the extension in the world of evaluation), I must know (a posteriori) its argument (the context of utterance). In other words, I must know what the world of utterance is. According to Chalmers and Jackson,

> if a subject possesses the concept "water," then sufficient information about the distribution, behaviour, and appearance of clusters of H_2O molecules enables the subject to know that water is H_2O, to know where water is and is not, and so on. This conditional knowledge requires only possession of the concept and rational reflection, and so requires no further a posteriori knowledge. . . . Possession of a concept such as . . . "water" bestows a *conditional ability* to identify the concept's extension under a hypothetical epistemic possibility. . . . Because all the relevant empirical information is present in the antecedent of the conditional, empirical information plays no essential role in justifying belief in the conditional. So . . . [this conditional] is a priori. (2001, 323–25)

The primary intension of a concept does not give us its extension, in a given world, but it does tell us how the context (i.e., the nature of a given world) determines this extension. The extension of the term "water" depends on the world of utterance, but knowledge of a physical description of the world of utterance (PH_2O or PXYZ etc.) puts the possessor of the concept "water" in a position to determine a priori the extension of the concept in that world.

As Chalmers and Jackson put it, "if a subject possesses a concept and has unimpaired rational processes, then sufficient empirical information about the actual world puts a subject in a position to identify the concept's extension. . . [A] 'water'-free description of the world can enable one to identify the referent of 'water'" (2001, 323).

Chalmers and Jackson seek to show that *P* (the full description of the real world in microphysical terms) allows us to infer a priori that

(*) Water covers most of the Earth.

The structure of this a priori deduction is as follows. *P* is supposed to contain the information that

(1) H_2O covers most of the Earth.

Then the conceptual analysis of the word *water* yields (this is an analytical and a priori truth) that

(2) water is the watery stuff that we are acquainted with.

Finally, the context of the utterance of (*) — that is, the world in which (*) is uttered — provides us with the information that

(3) H_2O is the watery stuff that we are acquainted with.[13]

(1), (2), and (3) together allow us to derive (*).

The possibility of such an a priori derivation of all macroscopic, common-sense, and scientific truths from a complete description of all microphysical states of affairs, through conceptual analysis alone, has been challenged on various grounds, notably by Block and Stalnaker (1999) and Byrne (1999).

First, (1) contains the macroscopic concept of the Earth. It is therefore necessary to justify the idea that one can derive (1) a priori from *P*, exclusively composed of truths in microphysical terms. This seems to be doubtful for reasons that were presented in Chapter 1 and to which we will return: the

13 This corresponds to the context of the actual world. In another possible world, the context would have determined, for example, this information: "XYZ is the aqueous substance that we are familiar with."

concepts that one uses to describe microscopic objects do not contain information about the macroscopic properties of the objects composed by these microscopic objects.[14] For this reason, the deduction of macroscopic properties from information about microscopic properties alone cannot be a priori.

Second, it is questionable whether the set of all microphysical truths, expressed in the language of "ideally completed physics," is well determined.[15] The concept of a completed or ideal physics is often used, for example, to define the concept of the law of nature.[16] However, the existence of scientific revolutions prevents us from conceiving of "completed physics" as a conservative extension of current physics. There is no reason to think that the concept of completed physics determines a single system of concepts and propositions, rather than a multitude of theoretical systems, all empirically adequate but incompatible with each other. Now, without a well-determined antecedent *P*, the implication $P \supset M$ has no well-determined meaning either, and the question of its a priori character cannot even be asked.

Third, in the remainder of this chapter, I will point out another major weakness of Chalmers and Jackson's argument. The epistemic status of (3) is problematic: Block and Stalnaker have pointed out that "the claim that H_2O is the (or even a) satisfier of the primary intension of 'water' is not a microphysical claim" (1999, 45). Proposition (3) is not part of *P*, so it cannot be used in the premises of a priori deduction in the same way as (1). Nor is it an a priori truth, so it cannot be used in the same way as (2). Block and Stalnaker offer no analysis of the nature and epistemic status of (3). It is important to fill this gap because the success of Chalmers and Jackson's cosmic hermeneutics project depends crucially on the status of (3). If it is true, as I will try to show, that (3) cannot play the role that Chalmers and Jackson ascribe to it, then we have no reason to think that macroscopic truths can systematically be deduced a priori from *P*. Specifically, I question the thesis that *P* conceptually entails (3) or, in Jackson's terms, that "a rich enough story about the H_2O way things are does conceptually entail the water way things are" (1998, 149).

---

14 See section 4 of this chapter.

15 This objection has been raised by Byrne (1999). See Chalmers and Jackson (2001, 334).

16 This is particularly the case with the so-called best system view advocated by David Lewis (1973, 73). See Kistler (1999b, Chapter 6; 2006d, Chapter 6).

3. Two Concepts of Reduction and Realization: Micro-Macro and Role-Occupant

In order to produce reductive explanations of macroscopic phenomena (with the exception of qualitative aspects of subjective experience, which Chalmers takes to be irreducible), Chalmers argues that it is sufficient to have (1) detailed knowledge of microphysical states of affairs and to have accomplished (2) the "functional analysis" (1996, 43) of macroscopic concepts. The former is empirical in origin, but the latter can be accomplished in a purely a priori manner.

Once the functional analysis[17] of the concept that describes a macroscopic phenomenon has been completed, all that remains to be done is to discover "how those functions are performed. . . . Once the relevant details are in, a story about low-level physical causation will explain how the relevant functions are performed, and will therefore explain the phenomenon in question" (Chalmers 1996, 44).

Chalmers uses the reductive explanation of heat as an example. Heat itself is a physical concept, but according to Chalmers its microreduction follows the same pattern as the microreduction of non-physical macroscopic phenomena, in particular psychological ones. According to the functional analysis of the macroscopic concept of heat, it "is the kind of thing that expands metals, is caused by fire, leads to a particular sort of sensation, and the like" (Chalmers 1996, 44–45). This analysis shows that heat — what was only implicit before the analysis — is "a causal-role concept," which characterizes itself "in terms of what it is typically caused by and of what it typically causes, under appropriate conditions" (Chalmers 1996, 45).

In general, the functional analysis of a concept shows that the concept describes a causal role. Accordingly, to complete the reductive explanation, it is sufficient to discover, in a second empirical step, what fulfills the role thus defined: it is discovered "that heat is realized by the motion of molecules" because "the motion of molecules is what plays the relevant causal role in the actual world" (Chalmers 1996, 45). However, as we will now see, the a priori deduction of (*) from *P* and the functional analysis of the concept of

17 It is analogous to what Kim (1998) calls the "functionalization" of macroscopic concepts. However, Kim specifies that it is the first step in functional reduction, whereas Chalmers says that a reductive explanation is "accompanied" in general (1996, 43) by such a functional analysis.

water is fallacious because it exploits an equivocation about the meaning of the concept of reduction, combined with an equivocation about the concept of realization. Once the ambiguity is removed and the two meanings of "reduction" and "realization" are distinguished, we will see what the conceptual analysis can and cannot achieve. It will also explain why the possibility of "cosmic hermeneutics" appears to be plausible at first sight.

Let us consider the case of heat. According to the analysis of this concept by Chalmers, heat is that which causes certain events and processes (e.g., raising temperature) and that which is caused by certain events and processes (e.g., combustion). The reductive explanation is then accomplished by finding out what fulfills this role (i.e., what "realizes" heat). According to Chalmers, it is thus possible to bridge the distance between a role concept (i.e., a second-order concept) and the first-order concept of what fulfills the role, alongside the distance between a macroscopic property concept and an underlying microscopic property concept such as molecular motion, in a single step.

However, there are in fact two steps to be taken.[18] The functional description defines a role in terms of interactions between macroscopic objects, a role that can be played only by a macroscopic property. The distinction between the role and the occupant of the role is independent of the distinction between the microscopic and the macroscopic: there are macroscopic roles fulfilled by macroscopic properties and microscopic roles fulfilled by microscopic properties. Two theoretical roles contribute to determining the identity of heat.[19] (1) The heat δQ lost by a closed physical system is equivalent to the work δW that it provides,[20] and (2) in a reversible process, the change δQ_{rev} in the amount of heat contained in a system is proportional to the change in its entropy (dS) and to its temperature. (In symbols, $\delta Q_{rev} = TdS$). However, only a macroscopic property (i.e., a property of macroscopic objects) can play these roles.

18 In response to Byrne (1999), Chalmers and Jackson (2001, 334n16) acknowledge that such a deduction must involve two steps. However, their argument for a priori deducibility does not take into account the step corresponding to the reduction from the macroscopic to the microscopic, the discovery of which, as I will show later, is always a posteriori.

19 The word heat designates the property that occupies the role, not the role itself, but it does so by way of a definite description of the role: heat is the property that has such and such functional and causal relations with such and such other properties.

20 Given that the total internal energy U is constant in an isolated system, $dU = 0$ and therefore $\delta Q = - \delta W$.

This point is worth looking at a little closer. The distinction between "macroscopic" and "microscopic" can be taken in a narrow or broad sense. In the narrow sense, an object is called "microscopic" in comparison to a given macroscopic object if it is smaller by several orders of magnitude than the latter. In the broad sense, each constituent part of an object can be called "microscopic" relative to the object as a whole and the object itself "macroscopic" in relation to that part. Heat is an essentially macroscopic property in the sense that it cannot belong to microscopic objects (in the narrow sense): a single atom cannot be hot. It is part of the conditions for the possibility of attributing the property of being hot to an object that the object has microscopic components, preventing it from being attributed to the individual microscopic components themselves.[21]

The first step in the microreduction of heat is to associate a categorical property with the role of heat: a macroscopic property designated by a first-order predicate is discovered, which plays the role, itself designated by a second-order predicate, typically in functional or dispositional terms. The discovery of the microproperties of the parts of the hot object that give rise to the macroproperty that plays the role occurs in a second and independent step: one step can be accomplished without the other.

Let us call them, respectively, RO reduction (RO for role-occupant) and MM reduction (MM for micro-macro). A reduction of the first kind, an RO reduction, leads to the discovery that a categorical property plays a previously determined role. For example, the concept of heat is primarily a role concept; this role can be made explicit by conceptual analysis. The development of thermodynamic theory led to the construction of the concept of heat as a form of energy equivalent to work: this concept was central to Carnot's (1824) theory of the heat engine. The RO reduction identifies internal energy, a categorical concept, as what fulfills the role of heat. The RO reduction is a conceptual reduction[22] that does not involve different properties; it consists

21 "Microscopic" components in the broad sense might themselves have components. What is crucial here is that heat cannot be ascribed to microscopic components in the narrow sense, such as isolated atoms or molecules.

22 It cannot always be accomplished a priori: this is possible only if one already knows the functional description and a categorical description of the property. Therefore, even RO reductions are not a priori in the sense that the reduction can be constructed by using only the categorical basis alone. The functional description of a property cannot by deduced a priori from any of its categorical descriptions.

of discovering that a property known by a categorical description plays a role characterized functionally or dispositionally.

In contrast, an MM reduction, typically the result of a later stage of scientific research on a natural property, brings different properties into relation: properties of a macroscopic object and properties of its microscopic parts. In the case of heat, Boltzmann and others discovered that laws involving heat could be derived from molecular models. This MM reduction of heat was discovered later than its RO reduction, in the 1870s. When I wrote earlier about the reduction of the property of being water and the property of having temperature *T* to the properties of the components of the objects having these properties and their interactions, I was talking about MM reductions.

Take the case of water: it has the functional or dispositional property of being transparent to light.[23] If water in its liquid state is exposed to light, then light passes through it, so that we can see through it. The reductive explanation of this property of water goes through two steps. First, the macroscopic dispositional property of transparency is RO reduced to the macroscopic property of having a certain absorption spectrum of electromagnetic radiation.[24] This property manifests itself in the form of transparency: water is transparent to the rays that it does not absorb.[25]

Second, the absorption of infrared in water is explained (in the form of an MM reduction) by the absorption of "parts" of light by "parts" of water. Individual photons are absorbed by individual molecules provided that their energy (and wavelength) correspond to the characteristic energy of one of the intramolecular vibrations accessible to the molecule given its geometry.[26]

---

23 Needham (2000) shows that certain macroscopic characteristics are part of the identity conditions of water.

24 Water absorbs rays whose wavelength falls in the centimetric range, then in the infrared (wavelength between 2 and 6 mm), then in the far ultraviolet (wavelength 1,650 Å). See Caro (1995, 86).

25 The property of having a certain absorption spectrum can be conceived as dispositional or as categorical. In Chapter 3, we will see that the distinction between dispositional and categorical is a semantic distinction concerning the meaning of predicates rather than a distinction between types of properties.

26 Symmetrical vibration of the two O atomic nuclei with respect to the H nucleus, antisymmetrical vibration (where the directions of movement of the O nuclei are opposite), or torsion; absorption in the centimetre wave range is explained by the absorption by the molecules of the energy required for rotations; absorption in the ultraviolet range is explained by the absorption — by the molecular electrons — of the energy required to pass into a molecular orbit that corresponds to an "excited" state of the electron.

When it is said that these microscopic mechanisms "realize" the transparency of water, the word *realize* can have two meanings, which might contribute to obscuring the difference between the two stages of reduction. It is possible to speak of "realization" to designate the two relations: one can say that part of the internal energy of a gas (δQ_{rev}) "realizes" the cause of the increase in entropy, as a function of temperature, according to the formula $\delta Q_{rev} = TdS$. In this context, the word *realization* refers to RO realization, a relation between what occupies the role and the role itself.

But there is another meaning of "realization" that expresses what might be called "micro-macro realization" or "MM realization": it is the relationship between the microscopic properties of the components of an object and a macroscopic property of that object to which the interaction between the components gives rise. It is in the sense of MM realization that Chalmers can say that the motion of molecules "realizes" heat: the motion of molecules is the microscopic property that MM realizes heat as a role-occupant (i.e., as a first-order macroscopic property).

The problem is that only RO realization can be discovered a priori, whereas MM realization is always discovered a posteriori. When both the role and the occupant are known, it can be discovered a priori that they are in a role-occupant relationship. In contrast, the discovery of a microreduction (i.e., the discovery of the microscopic properties and interaction laws that determine the macroscopic property together), is always a posteriori.

The choice to call both of these relationships "realization" can be misleading. In reality, the only thing that they have in common is that each corresponds to one of the two reduction relations that I have distinguished. However, the differences are important: RO realization corresponds to a relation between concepts, just like RO reduction, whereas MM realization corresponds — like MM reduction — to a relationship between distinct properties, microscopic in one case and macroscopic in the other. An MM reduction describes how the microscopic properties of the parts that make up an object naturally determine its macroscopic properties, whereas an RO reduction consists of the discovery of a categorical description of a property first conceived of in a functional way.

4. Multi-Realizability

The concept of realization, like that of implementation, allows us to conceive of the possibility that a property can be realized in different ways: such a property is "multi-realizable." The thesis that mental properties are multi-realizable was introduced in the philosophy of mind in the context of machine functionalism and the analogy of the mind with computer software. Just as software can be "implemented" in different ways in different machines, so too cognitive states can be implemented in brains with different neurophysiological properties.

The analysis of multi-realizability has enhanced the confusion between the two kinds of realization: macroscopic roles in general can be "RO realized" by different categorical macroscopic properties, so that these roles are RO multi-realizable. But macroscopic properties are also often MM multi-realizable too, in the sense that objects can share macroproperties determined, in different cases, by different microscopic properties of their parts.

RO realization allows for the possibility that a single causal role can be occupied by different occupants.[27] Many biological functions are RO multi-realized in the sense that they are performed by different categorical properties in different biological species. The function of enabling an organism to see (i.e., to give it access to the information contained in the light waves that reach the surface of its body) can be fulfilled by several properties. The property of being a mammalian eye and the property of being an arthropod compound eye are two first-order structural properties that perform the function of enabling an organism to see. Either can play the role of giving the organism access to the information contained in light. Being an antibody is a functionally designed (microscopic) property that can be achieved by "millions of different chemical structures" (Kincaid 1990, 585), which are also microscopic. This example illustrates the above-mentioned fact that an RO reduction can be achieved only after the independent discovery of both a functional description (the determination of the role that antibodies play)

27 The reverse is also true. Morange (1998) mentions numerous examples of biological molecules, and in particular genes, that play several roles assumed to have been acquired successively and independently of each other. Tompa, Szasz, and Buday (2005) and Tobin (2010) analyze the case of so-called moonlighting proteins that fulfill several functions, in an analogy to people having a second job at night, in addition to their main job.

and a categorical description (the RO reduction consists of the discovery that a property that satisfies the latter also satisfies the former).

The natural determination of the properties of a macroscopic object by the properties of its parts gives rise to multi-realizability in a very different sense from that of RO multi-realizability: consider, for example, the property of being a hemoglobin molecule. Its overall structure, or "conformation," allows the molecule to play its biological role of transporting oxygen. This way of speaking tends to obscure the fact that there are many different types of hemoglobin molecules in different biological species that differ in their parts: that is, in the amino acids that make up the sequence of each of the four proteins that make up the four subunits of the molecule (a "tetramer"). The amino acid sequence is known as the "primary structure" of the hemoglobin molecule. Each type of hemoglobin has its own primary structure and differs from the other types in some of its 140 amino acids. Only 9 of 140 positions are occupied by the same amino acids in all hemoglobin species. The chemical properties of these nine amino acids and their interactions determine the "conformation" of the molecule.[28] It is this overall structure of the molecule that allows hemoglobin to play its role in all biological species. As Rosenberg states, "it is the quaternary structure that produces haemoglobin's remarkable functions" (1985, 77).

However, this overall structure can be determined naturally by a large number of different properties at the level of the parts (i.e., by all of the primary structures that have in common the nine amino acids at the "strategic" positions). The existence of a single overall structure common to all of these molecules justifies speaking of the (kind of) hemoglobin molecule in the singular. However, taking into account the different microscopic structures also justifies speaking of hemoglobins in the plural (Rosenberg 1985, 77). Each of these microstructures naturally determines the same overall structure. The natural determination thus establishes a "many-to-one" relationship between the microstructures and the overall structure. Using the term "macroscopic"

---

28 This determination goes through two intermediate steps: the chemical properties of these nine amino acids and their interactions determine where the chain folds or overlaps, giving rise to the "secondary structure," which has the effect of bringing together distant amino acids given their positions in the chain, giving rise to new interactions that determine the "tertiary structure." Finally, the "quaternary structure," which characterizes the overall structure of the molecule, is determined by the interactions between the four subunits that come together in a stable conformation. I will consider this example again in Chapter 5.

in a broad sense, so that it generally characterizes the properties of a whole in relation to the properties of its parts, it can be said, at the risk of confusion with RO multi-realizability, that the macroscopic property of being a hemoglobin molecule can be MM realized by many different microstructures.

The function of transporting oxygen in an organism is not only MM multi-realized but also RO multi-realized (Kurtz 1999). Indeed, in some marine invertebrates, such as brachiopods, hemerythrins perform the oxygen transport function, and in some arthropods and molluscs that role is played by hemocyanins, in which oxygen is bound to a pair of copper atoms rather than being bound, as in hemoglobin, to a heme group around an iron atom (Kurtz 1992; van Holde and Miller 1995; van Holde, Miller, and Decker 2001).

The crucial point for my argument — which might be blurred by the confusion between RO realization and MM realization — is that the discovery of MM realization (i.e., the natural determination of the properties of a whole by the properties of its parts) is always a posteriori. Even if we had an absolutely complete description of a situation in microscopic terms, together with the complete set of microscopic laws that apply to it, we would still not know the MM reduction of the macroscopic properties exemplified by the objects composed from the microscopic objects that appear in the microscopic description.

The reason is that not all of the laws necessary to deduce the macroscopic properties belong to the level of the reducing theory, nor can they be deduced from it. In the important case of the reduction of thermodynamic — and thus macroscopic — concepts, such as heat, entropy, or temperature, their MM reduction to mechanical concepts that apply to the microphysical components of the systems to which the thermodynamic concepts apply depends on the introduction of the concept of an “ensemble” of systems that has no meaning at the microscopic level. The construction of macroscopic concepts such as temperature requires the use of new conceptual tools that have no microphysical equivalent and cannot be constructed with the concepts appropriate for describing microphysical objects and states of affairs.[29]

29 See Sklar (1993) and Chapter 1. Chalmers and Jackson are hesitant about the need to include laws, in addition to particular microphysical facts, in the reduction basis. According to Chalmers, “high-level facts are entailed by all the microphysical facts (perhaps along with microphysical laws)” (1996, 71). Surely, no reduction of a macroscopic phenomenon can be accomplished without using the laws that apply to the microscopic entities mentioned in the reducing theory: without the laws governing the interactions between the molecules of a gas, it is

The same is true of the analysis of other cases of successful MM reduction: the reduction of classical genetics to molecular biology provides genetic explanations of macroscopic phenomena, but molecular biology does not claim to construct concepts applying to macroscopic phenomena.[30] In the reductive explanation of certain elementary forms of learning — such as habituation, sensitization, and classical conditioning — no effort is made to eliminate essentially macroscopic concepts such as stimulus, reflex, conditioning, and withdrawal behaviour in favour of microscopic concepts. Even after the reductive explanation of these phenomena, psychological concepts continue to provide the framework within which they are described (see Chapter 1).

The advocate of the thesis of the a priori deducibility of macroscopic truths from a complete (hypothetical) microphysical description *P* of the world can make two rejoinders. A first rejoinder is to incorporate all of the laws necessary for reduction into the set of premises *P*, including those that are not purely microscopic. However, this would trivialize the thesis of a priori implication (i.e., deprive it of its content), insofar as the content of the premises *P* would no longer be exclusively microscopic.[31]

A second rejoinder is to argue that the fact that one can reduce, for example, thermodynamic laws only by making use of irreducibly macroscopic concepts, such as the concept of an ensemble, only shows the inability of *current* physics to accomplish this reduction from a purely microscopic basis, whereas what is at stake is the *possibility in principle* of such a reduction. Here one admits that the actual reductions accomplished in the history of science *start* from the knowledge of macroscopic properties and that their

impossible to deduce the properties of the gas. However, though the microscopic laws are necessary, they are not sufficient for deducing macroscopic facts: only our prior knowledge of macroscopic phenomena guides us in the construction of concepts adequate to their description.

30 See Kitcher (1984), Schaffner (1993), and Morange (1998) for many illustrations of this fact, in the context of determining the macroscopic properties of organisms from the properties of their genes.

31 Chalmers and Jackson allude to the problem raised here, that no MM reduction can be accomplished a priori, when they point out that "the only worry" about the truth of their thesis that describing the world in microphysical terms implies all descriptions of the world in macroscopic terms "might concern the status of bridging principles within physical vocabulary" (2001, 331). Rather than confronting this difficulty, they propose to "bypass" it "by stipulating that the relevant physical principles are built into P" (331). This indeed solves the problem but at the cost of abandoning the thesis initially defended and criticized here, according to which a (purely) microphysical description a priori implies all macroscopic truths.

elaboration depends on the prior knowledge of the macroscopic properties to be explained.[32] However, while admitting this about the *actual* discovery of MM reductions, the defender of a priori reducibility could simply postulate that it is still *possible in principle* to deduce all macroscopic truths from a purely microscopic basis. A future molecular biology will shed the conceptual framework bequeathed to it by classical genetics to become a purely microscopic science, and a future neuroscience will construct reductions that make no use of macroscopic cognitive concepts.

Indeed, no logical inconsistency seems to prevent such a possibility. Meanwhile, the burden of proof lies with those who assert a possibility that does not correspond to actual scientific discoveries of MM reductions. As long as historical reductions do not confirm the existence of inferences of macroscopic states of affairs from purely microscopic premises, it seems to be gratuitous to assert that such a feat is nevertheless possible in principle.

5. Conclusion

The deduction of macroscopic common-sense truths from *P*, the hypothetical complete description of all microscopic states of affairs, necessarily goes through two steps: the first is the discovery of an RO reduction: that is, the discovery of the property that fulfills (or properties that fulfill) a certain functional role. Conceptual analysis allows us to discover which properties play the roles corresponding to common-sense concepts such as "water" or "heat" or scientific concepts such as "oxygen carrier." However, the RO reduction is not, for all that, an a priori deduction from the microscopic description *P*, insofar as both of these properties and the functions that they perform belong to the macroscopic level.

The explanation of macroscopic phenomena in microscopic terms is the subject of a second reduction step, which I have called MM reduction. Historically, and as we have seen in Chapter 1, the premises of MM reductions were not purely microscopic. First, some laws, in particular statistical laws relating macroscopic properties to the properties of microscopic constituents of matter, are irreducible to the laws governing microscopic properties and their interactions. The historical cases of reductions of biological or cognitive phenomena also involved macroscopic concepts not built upon

32 "Building a model . . . is not a matter of deduction" (Holland 1998, 9).

a microscopic basis. Second, these historical reductions were accomplished only through prior knowledge of the macroscopic phenomena to be reduced: they proceeded by constructing a model of the microscopic phenomena under the constraint of its adequacy to the macroscopic phenomena, known beforehand.

The deduction of a macroscopic truth expressed with common-sense concepts from *P*, as envisaged by Chalmers and Jackson, must have two parts, one corresponding to an RO reduction and the other to an MM reduction. Since neither the historical RO reductions nor the historical MM reductions took the form of deductions from the mere knowledge of microscopic phenomena, the burden of proof is on those who proclaim a principled possibility that does not correspond to the reality of historical reductions.

Chalmers and Jackson's argument that macroscopic common-sense truths, such as (*), can be deduced a priori from a complete description of the world in microscopic terms is fallacious because it relies on an equivocation: the concept of reduction is sometimes understood in the sense of RO reduction and sometimes in the sense of MM reduction. Contrary to what Chalmers and Jackson claim,

(3) H_2O is the aqueous substance that we are familiar with

cannot be deduced from *P* a priori, just with the help of conceptual analysis. One of the steps in the reductive explanation of (3) is a local and a posteriori MM reduction that allows us to deduce, from a microphysical description and various micro- and macrophysical laws, the macroscopic properties of the substance composed of H_2O molecules: the facts that this substance is liquid at ambient temperature and pressure (near the surface of the Earth, in summer, not too close to the poles), has a reduced viscosity, is transparent to light, et cetera.

Proposition (3) has a hybrid character, partly microscopic ("H_2O"), partly macroscopic ("the watery substance"). For Chalmers and Jackson's argument to be valid, it would have to be purely microscopic, on the one hand: then it would be plausible that it is a priori derivable from a complete microphysical description of the world. On the other hand, it would have to be purely macroscopic: to be the object of the purely a priori discovery that some macroscopic property plays a certain macroscopic role, both the conception of the role and the conception of its occupant would have to be macroscopic, because the natural determination of the macroscopic by the microscopic cannot be the object of a priori discovery.